EQUITY, HEALTH AND HEALTH CARE:
AN ECONOMIC STUDY WITH REFERENCE TO PORTUGAL

Thesis submitted for the degree of Doctor of Philosophy

by

João António Catita Garcia Pereira

University of York
Department of Economics and Related Studies

January 1995

For my parents and sisters,
António, Aira, Fátinha and Zaíta

Abstract

This study investigates the extent to which equity in the distribution of health and health care has been achieved in Portugal. It draws on economic methodology from various sub-areas of the discipline, namely health economics, inequality measurement and economic philosophy. The study is divided into five main parts. The first provides background information necessary to an understanding of the health-equity problem in Portugal. The second part addresses the issue of normative specification of equity objectives. Part 3 presents methodology designed to measure inequity in the domain of health, focusing specifically on indices derived from the concentration curve approach. Part 4 is devoted to empirical analysis of income-related inequity in Portuguese health and health care. Finally, Part 5 provides further conceptual analysis with a view to future empirical research.

The study's core contribution is in developing the concentration curve approach to measurement of health domain inequity. The normative and statistical properties of indices used in previous work are clarified and a number of indices that are new to health economic research are developed. The new measures include generalized concentration indices that allow for a representation of alternative social judgements on equality preference, and an index of horizontal inequity in the delivery of care based on the correspondence of rank positions in the utilization and morbidity distributions. The study also presents a rationalization of a family of indices that measure the level of social welfare associated with the health distribution. These indices permit a qualified trade-off between distributional and aggregative goals, and may be seen as a step toward integrating equity and efficiency measurement. At the empirical level, the thesis provides a wider range of applications than is currently available; and examines extensively the impact of methodological choices on the degree of measured inequity, an area where past work has been largely silent.

Other themes developed in the thesis are the application of Sen's capabilities framework and Grossman's demand for health model to the study of health domain inequity.

The empirical analysis draws on three separate portuguese informational bases: a large scale health interview survey, routine mortality statistics and two family budget surveys. The results show, *inter alia*, that the distribution of ill-health is generally unfavourable to poorer income groups; that the degree of inequity in infant mortality is currently much lower than it was in the 1970's; and that, throughout the 1980's, the distribution of family health care payments evolved from being overall progressive to overall regressive.

Contents

Acknowledgements

I am deeply indebted to my supervisor, Alan Maynard, for his expert academic guidance, patience and encouragement. These qualities, which I have drawn on generously, were decisive in the preparation of the thesis. On a wider plane, I wish to acknowledge the important role of his teaching in improving my understanding of health economic issues.

I would also like to thank the other members of my advisory panel, Tony Culyer and John P. Hutton. Their special talents of clarification, of extracting the essential and their ability to teach and criticize in a constructive and motivating manner have contributed immensely to the research now presented.

Other individuals, both within York University and elsewhere, have proven indispensable in giving me advice and assistance. At York, I was able to call upon Peter Lambert's unparalleled knowledge of conceptual issues related to inequality indices. Though he was not a member of my advisory panel, he always gave freely of his time, providing detailed comments and suggestions. I also received helpful comments from Andrew Jones, Alan Williams and Owen O'Donnell; and from former students Rosella Levaggi and Muradali Ibrahimo.

Outside York, I owe a special debt of gratitude to co-researchers on an EU project on 'Equity in the Finance and Delivery of Health Care'. In particular, I would like to acknowledge the contribution of the project leaders, Eddy van Doorslaer, Adam Wagstaff and Frans Rutten. Even a cursory glance at the thesis will show that it owes a great deal to the seminal work which they have carried out.

I would also like to thank Julian Le Grand, Gavin Mooney and Juan Rovira, who made stimulating comments on the earlier parts of the thesis.

In Portugal, I thank my colleagues - Maria do Rosário Giraldes, Francisco Ramos and Carlos Costa - at the Department of Health Economics of the National School of Public Health, New University of Lisbon, for their encouragement and for supporting a heavier workload during my absences. I am particularly indebted to the former head, António Correia de Campos, who provided the initial stimulus for the research and made many useful comments on draft chapters.

I had many helpful discussions with Carlos Gouveia Pinto which assisted and elucidated my work. Carlos Farinha Rodrigues helped me in the preparation of the Family Budget Survey data bases and provided useful insights into their operation. Fernando Galvão de Melo and Pedro Pita Barros made valuable comments which allowed me to improve the conceptual analysis. My thanks go to them.

Parts of the research were presented at the following scientific meetings: the VII

and XI Jornadas de Economia de la Salud (Gran Canaria and Alicante); 1st Workshop on Econometrics and Quantitative Methods in Health Care (York); 2nd European Congress of Health Economics (Paris); and IV Encontro da APES (Coimbra). I wish to record my thanks to the organizers and participants of these meetings for their stimulating comments. A similar acknowledgement is also due in relation to the two seminars I gave at the Department of Economics in York.

Naturally, all remaining errors and misconceptions are my own.

Financial support for the research came from a number of sources. The initial funding was provided by the British Council. Subsequently, I received short term awards from the Fundação Luso-Americana para o Desenvolvimento, JNICT-Invotan and the Fundação Gulbenkian. The investigation reported in Chapter 7 was funded by the Junta Nacional de Investigação Científica e Tecnológica (Project no. PCSH/C/ECO/225/91).

Other institutional support was provided by the National School of Public Health in Lisbon, which generously granted me study leave in order to pursue the thesis. I wish to record my gratitude to all who helped and encouraged my work, in particular the two Directors during the period of research, António Lobato Faria and José Caldeira da Silva.

I also wish to thank the Department of Health Studies and Planning (DEPS) of the Ministry of Health and the National Statistical Institute (INE) for granting me access to the National Health Survey and Family Budget Surveys data bases; and the Portuguese Health Economics Association (APES) for facilitating access to bibliographical sources.

Throughout my stay in York I received the warmest attention and support from all departments. A special mention is however due to all who work at the Centre for Health Economics for their friendliness and efficiency.

Finally, I wish to thank Margarida Moreira for her constant support and encouragement. I fear that the many hours she gave up to help with data input will have left her with an impression that economics really is a "dismal science". I know, however, that if any good comes of the study, hers will be the greatest happiness.

Declaration

Part of the material presented in the thesis has previously been published in the course of the period of registration.

Chapter 2, section 4, as a sub-section of 'Chapter 11: Portugal', In: E. Van Doorslaer, A. Wagstaff and F. Rutten, *eds.*, Equity in the Finance and Delivery of Health Care: An International Perspective. Oxford University Press, Oxford, 1993 (in co-authorship with C. Gouveia Pinto). The relevant part of the chapter may be attributed to both authors in equal parts.

Chapter 3, sections 2 to 5, as 'What does equity in health mean?' Journal of Social Policy, Vol. 22, no. 1, 1993.

Chapter 8, section 3, as 'On the measurement of social welfare associated with health' In: E. Levy and A. Mizhari, *eds.*, Colloque Européene: De l'Analyse Economique aux Politiques de Santé. CREDES, Paris, 1993.

Chapter 1

Introduction

1.1 Purpose, scope and limitations

This study has two main objectives. First, it aims to *investigate if, and to what extent, there exists inequity in the distribution of portuguese health and health care.* As in other countries, policy makers in Portugal have repeatedly shown concern that fairness obtains in the field of health. In 1979, a National Health Service was created with the expressed aim of guaranteeing access to health care for all citizens, independently of their social and economic status (Assembleia da República, 1979). Subsequent governments have maintained this and related commitments. For example, in 1990 a Law setting out basic principles of health policy considered "equity" to be a "fundamental objective" (Portugal, 1990); and in 1986, the Secretary of State for Health, whose party has held power over the last decade, declared that "it will not be as a result of the Government's inaction that a 'Portuguese Black Report' is published in the year 2000" (Baptista Pereira, 1986). Despite these manifestations, there have been relatively few studies that measure the extent to which health-equity objectives are being achieved.[1] Furthermore, the research that exists is limited in a number of respects. Empirical investigation is seldom based on sound theoretical analysis, making it difficult to extract policy implications; measurement techniques are often crude and uninformative; and the range of questions that are addressed appear to be simply guided by ready availability of empirical elements.

The present study aims to redress the lack of suitable information about health domain inequity in Portugal, by developing and applying a research strategy based on

[1] I use the term "health-equity" as shorthand for "equity in the domain of health", thus englobing objectives that pertain to the distribution of either health or health care.

economic methodology. It begins by tracing out a normative framework that clarifies concerns expressed in portuguese policy statements. From this framework an agenda of positive analysis is developed. Actual empirical application is circumscribed to certain key aspects of the agenda, namely the economic distributions of morbidity, infant mortality and family payments towards the health service. These analyses draw on three distinct informational bases: a large scale health interview survey, routine mortality statistics and two family budget surveys. They provide evidence previously unavailable to portuguese policy makers and also some elements that may be useful in other contexts (*eg.* a before and after comparison of the degree of inequity, following changes in the structure of health care financing of the type being considered by other countries). Those aspects identified in the normative framework as requiring positive analysis, but for which empirical measurement is not carried out, are not left completely unattended. The final part of the thesis specifies conceptual frameworks that permit a further two issues to be addressed in future research: namely, the level of social welfare associated with the health distribution and the degree of horizontal and vertical inequity in the delivery of care.

The second, but no less important, objective of the thesis is *to contribute to economic analysis of inequity in the domain of health.* A number of developments are made, which complement previous work in the health economics, income inequality and economic philosophy literatures. The core contribution is the application of the concentration curve approach to measurement of health-inequity. The key features of this approach are: (i) a focus on health and health care inequalities related to economic status, which appears to be the equity issue which most preoccupies observers, and (ii) the construction of index numbers that quantify the *degree* of inequity and facilitate comparisons of the large amounts of information that are often employed in distributional studies in the field of health. The methodology was introduced to health economic research by Wagstaff, Van Doorslaer and Paci (1989) and later refined in, *inter alia*, Wagstaff *et al* (1991c) and Van Doorslaer *et al* (1993). In reviewing this work, Henry Aaron (1992) argued that "as with any pioneering effort, observers can find issues that are left hanging or that are subject to challenge. Indeed, the essence of seminal contributions is that they spawn additional research to correct, modify and extend the initial effort" (p. 467). The present thesis takes up this challenge by

providing both conceptual and empirical extensions.

At the conceptual level, the study clarifies the normative and statistical properties of indices used in previous work. It also presents a number of indices that are new to health economic research. This is the case of the so-called generalized concentration measures, which incorporate alternative social judgements on the degree of equality preference; and of the new index of horizontal inequity in the delivery of care, based on the correspondence of rank positions in the utilization and health status distributions. At the empirical level, the thesis provides a wider range of applications than has hitherto been the case (*eg.* the measurement of inequity in infant mortality making use of geographical and economic data, and a time comparison of inequity in health care financing which avoids problems of data comparibility that pervade cross-national studies). The study also computes generalized concentration measures, which as far as I am aware have not been the subject of empirical application elsewhere in the literature. Perhaps the most important contribution is, however, the examination of the impact of methodological choices on the degree of measured inequity. Given the notoriety of measurement problems in health-inequity analysis, it is surprising that previous research has been largely silent on this issue. The present study carries out sensitivity analyses on a number of controversial choices (*viz.* family equivalence scales, ability to pay proxies, aggregation of inequality at different points of the distribution, etc.). Such analyses improve the robustness of conclusions that are drawn and provide useful information to future empirical work, whether in Portugal or abroad.

Other contributions to the application of economic methods in the study of health-inequity are also made. First, a normative framework based on Amartya Sen's (1982) concept of capabilities is proposed as a means of organizing empirical research and policy formulation. This proposal comes in the sequence of various other economic contributions aimed at disentangling the complex definitional issues raised in health-equity analysis [*eg.* Mooney and McGuire (1987), Le Grand (1987, 1991a), Culyer (1990, 1993) and Culyer and Wagstaff (1993)].[2] These studies have contributed to the

---

[2] As Gavin Mooney argued in an early article, without more explicit consideration of normative issues the debate on equity in health and health care would remain "confused and confusing" [Mooney (1983, p. 179)].

emergence of what is arguably a new paradigm in economics: the 'extra-welfarist' perspective (Culyer, 1989). The development of Sen's notion to the health sphere can be seen as a further addition to this line of enquiry.

Second, the thesis provides a rationalization of a family of indices that measure social welfare in the domain of health. An important development in recent research has been the attempt to integrate equity and efficiency measurement in health policy evaluation [*eg.* Culyer (1989); Wagstaff (1991)]. This work has concentrated on probing conceptual issues. The research reported in this study aims to provide a bridge to empirical application. Drawing on the concentration curve approach and little known work by Kakwani (1986) and others, indices are developed that allow for a qualified trade-off between equity and health maximization objectives.

Finally, in an appendix, the study proposes a modest extension to existing demand for health models. Building on the work of Grossman (1972), Muurinen (1982a,b) and Williams (1988a), a theoretical framework is presented which aims to adapt those models to better examine the relationship between economic and social inequality and inequality in the distribution of health. The distinguishing feature of the model is the introduction of an independent stock of education capital which permits a clearer understanding of the relationship between wealth and knowledge and of their impact on health and health care. This research seemingly shifts the ambit of the analysis from measurement of the extent of inequity to explanation of how particular distributions come about. However, it is an integral part of positive analysis suggested by the normative framework, which places emphasis on inequity in the transmission process from goods (such as health care) to health outcomes.

Inevitably, the thesis has a number of limitations. One of these is that empirical analysis is not carried out for some issues identified by the normative framework. It should not be presumed that these questions (examined conceptually in the latter part of the thesis) are less important than those for which empirical measurement is provided. Their exclusion reflects the delicate trade-off between exiguity of space and the need to examine particular questions in sufficient detail. Given that a choice has to be made, empirical investigation is aimed at those issues that are currently the subject of debate

in Portugal but for which available information is sparse.

Another limitation is that the study concentrates almost exclusively on violations of fairness that are related to economic status. There are important reasons for this, provided by both the normative framework and portuguese policy statements. However, there are other potential sources of inequity, such as those related to place of residence, age, race or gender. The examination of these might have implications for the policy conclusions to be drawn. Thirdly, despite an attempt to introduce tools that simultaneously measure the attainment of distributional and maximizational goals, the question of how health-equity objectives might be traded-off against other desiderata is left largely unresolved. Finally, the thesis has limitations in terms of the data it is able to call upon. Some issues which appear conceptually straightforward, become hazed once attention is turned to data analysis. Therefore, a significant amount of space is given over to identifying data deficiencies and in signalling due caution in the interpretation of results.

1.2 Outline of the study

The thesis is divided into five main parts. Part 1 provides background information necessary to an understanding of the health-equity problem in Portugal. The second part addresses the issue of normative specification of equity objectives. Part 3 presents methodology designed to measure inequity in the domain of health, focusing specifically on indices derived from the concentration curve approach. Part 4 is devoted to empirical analysis of income-related inequity in Portuguese health and health care. Finally, Part 5 provides further conceptual analysis with a view to future empirical research.

Given that the thesis focuses on the problem of health domain inequity in Portugal, it begins by presenting relevant background information about that context. The first three sections of Chapter 2 - where the information is presented - draw attention to distinguishing features of portuguese economy and society, the health of its population and the finance and delivery of health care. The description does not aim to be

comprehensive, concentrating instead on key issues that characterize the portuguese situation when compared to, for example, the european context. The final two sections examine the various health-equity objectives imprinted in legislative and policy statements and the manner by which portuguese authors have previously addressed the problem of inequity in health and health care. Generally, the chapter serves to identify key questions that are addressed in later conceptual and empirical analysis.

Chapter 3 deals with the specification of equity objectives in health policy. Prior to empirical analysis one needs to establish which attributes ought to be measured. This entails asking the question "what does equity in the domain of health signify?" In the present context, the answer ought properly to be given by portuguese policy statements. However, as in other countries, these are insufficiently precise (and sometimes confusing) to allow positive analysis to be undertaken. Consequently, the chapter begins by critically appraising a number of distribution principles that might suitably be applied to the sphere of health and shed light on concerns expressed by portuguese policy makers. Six well-established approaches to defining equity - egality, entitlement, the decent minimum, utilitarianism, Rawlsian maximin and envy-free allocations - are discussed. Each is found wanting in terms of previously laid out assessment criteria. More recent formulations suggested by health economists are also reviewed. Though these provide useful insights, they too are shown to fall short of the basic requirements that a definition of health-equity ought to meet. An alternative conception based on Sen's notion of equality of capabilities is then put forward. After examining its application to health space, a simple diagrammatic framework is used to highlight questions that ought to be the correct focus of empirical analysis. Two main issues are identified: the actual level of the functioning 'good health' that particular individuals obtain, described as their *achievement*; and the experience of individuals in the space of commodities that are instrumental to the attainment of the desired functioning, their *advantage*. The chapter then proposes the manner by which these attributes are to be measured, adopting an indicative pluralistic approach that seems in tune with the multiplicity of health-equity objectives implied by portuguese legislation.

Having discussed the normative basis of the analysis, the study moves on to the establishment of tools for measurement. This task is divided between Parts 3 and 5,

with the distinction that those presented in the former are the subject of extensive empirical application in Part 4, whereas the latter are examined only at the conceptual level. Chapter 4, which makes up Part 3 of the study presents various indices which permit measurement of the two main issues of analysis: inequity in health and inequity in health care payments. Taking income-based indices of inequality as the point of departure a choice is made to focus on concentration curve based measures, namely the illness concentration index and the Suits and Kakwani progressivity indices. Generalized versions of these measures, allowing for the parameterization of alternative social judgements concerning the degree of equality preference, are also presented. Finally, the chapter discusses methods of estimation and provides empirical information on the statistical bias inherent in linear approximation methods.

Moving to Part 4 of the study, Chapter 5 provides estimates of inequity in morbidity based on data from the 1987 National Health Survey. This is the only nationally representative data source which allows data on illness to be systematically linked to the demographic and socio-economic characteristics of individuals. Strong evidence is found showing that the distribution of ill-health is generally unfavourable to poorer income groups. There are, however, some exceptions to the rule (*viz.* certain types of morbidity and particular age groups). The chapter also establishes the relationship between equivalence adjustments to the income variable and the degree of measured inequity. It is shown that this issue cannot be overlooked in future research.

Chapter 6 measures the time-trend of inequity in infant mortality. Given qualitative limitations in the socio-economic data available on birth and death certificates, the analysis relies on geographical observations ranked by an index of economic position. The results suggest that the degree of inequity is currently much lower than it was in the 1970's. However, there are some noticeable counteracting trends, such as in the distribution of post-neonatal deaths where the degree of inequity appears to be increasing.

Chapter 7 measures inequity in the distribution of health care payments among families with differential ability to pay. The analysis makes use of two Family Budget Surveys carried out in 1980/81 and 1989/90. The main result is that, throughout the

1980's, the distribution of health care payments evolved from being overall progressive to overall regressive. This change is shown to be robust in terms of a wide range of methodological choices. This part of the analysis is particularly relevant to research in other countries since it explores questions over which there is uncertainty with regard to empirical specification.

Chapters 8 and 9 explore measurement tools for examining other health-equity issues in future research. Chapter 8 tackles the question of social welfare associated with the health distribution. A mathematical rationale is provided for a family of indices that take due account of the size of health output *and* the way this output is distributed among different income groups. The measures permit a qualified trade-off between distributional and maximizational goals on the basis of variation in a single parameter. Furthermore, the analysis shows that health concentration indices are ethically-admissible measures in the context of social welfare. The chapter also includes empirical illustrations drawing on the health survey data used in Chapter 5.

Chapter 9 discusses indices of concentration aimed at measuring vertical and horizontal inequity in the delivery of health care. It builds on the work of Wagstaff and van Doorslaer, reinterpreting their analysis as measurement of the vertical equity component in the goal of health care utilization according to need. The claim that their method examines horizontal inequity can only be accepted under restrictive data-related assumptions that are unlikely to command widespread agreement. Accordingly, a more general index of horizontal inequity is developed, based on the correspondence of rank positions in the utilization and morbidity distributions.

Chapter 10 provides an overall conclusion to the study. A summary of the main findings is presented, with attention being drawn to the research and policy implications of the results. The chapter also identifies areas for future investigation.

Finally, a number of Appendices are included that present complementary material to the research reported in the main body of the thesis. Appendix 1, however, is of a different character since it provides a self-contained analysis of the question of modelling health related behaviour. This work is an implicit requirement of the

underlying ethical theory chosen for the study, which suggests that measurement of health domain inequity must also consider the *process* whereby goods, such as health care, are transformed into health functioning. However, it uses techniques that are somewhat different to those employed throughout the thesis and is quite evidently of an exploratory nature. Unlike Chapters 8 and 9 where the route to empirical application is straightforward, further specification remains to be accomplished in future research. The appendix begins by appraising Grossman's health investment model as an aid to health-equity research, and then presents an alternative specification that aims to remedy the identified weaknesses. The new model is solved by means of optimal control theory and has the potential for allowing a clearer understanding of the relationship between income and education and of their impact on health and health care.

Part 1

BACKGROUND

Chapter 2

Health, Health Care and Equity in Portugal

"It will not be as a result of the Government's inaction that a Portuguese 'Black Report' is published in the year 2000."
A. Baptista Pereira, Secretary of State for Health (1986)

2.1 Introduction

This chapter presents background information relevant to the study of health domain inequity in the particular context of Portugal. Besides providing the reader with basic descriptive elements that are little-known outside the country's borders, the chapter also helps to identify specific research questions that need to be addressed and allows later empirical results to be usefully interpreted.

The organization is as follows. Section 2.2 describes the main contours of Portuguese demography, economy and society. Data are presented in a form which permits ready comparison with other European countries. In section 2.3, statistics on population health (mainly mortality) are examined. It is shown that despite important changes in recent decades, Portugal continues to reveal health patterns that are distinct from those of its european partners. Section 2.4 describes basic aspects of the finance and delivery of health care. Special emphasis is given to features which distinguish the health care system from those in comparable countries, rather than providing a detailed examination of administrative structures, coverage, resources and utilization patterns. It is argued that the claim that portuguese health care conforms to the NHS model is in some respects misleading. Particular attention is then paid, in section 2.5, to the legal and political specification of health-equity objectives in an effort to extract the precise yardsticks by which policy should be monitored. Section 2.6 examines how Portuguese

researchers have tackled the measurement of inequity in health and health care and reveals the type of information presently available. Finally, section 2.7 brings together the various strands of information invoked in the chapter.

2.2 Demography, economy and society

Portugal is a small country on the periphery of Europe with a population not much larger than London's. Throughout the post-war period it has lagged behind Western Europe in terms of economic and social development and is generally acknowledged to be the second poorest member of the European Union (formerly EC), which it joined in 1986.

Table 2.1 reports various demographic, social and economic indicators for Portugal and contrasts them with corresponding figures for the UK and the twelve EU countries as a whole. Since 1970, Portugal's population has grown by 18 per cent, to over 10 million. With birth and death rates not much different from those of other EU nations, it seems that natural demographic phenomena explain only a part of the increase. Far more important are the influx, between 1974 and 1976, of over 600 000 residents from Portugal's former colonies and the regular stream of returning emigrants from N.W. Europe. A further characteristic is that, despite recent population growth being circumscribed to the western-most urbanized regions of the country, Portugal's residents continue to live predominantly in rural settings: indeed more so than in any other country within the World Health Organization's European Region, including Ireland, Greece and Turkey (WHO, 1985a).

Portugal's economy registers a relatively poor performance in the European context and is reckoned to be "structurally deficient" (OECD, 1984). Purchasing power parity comparisons of GDP per head show the figure for Portugal to be 51 and 54 per cent of the UK and EU averages, respectively. Added to this, the cost structure of GDP reveals strong characteristics of a less-developed economy, namely low proportions attributed to wages and salaries and to investment. The rate of inflation too has systematically overshot trends in richer countries.

Table 2.1: Background data on Portugal, United Kingdom and the EU

1990 or nearest available year

Indicator		Unit	Portugal	U.K.	E.U.
1. Population					
Total population		1000	10 337	57 409	327 137
Population density		Per Sq Km	112.1	234.5	144.7
Population aged 65 +	M	Per cent	5.3	6.2	6.0
	F		7.7	9.4	8.7
Birth rate		Per 1 000	11.7	13.8	11.0
Death rate		Per 1 000	9.5	11.4	11.2
2. Economy and Finance					
GDP per capita		Current PPP's	9 452	18 402	17 229
Cost structure of GDP		Per cent			
- Compensation of employees			44.1	55.6	50.7
- Taxes minus subsidies			13.6	14.2	11.1
- Consumption fixed capital			4.3	11.0	11.6
- Net operating surplus			37.9	19.0	26.7
Consumer price index		1985 = 100	170.9	133.3	123.0
3. Employment & Living Standards					
Employment by sector		Per cent			
- Agriculture			18.9	2.2	7.0
- Industry			35.3	29.5	32.5
- Services			45.7	68.4	60.6
Average length working week		Hrs. per week	41.3	37.3	37.6
Unemployment		% labour force	8.8	11.4	10.8
Average wages of manual workers in industry		Gross hourly, current PPS	6.8	14.8	-
Current expenditure on social protection		Per cent of GDP at market prices	17.0	22.1	25.3
Private motor vehicles in use		Per 1000 popul.	227	366	394
Number of students in full-time education		Per cent of 5-24 age group	62	62	72

Source: EUROSTAT, Basic Statistics of the Community, 28th Ed., Office Official Publications of European Communities, 1991.

In the years since joining the EU, Portugal's economic performance has improved: inflation and unemployment have decreased and there has been a period of sustained growth. This improvement, however, "owes much to factors that may not be repeated and policies that cannot be sustained" (OECD, 1988). Structural problems in the economy remain, many of which are likely to affect overall production of health. Nineteen per cent of the work force continues to be engaged in agriculture, largely in the North of the country where land is fragmented into small holdings of scattered plots from which proprietors eke out a living at subsistence level. Average wages in industry are only 46 per cent of those in the UK when expressed in purchasing power standards. There are wide disparities in income between regions and population groups, while poverty is more pronounced than in other European countries (O'Higgins and Jenkins, 1989).

A picture of the distinct make-up of Portuguese society can be gathered by applying the General Household Survey classification of socio-economic grouping (SEG) to data drawn from the Portuguese National Health Survey (see Table 2.2).[1] The percentage of persons in each SEG is radically different between the two countries. The distribution in Portugal is heavily weighted towards manual work with relatively low proportions in the higher grades that tend to find employment in the service sector. The strength of SEG IV reflects higher percentages of subsistence farmers and small-shopkeepers rather than skilled industrial workers.

Official statistics may, however, overestimate the degree of inequality. There is a flourishing underground economy and double employment is pervasive particularly among families in the rural communities of western-most districts, who complement industrial wages with subsistence agriculture or small-scale entrepreneurship. In themselves, these factors have an important bearing on family health production and serve to counteract low levels of collective social provision, which in expenditure terms amounts to only 17 per cent of GDP compared with an average of 25 per cent in the European Union.

---

[1] The National Health Survey is a cross-sectional household interview survey covering the non-institutionalized civilian population. It is described in detail in Chapter 5, where it is used to examine the extent of income related inequity in morbidity.

Table 2.2: Socio-economic group of head of household in Portugal and the U.K., 1987 (Percent)

Socio-economic Group	Portugal		U.K.	
	Male	Female	Male	Female
Professionals, employers and managers (I & II)	11	5	30	9
Intermediate and junior non-manual (III)	14	13	15	38
Skilled manual & own-account non-professional (IV)	49	37	37	9
Semi-skilled and unskilled manual (IV & V)	26	45	18	37
Never worked	na.	na.	0	6
Sample size	11173	2560	7610	2525

Sources: Computed from: National Health Survey, 1987, MS-DEPS, Lisboa.
OPCS, General Household Survey 1987, HMSO, London, 1989

The general impression from the figures reported here is that despite being a country fully integrated in the european economic system, Portugal remains very much a peripheral society. The population has distinctive features, close no doubt to much of Spain's, Southern Italy's, Ireland's or that of Greece, but radically different to the rest of Western Europe; patterns of employment and production are outmoded, inefficient and sometimes ambiguous; and living standards, in general, are poor and reveal extreme inequalities. These important factors must be borne in mind in an analysis of the equity attributes of the health care system, since they inevitably influence the nation's production of health.

2.3 Health of the population

In recent decades, portuguese mortality patterns have undergone profound changes. Though in many respects they now approximate the experience of the more developed parts of Europe, there are still signs of late socio-economic development. Evidence of

Table 2.3: Indicators of length of life in Portugal and the UK

	1970		1980		1990	
	UK	Por	UK	Por	UK	Por
Infant Mortality						
Infant	18.5	55.1	12.1	24.3	8.4	11.0
Neonatal	-	25.4	7.7	15.4	4.4	7.0
Post-neonatal	-	32.6	4.4	8.9	4.0	4.0
Perinatal	23.8	37.0	13.9	23.9	8.3	12.6
Life Expectancy at Birth						
Men	68.6	64.1	70.4	67.5	72.7	71.1
Women	74.9	70.3	76.5	74.6	78.2	78.2
Potential Life Years Lost						
All causes	7922	16028	6309	10746	4929	7755
Tuberculosis	34	446	11	110	5	48
Lung cancer	408	102	313	156	220	165
Ischaemic heart disease	1315	422	1279	463	854	359
Road accidents	603	1083	531	1317	420	1154

Notes: 'Potential Life Years Lost' are defined as the years lost due to premature death before age 65 (OECD, 1993). The figures shown refer to males and are expressed per 100 000 population.

Sources:
- Ministério da Saúde, Portugal - Saúde - 1990, DEPS, Lisboa, 1992.
- OECD/CREDES, OECD Health Data. A software package for the international comparison of health care systems. OECD/CREDES, Paris, 1991.

this can be seen in Table 2.3 which compares, for Portugal and the UK, the post-1970 evolution of infant mortality rates, life expectancy at birth and potential life years lost. In Portugal, this period has witnessed the most impressive gains in lives previously lost prematurely. Take, for instance, the case of infant mortality. In 1970, one in eighteen portuguese children died in their first year of life. In the ten years to 1980 this figure was reduced by 56 per cent and then by a further 55 per cent to 1990. Despite this impressive decline, the infant mortality rate remains 31 per cent higher than that of the UK, which is close to the EU average. Some commentators have also drawn attention to the existence of extreme regional inequalities. Northern districts such as Bragança and Vila Real show rates twice as great as those in the south where they are generally

below 9 per thousand.[2] Disparities are even wider for post-neonatal mortality, the component of infant deaths most often associated with socio-economic environment. In this case almost three times as many children die in the two northern districts (Carrilho, 1985).[3]

Table 2.4 provides a comparison of 'potential life years lost' (PLYL) for males in European Union countries. The aggregate figures shown in column 9 indicate that rates of premature death are highest in Portugal. There is a 23 per cent excess with regard to the second placed country (Luxembourg) and 87 per cent *vis-à-vis* the country with the lowest rate (Netherlands). Disaggregation by selected causes of death shows a distinct pattern for Portugal in that it tends to be placed at the extremes of variance. The number of male PLYL due to tuberculosis, which the previous table showed to have declined rapidly in Portugal in the last two decades, are still much higher than in other countries (*eg.* more than six times greater than France, Greece or Italy). Premature deaths due to cancer tend to be comparatively low. Lung cancer PLYL, for example, is at the bottom of the EU league and even then only 73 per cent of the next lowest country (Ireland). However, as the figures in the previous table suggest, the situation in Portugal has been steadily worsening over the past two decades. This trend is likely to continue given that consumption of tobacco is still increasing and because of the long interval between exposure to tobacco and development of cancer.

Premature deaths attributable to diseases of the circulatory system show a mid-ranking position in the EU table, but their composition is striking. PLYL due to ischaemic heart disease are the second lowest in the Union; but PLYL due to cerebro-vascular disease are the highest (65 percent higher than Spain, the second placed country). Though the asymmetries may in part be explained by variation in diagnostic practice between countries, they also reflect differences in diet and exposure to harmful substances such as tobacco.

---

[2] A map of Portugal, showing geographic divisions by regions and districts, is presented as Appendix 2. Reference is made to these delimitations throughout the thesis.

[3] Given that persisting inequities in infant mortality and its components have recently been the source of some concern in Portugal [see, *eg.*, Leitão (1987)], Chapter 6 is dedicated to analysing this problem. The objective is to examine whether the rapid decline in overall rates has been distributed fairly between rich and poor.

Table 2.4: Potential life years lost in EU countries, Males

Selected Causes of Death, 1988
Per 100,000 male population

	(1) Tuber-culosis	(2) All Cancers	(3) Lung Cancer	(4) Diseases Circulat. System	(5) Ischaemic Heart Disease	(6) Cerebro-Vascular Disease	(7) Liver Disease/ Cirrhosis	(8) Motor Vehicle Accidents	(9) All Causes
Belgium	10	1139	394	1020	527	159	149	904	5943
Denmark	5	1002	244	962	642	143	183	533	5157
France	10	1353	316	721	301	142	234	787	6008
Germany	10	1089	262	1005	531	141	252	552	5266
Greece	12	923	288	1044	554	212	66	753	5321
Ireland	11	956	217	1250	902	120	32	599	5183
Italy	9	1180	331	810	403	163	235	578	5196
Luxembourg	12	1253	369	1069	593	152	292	1348	6571
Netherlands	5	982	288	903	566	105	69	345	4322
Portugal	61	1033	160	1009	391	345	371	1193	8092
Spain	36	1070	240	1108	408	239	271	629	5758
U.K.	5	967	235	1247	925	150	61	409	5087

Notes: - The values for Belgium and Spain refer to the year 1985.
- Full designation of causes of death with respective ICD number:

(1) Tuberculosis (01)
(2) Malignant neoplasms (08-14)
(3) Mal. neoplasm of trachea, bronchus & lung (101)
(4) Diseases of the circulatory system (25-30)
(5) Ischaemic heart disease (27)
(6) Cerebro-vascular disease (29)
(7) Chronic liver disease and cirrhosis (347)
(8) Motor vehicle traffic accidents (E471)

Source: OECD/CREDES, OECD Health Data. A software package for the international comparison of health care systems. OECD/CREDES, Paris, 1991.

Two other specific causes of death have caused alarm in recent years. Firstly, mortality due to liver cirrhosis which, in common with other wine producing countries, is extremely high. The figures for Portugal are once again the highest (more than four times greater than than the UK's, for example). Secondly, deaths due to motor vehicle accidents, where Portugal is only surpassed by Luxembourg in terms of male PLYL. It is also noticeable that, unlike the UK, the situation has worsened over the past 20 years (see Table 2.3).

Although mortality statistics are useful for comparative purposes they provide an incomplete picture of the health of a population. Indicators of morbidity for the country as a whole only became available recently through the National Health Survey for 1987. Figure 2.1 reports some information computed from the Survey which complements the data on mortality.[4] The prevalence of disability (days off work or school and restricted activity) is broken down by the main types of illness. The National Health Survey is useful for this purpose in that a rigorous attempt is made to classify illness by pathology, as well as its chronic, acute or symptomatic nature. In 1987 the principal reported pathology referred to illness of the muscoskeletal system with chronic conditions being the most important. The strength of this type of illness is not apparent if one considers mortality statistics as proxies for health status. Chronic circulatory and acute respiratory illness also accounted for important proportions of disability, while lesions and digestive system illness show up strongly as they do for mortality.

Despite being the only nationally representative data source with information on morbidity, the 1987 National Health Survey has scarcely been used in past research. One of the main contributions of the present thesis is to draw on the data base to examine income related inequity in the distribution of morbidity (Chapter 5). In so doing, new information is produced that helps to understand the health profile of the portuguese population.

---

[4] Due to differences in practice between health surveys in different countries it is not advisable to provide a cross-national comparison.

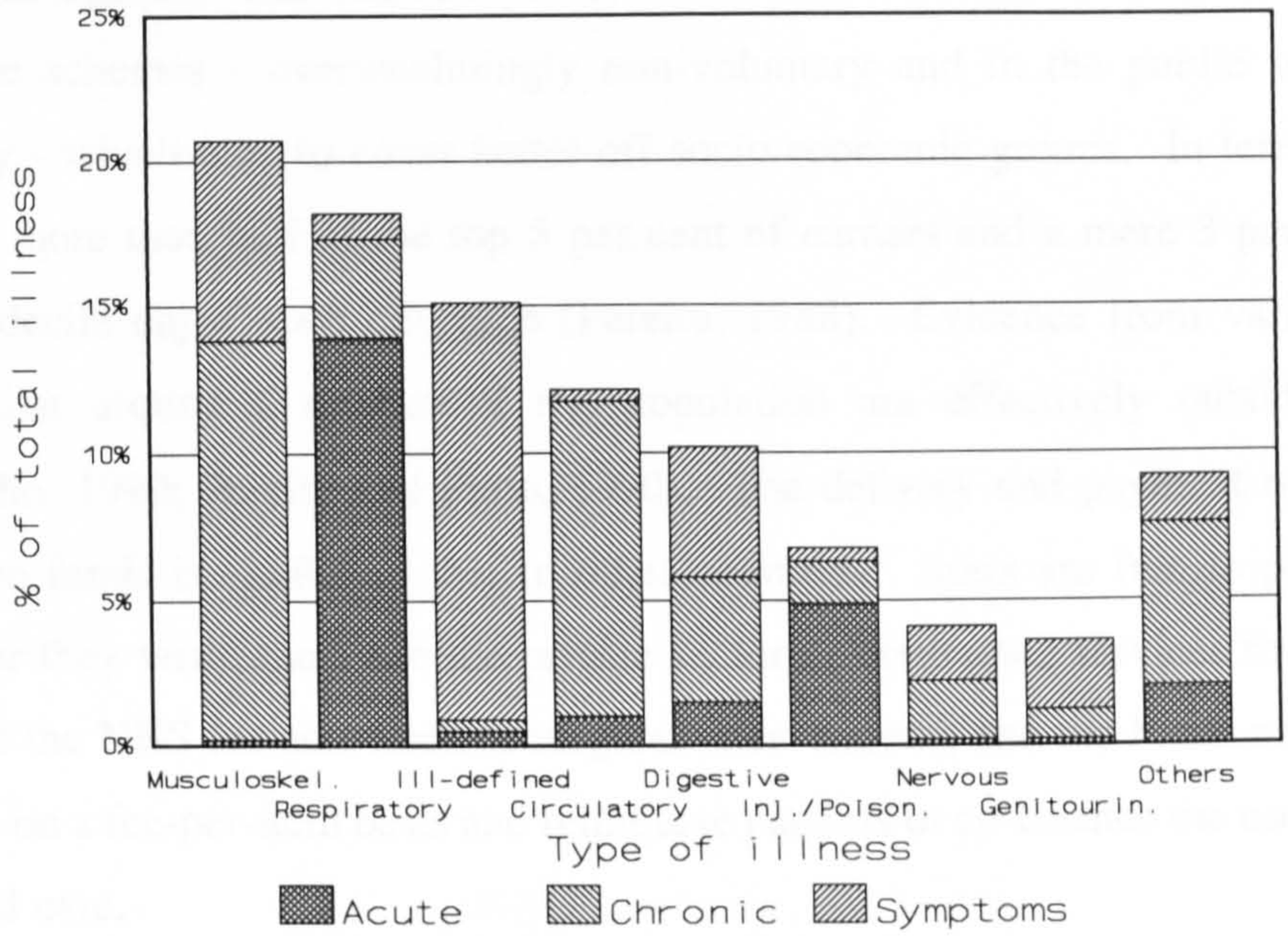


Figure 2.1: Morbidity indicators. Type of illness reported by persons with disability in a two-week reference period

(National Health Survey, 1987 - Own computation)

2.4 Finance and Delivery of Health Care

The portuguese health care system is often described as conforming to the classical National Health Service model (*eg.* WHO, 1981; Urbano *et al*, 1993). This model is characterized by universal coverage of the population, generality of benefits, national tax financing and national ownership or control of factors of production (OECD, 1987). In 1979, a National Health Service was indeed created with a political commitment that it become the preponderant mode of health care financing and provision. Yet the available evidence suggests that the system as a whole departs significantly from those of other countries usually accepted as conforming to the NHS model (*eg.* the UK, Denmark or Italy).[5]

---

[5] Of course, no specific country's health system is fully described by the NHS model. The point made in this section is that finance and delivery of health care in Portugal departs from the classical paradigm in more respects than other systems. The importance of this observation is the implications it holds for the nature of incentives to economic and social agents and, ultimately, for the distribution of health and health care.

Although the NHS claims to be *universal* there coexist a number of occupational insurance schemes - overwhelmingly non-voluntary and in the public sector of the economy - which tend to cover better-off socio-economic groups. In terms of family income, more than half of the top 5 per cent of earners and a mere 3 per cent of the bottom decile enjoy such coverage (Pereira, 1988). Evidence from various sources shows that around a quarter of the population are effectively outside the NHS (Freixinho, 1990; Pereira and Pinto, 1990). The delivery and payment of care in the insurance funds is similar to that in other countries: users are free to purchase care wherever they wish; most use the private sector or contracted services for ambulatory care and the NHS for non-elective surgical interventions; and the funds pay contracted services on a fee-per-item basis and reimburse patients or co-finance the use of privately provided care.

The insurance part of financing is also similar to that in other countries in that employees contribute a small proportion of their income, but with an important qualification. This is that, effectively, an important proportion of expenditures are part-financed by state taxation, due to the insurance funds operating overwhelmingly in the public sector of the economy. The major occupational scheme (ADSE) is destined for public servants and has the extraordinary implication of providing incentives for NHS workers not to use the NHS. A significant proportion of other insured individuals are covered by schemes run by public sector bodies or nationalized industries. Since employee contributions are generally insufficient to cover expenditures, the deficits are covered by taxation or foregone revenue, and effectively, by other sectors of the economy with greater proportions of lower paid workers. Private insurance, as such, is a negligible part of total health care financing. The precise distributional implications of such financing arrangements are not known.

One of the empirical exercises undertaken in this study is to measure how family payments to the NHS, insurance schemes and directly to providers are distributed according to ability to pay (Chapter 7). This analysis is important since there is an ongoing debate in Portugal over the future of health care finance, with both leading political parties apparently committed to enhancing the role of direct payments and private insurance (Mendo, 1993; Campos, 1990).

With regard to the NHS providing a *general* service of health care to patients the evidence is inconclusive. There exists a perennial under-utilization of equipment, either because of shortages in the supply of human resources or laxity in administrative controls on providers who work simultaneously for the NHS and the private sector. Physicians are plentiful by international comparisons but there are extreme shortages in some specialities (*eg.* dentistry and ophtalmology). Nurses are few, with scarcely one for each doctor, compared to a ratio of 6:1 in the UK (OECD, 1985). There is also a wealth of evidence showing an unequal spread of human and material resources throughout the territory (Campos, 1987).

From an equity perspective the fulfilment of the 'generality' objective requires that all types of care are readily available within the NHS. If there are extreme shortages of personnel or equipment in the provision of a particular mode of care and these imply prolonged and persistent queuing which leads people to seek treatment in the private sector then arguably the objective is not being met. There is a widely held belief that this situation is in fact the case, but admittedly very little evidence.

An indication that the NHS may not provide the sufficiently wide range of services it promises is indicated in Table 2.5. It shows that the NHS is predominant in the provision of hospital stays and GP and mother and child care but takes a minor role in specialist and dental consultations as well as diagnostic services, where it commonly reimburses private providers. Although such information raises more questions than it answers, it is fair to assert that private provision plays an important role in the delivery of health care in Portugal and that it does so where the NHS has willingly or unwillingly failed to carry out its intended general role.

The idea that the Portuguese health care system is *free at the point of use* and overwhelmingly financed by taxation is not totally borne out by the evidence. Table 2.6, shows the percentage of total and public health expenditure in GDP for the twelve EU countries. When public expenditure is expressed as a proportion of all expenditures, Portugal is shown to have the lowest share in the Union. In 1980, almost 30 per cent of all expenditures were out-of-pocket (rising to almost 40 per cent in 1990), figures which are significantly higher than those of other countries acknowledged to conform

Table 2.5: Health care utilization by sector in Portugal
Per cent. 1987

Type of care	NHS	Private
All consultations	67.0	33.0
GP consultations	76.5	23.5
Dental consultations	15.5	74.5
Specialist consultations	47.8	52.2
Family planning consultations	61.7	38.3
Ante-natal consultations	61.9	38.1
Child delivery	87.6	12.4
X-rays	47.5	52.4
Laboratory tests	29.5	70.5
Hospital stays	72.8	27.2

Notes: - All consultations refer to the last visit within a three month reference period, except for family planning consultations which are the last consultation with no reference period, and ante-natal visits, which refer to the last child born and currently under 5 years of age.
- Diagnostic utilization refers to the last use in a three month reference period.
- Hospital stays have a one year reference period and are taken from the 1985 National Health Survey, which was circumscribed to the Lisbon region.

Source: Ministério da Saúde, Inquérito Nacional de Saúde - 1987 (National Health Survey), DEPS, Lisboa. Own computation.

largely to the National Health Service model (*eg.* the UK or Italy). One would expect with this evidence, and other things being equal, that money prices of health care have a stronger rationing role in Portugal than in other countries where care is designated to be free at the point of consumption. It is true that a high share of out-of-pocket expenditures may simply reflect the strength of the insurance funds but there is evidence that NHS users also face significant money prices. On the one hand, the figures on which the proportion of public expenditure, shown in Table 2.6, is based include the public servants insurance fund (ADSE), thus leaving a small proportion of the population to account for a relatively large percentage of private expenditure. But more importantly, it is known that most individuals using the NHS face flat-rate co-payments

Table 2.6: Total and public health expenditure as a percentage of GDP in EU countries

	1980			1990		
	Total	Public	Public as % of total	Total	Public	Public as % of total
Belgium	6.6	5.4	*81.8*	7.4	6.1	*82.4*
Denmark	6.8	5.8	*85.3*	6.2	5.2	*83.9*
France	7.6	6.2	*81.6*	8.9	6.6	*74.2*
Germany	7.9	6.2	*78.5*	8.1	5.9	*72.8*
Greece	4.3	3.5	*81.4*	5.3	4.0	*75.5*
Ireland	8.5	7.8	*91.8*	7.1	5.8	*81.7*
Italy	6.8	5.6	*82.4*	7.6	5.9	*77.6*
Luxembourg	6.8	6.3	*92.6*	7.2	6.5	*90.3*
Netherlands	8.2	6.5	*79.3*	8.1	5.9	*72.8*
Portugal	5.9	4.2	*71.2*	6.7	4.1	*61.2*
Spain	5.9	4.4	*74.6*	6.6	5.2	*78.8*
United Kingdom	5.8	5.2	*89.7*	6.1	5.2	*85.2*

Source: OECD/CREDES, OECD Health Data. A software package for the international comparison of health care systems. OECD/CREDES, Paris, 1991.

for consultations and diagnostic tests and pay a large and rising proportion of the cost of drugs.[6] The latter payment varies with the therapeutic value of the drug in question with exemptions operating only in relation to the product (*ie.* if it is perceived to be life-saving) rather than patient characteristics (*eg.* age or income).

The final piece of the puzzle in explaining patterns of finance and delivery of care is provided by information on the ownership and control of the factors of production. With regard to human resources the NHS has guaranteed either by design (in the case of doctors) or shortfalls in supply (nurses) the full employment of the principal providers of health care. These same professionals, however, are not required to exercise their

[6] In 1980 NHS consumers paid on average 29 per cent of a drugs cost, a figure which had risen to 40 per cent by 1985 (Pinto, 1988).

duties on a full-time basis and tend overwhelmingly to work for the NHS in the morning, as salaried civil servants, and in private practice in the afternoon, on a fee-per-item of service or contractual basis. Autonomous market or NHS provision is negligible. Individuals who consult doctors in a private setting, either because they have insurance coverage or are willing to pay the rates set by the market, will be seen by the same practitioners who in the morning provided the same types of care in a public institution. The incentives generated by these circumstances go some way to explaining the utilization and expenditure patterns previously described. Due to laxity in regulation, doctors are motivated to supply minimum standards of care in NHS work-settings in order to augment the potential market share of private practice.

The NHS owns a sizeable majority of physical resources involved in the delivery of care, though as we have seen, provision in a private setting is far from negligible. Eighty eight per cent of hospital beds are in the public sector and there is a comprehensive network of integrated health centres and extensions in primary care. The NHS legislation decreed that private practice should complement public provision, in the sense of operating in areas where the latter was deficient, but all available evidence points to the contrary. In the hospital sector, for example, private provision is heavily concentrated in those regions where NHS supply is more extensive, while a comparative analysis of case-mix shows that it tends to produce routine, low-cost treatments where there is no obvious shortage of supply in the public sector (Campos, 1987). It is in ambulatory care, however, where financing is open-ended that we find the more striking departure from the NHS model. The provision of medical acts arising from NHS GP visits is dominated by the private sector. The private supply of pharmaceutical drugs is, of course, a feature of many NHS type systems, but in Portugal a large and rising proportion of diagnostic tests and treatments are contracted from the private sector, rather than being carried out in NHS hospitals.

In summary, although Portugal is commonly believed to have a system of the NHS type, the incentives built in to this structure are such that it tends to operate in a fashion not dissimilar to countries where there is collective provision of a basic level of care complemented by private individual purchase.

2.5 Health-equity policy objectives

Having outlined the principal characteristics of society, health and health care in Portugal, attention is now turned to health-equity objectives imprinted in legal and policy statements. This question has been examined in detail by Pereira (1990a,b), where it is argued that three seemingly distinct objectives have emerged. In the Constitution (Portugal, 1989), there seems to be implicit an objective in terms of *equal opportunity to maximize health potential*, through the access of all citizens to health promoting goods. This rather ambitious goal has not been clarified in subsequent documents, where the emphasis has been on health care, rather than health *per se*. Nevertheless, the Constitutional text is important to the extent that it recognizes that the key objective is the "right of all to the protection of their health" (*ibid.*) and that this right depends on equal access to both health care *and* other health promoting commodities.

A second objective is to be found in the National Health Service Law (Assembleia da República, 1979). Article 4 says explicitly that "access to the NHS is guaranteed to all citizens, independently of their economic and social status" (*ibid.*). Given that in the justification for the Law it is also suggested that the NHS should have "universal characteristics, in which all citizens have access, in equality of circumstances" (*ibid.*), the equity objective seems best interpreted in terms of *equality of access to NHS care, irrespective of social and economic circumstances*.

The third objective identified in Pereira (1990a,b) was similar to the preceding one, but appeared to widen the ambit of application to *health care in general (rather than simply the NHS)*. The distinction is related to the government's intention of subsidising the private insurance market so that patients are able to freely choose between the public and private sectors. Thus, a Law setting out basic principles of health policy declared that "it is a fundamental objective that citizens obtain equality of access to health care, irrespective of their economic condition or the place where they live" [Portugal (1990, p. 3452)].

Recent policy statements have added yet other equity objectives. For instance, a

document defining policy options for 1993 stated that the health care system should "aim to guarantee ... delivery of care according to population needs ... whether in the ambit of the NHS or in the private sector" [Portugal (1992, p. 5980)]. The equity objective in this case appears to be *delivery of health care in relation to medical need.* There have also been, for the first time, explicit pronouncements with regard to health care financing. The present Minister of Health recently stated that: "In general, the poorest members of the population ... have to be supported entirely by the State, and richer individuals - for reasons of social justice, equity and social solidarity - have to make a financial effort towards the maintenance of [health] services. ...That, for me, is the great principle of health care financing" [Mendo (1993, pp. 10-11)]. This seems to be a commitment to the egalitarian principle that *payments for health care should be related to ability to pay.*

Therefore, portuguese legislative and policy statements point to at least five different definitions of equity in the domain of health.[7] Though these are obvious yardsticks for monitoring the extent of health-inequity in Portugal, they raise a number of questions for applied work. For example, how may the definitions be given suitable empirical content? How are terms like 'related to' and 'equality' to be interpreted? How should key variables like health, access or economic status be defined? Prior to these issues, however, there are others that require investigation. For example, are some definitions more compelling than others? Can the concerns expressed in the different objectives be organized according to some underlying ethical base? If so, what does the normative framework suggest with regard to positive analysis? It is this second group of questions that command our attention in Chapter 3, leading to the adoption of Sen's capabilities framework as the normative base within which to organize applied research.

2.6 Previous Portuguese studies

Research aimed at measuring inequity in health and health care in Portugal is very limited and dates back no further than the beginning of the 1980's. Virtually all applied work is directed at examining the issue of geographical inequalities [see, *eg.*

---

[7] As is well known, such diversity is common to other health systems [see, *eg.*, Le Grand (1982) and Mooney (1983)].

Carrilho (1985), Kannisto (1986), Pereira *et al*, (1987), Giraldes (1988a,b) and Mantas *et al* (1991)]. The evidence from these studies points to the existence of widespread disparities in both health and health care. For example, Pereira *et al* (1987) showed that despite an accentuated decline in child death rates between 1972 and 1982, regional inequalities had scarcely improved and actually worsened in the case of perinatal mortality. The variation in death rates was found to be negatively correlated with indicators of wealth and income (though only post-neonatal mortality showed a statistically significant association), and with indicators of health care resources, utilization and expenditure. The same study also carried out an analysis of the spatial distribution of public health care. This revealed coefficients of variation in the order of 2.1 for hospital stays, primary health care consultations and prescribed medicines. As one would expect, NHS expenditure per capita showed similar disparities: on average almost twice as much is spent on patients in northern areas than those who live in the south. The variation in the number of prescribed diagnostic tests per head was conspicuously large: nine times greater in the Lisbon region than in the northern district of Vila Real.

Giraldes (1988a,b) carried out a comprehensive study of the allocation of NHS expenditure in the primary care sector, with a view to establishing an explicit equity-promoting distribution mechanism. If the mechanism were to be applied, Giraldes shows that a substantial reallocation of current expenditure from southern to northern districts would have to be undertaken. The picture for capital expenditure shows no clear North/South divide but also points to a need for extensive redistribution of investment resources.

A recent international review of applied research on socio-economic inequalities in health and health care (Mielck and Giraldes, 1993) found only two such studies in Portugal.[8] Lucas (1986) examined inequalities in mortality, morbidity and health care utilization according to occupational class. In the first case he used a sample of total deaths in 1981 to estimate male adult mortality rates. The results reveal a rising gradient from manual to non-manual workers, which is at its steepest in younger men.

---

[8] The issue of socio-economic inequalities has of course been the main focus of research in other countries and, as the previous section showed, often crops up in portuguese policy documents.

In the 25-34 age group an agricultural worker faces a risk of death 4.8 times greater than a high-level public or private administrator. This pattern was also found for virtually every cause-specific death rate. In the case of morbidity, data drawn from the 1983 National Health Survey (confined to the Lisbon Region), showed occupational class gradients for absence from work and confinement to bed due to illness in a two-week reference period. Manual workers were three times more likely to report illness than professionals, employers and managers. This pattern is more pronounced in younger age-groups and applies for both men and women. Finally, based on the same survey, Lucas showed that health care utilization by occupational class was roughly uniform for both men and women (*eg.* physician visits by manual workers only exceeded those of non-manual by between 10 and 30 per cent in different age groups). He argued that if these figures were corrected for need, as measured by self-reported morbidity, a class gradient in favour of higher occupational would emerge (though no actual evidence was presented).

Pinto (1988) measured the distributive impact of public provision of health care in the Lisbon region. The study used two data sources - a sub-sample of the Family Income and Expenditure Survey (1980-81) and the National Health Survey (1985). It computed economic benefits accruing to five socio-economic groups (determined by the educational level, occupation and main source of income of the head of the family). Comparing both years of reference, Pinto found roughly similar results: the incidence of benefits was characterised by a U-shaped pattern, signifying that the lowest and highest socio-economic groups derive greater amounts of economic benefit from the NHS. The 1985 data were also corrected for 'need', by considering only those individuals who reported illness in a two week reference period (see Table 2.7). Once again the U-shaped pattern was found, and it was shown, by disaggregating the types of care consumed, that the co-financing of drugs by the NHS was the principal factor in determining the overall incidence profile. Unfortunately, the author failed to age- and sex-standardize the benefit/need ratios, so that, in effect, the socio-economic groups are not strictly comparable. Given that there are a disproportionately higher number of older people in lower socio-economic groups, an age- and sex-standardized distribution might well have thrown up a pattern more clearly favourable to better-off individuals.

Table 2.7: Economic Incidence of Benefits in NHS care by socio-economic group, 1985

SEG	No. in sample.	Aver. benefit p/adult equivalent	Incidence index
V. (lowest)	383	53.1	159
IV.	2174	28.9	86
III.	724	19.6	59
II.	277	37.4	112
I. (highest)	11	36.4	109
Total / Aver.	4069	33.4	100

Notes: Estimated benefits = average costs of consumed care + reimbursement by NHS of privately consumed care.
Source: Pinto (1988)

Besides being relatively few in number, portuguese studies dealing with health domain inequity are limited in three major respects. The work reported in this thesis is largely aimed at redressing these shortcomings. First of all, existing research is seldom based on sound normative analysis, either of explicit objectives of portuguese health policy or of a more general nature (*eg.* drawing on the wealth of knowledge available in the sub-branch of welfare economics). Theoretical analysis of concepts such as inequality, health or access - which has underlined the inequality in health debate of recent years - has scarcely been touched on, with researchers adopting a pragmatic attitude that in a country with an under-developed information system the priority must be to extract the maximum descriptive elements available. Indeed, in some cases the work is guided by the simple objective of making available the type of information reviewed by the Black Report. The result is that it is often difficult to extract policy implications from the available research.

Secondly, the measurement techniques that are employed are often crude and uninformative. For instance, by far the most used inequality measure is the *range* which as Wagstaff *et al* (1991a) argue does not consider inequality in intermediate groups and takes no account of sizes of the groups being compared. Standardization for relevant intervening variables, such as age and sex, is also rarely found; and the

potential of econometric techniques to explain the persistence of inequalities has not been tapped. Given the controversy that surrounds measurement in the field, it is also surprising that no sensitivity analyses aimed at improving the robustness of conclusions have been reported.

A final problem is that existing research does not seem to have made adequate use of available data sources. For instance, the lack of country-wide information on health and health care inequality related to socio-economic status does not seem justified, given that a national health survey has existed since 1987. Furthermore, there have been scarcely any attempts to measure the evolution of health domain inequity, even though comparable data exists for recent periods.

Since Mielck and Giraldes' (1993) review, a further study concentrating on socio-economic inequality has been published. Pereira and Pinto (1993) measured inequity in the finance and delivery of health care, as part of an international research effort now known as the *ECuity study*.[9] This research interpreted equity in financing as the distribution of health care payments according to ability to pay and equity in delivery, as the distribution of access to care (proxied by utilization) according to need, rather than economic status. As the previous section showed, both interpretations are reflected in portuguese policy statements. Pereira and Pinto used data from the Family Income and Expenditure Survey (1980-81) and the National Health Survey (1987), and computed a variety of summary measures based on the concentration curve approach. They concluded that the "financing of portuguese health care is ... slightly progressive" and that in the delivery of care there is "inequity favouring the rich" (pp. 197-198). The research reported in this thesis is, in part, a development of the analysis carried out in Pereira and Pinto (1993). For instance, the finance-side analysis is updated with a newly available sample survey and improved by considering various empirical specifications that minimize the uncertainty surrounding theoretical concepts. The analysis of inequality in *health*, produced in the ECuity study as a by-product of measurement of inequity in the delivery of care, is here given a central role. These extensions mean that a number of new results are produced. Finally, the measurement

---

[9] The term *ECuity* derives from the fact that the study was financed by the EC, now European Union.

tools used by the ECuity researchers are brought under close examination leading to improved understanding of their normative and statistical properties and opening up new avenues for application. Thus, although the Pereira and Pinto (1993) study was able to amend some of the more obvious limitations of previous portuguese research, this thesis tries to show that the methods it used can be further improved and thoughtfully applied to data sources in Portugal, thereby enhancing knowledge of the extent of health-inequity in that country.

2.7 Concluding comments

In Portugal, as in other countries, the problem of equity in health and health care has attracted growing attention by researchers and policy makers. The work reviewed above shows that, in Portugal, inequalities in health care utilization and health outcomes do exist and that the patterns are similar to those found in other European countries. Though this research has been useful in highlighting the problem, it suffers from a number of limitations. Like much of its counterpart internationally, it is not based on the sound theoretical analysis. It also lacks technical rigour and is often guided by the ready availability of empirical elements. It is hoped to show, in subsequent chapters, that economic analysis has an important contribution to make in the resolution of these drawbacks.

A particular question which requires urgent consideration is the specification of equity objectives themselves. Section 2.5 showed that there are distinct goals proposed in various legal and policy statements. This presents obvious but not irreconcilable problems for researchers. In Chapter 3 the usefulness of economic analysis in clarifying societal equity targets is shown. This normative work also provides a guide to eventual monitoring of objectives through positive analysis. It offers, therefore, much of the theoretical support which previous work has lacked.

Finally, much of this chapter has been taken up in describing Portugal's socio-economic environment and the observed patterns of its health and health care systems. The main reason for this is that Portugal reveals distinct characteristics in relation to the

european countries to whom it is most often compared. The features that have been highlighted (*eg.* a predominantly rural population, the transitional nature of health problems, the importance of direct payments in health care finance, etc.) need to be borne in mind when interpreting later empirical results. Indeed, as will be seen, they also have a role in determining the very questions that are addressed.

Part 2

NORMATIVE SPECIFICATION

Chapter 3

What does Equity in Health Mean?

" So I must confess that the outcome of the discussion is that I know nothing. After all, if justice still remains undefined, I can hardly know whether it is in fact a virtue or a vice. Nor can I know whether the just man is in fact happy or miserable"
Plato, The Republic

3.1. Introduction

A significant feature of the health inequality debate is that, until recently, it tended to produce a wealth of empirical facts while at the same time disregarding the precise specification of equity objectives. This was unfortunate, given that normative investigation is an essential prerequisite for understanding the reasons why people care about social justice in the field of health; the extent to which specific types of inequality are compatible with equity; how the concept should be measured; and how rational policies may be formulated and monitored.

Economists were early to see that without more explicit consideration of normative issues the debate on equity in health and health care would remain "confused and confusing" (Mooney, 1983). A limited amount of work was initially undertaken on specifying egalitarian health policy objectives (Le Grand, 1982; Mooney, 1983). Notions such as equality of public expenditure on health services, equality of access to health care and equality of health itself were discussed, yet there was a failure to locate these specific objectives in theories of society and public policy or to relate them to existing economic definitions of equity.

More recently, a number of authors have reassumed the difficult challenge of disentangling normative issues in the health-equity debate (*eg.* Mooney and McGuire,

1987; Le Grand, 1987; Culyer, 1990). Their work is an encouraging, if belated contribution, clarifying the definitional steps that should precede a rigorous positive analysis of the problem. The present chapter, whilst addressing itself primarily to the interpretation of portuguese policy objectives, tackles these same questions. In essence, it searches for rules that are capable of shedding light on equity concerns in the field of health and that may serve as an ethical guide to later empirical measurement. The main argument is that Sen's notion of equality of capabilities, strangely ignored by health economists in the past, provides such foundation. Moreover, it seems a useful interpretation of the Portuguese Constitutional objective, from which presumably other national targets pertaining to the distribution of health care are derived.

The chapter is organized as follows. Section 3.2 sets out criteria for assessing the various distribution rules under review. Section 3.3 critically appraises six well-established approaches to defining equity - egality, entitlement, the decent minimum, utilitarianism, Rawlsian maximin, and envy-free allocations. Each of these is found wanting in some respect when applied to the health sector. Section 3.4 turns attention to two alternative formulations recently proposed by health economists - equity as choice and health maximization.[1] Despite providing some interesting insights these too have limitations. Section 3.5 describes the capabilities approach and applies it to the specific domain of health. Section 3.6 discusses measurement implications of the Sen approach and proposes an agenda of positive analysis for studying inequity in health and health care based on the underlying ethical framework. This agenda is followed in remaining chapters. Finally, section 3.7 concludes the analysis.

3.2 Assessment criteria

Suggestions with regard to standards for evaluating equity proposals abound in the literature [see, for example, the contributions of Rawls (1971), Pazner and Schmeidler

---

[1] It will be noted that the discussion omits the important recent contribution of Culyer and Wagstaff (1993) and other related papers by the same authors. The reason for this is that the Chapter was prepared, and published, before these articles appeared. I have opted to leave the discussion more or less in its original form. It may be noted, however, that Culyer and Wagstaff's analysis has some points in common with Sen's specification (*viz.* the emphasis on ultimate entities of concern and the implication that empirical analysis must consider inequality in the distribution of health and in the distribution of those factors, such as health care, that contribute to it).

(1978), Le Grand (1984) and Baumol (1986)]. Some writers stress that we should merely look for clarity and specificity while others propose complex mechanisms such as the Rawlsian "social contract", whereby from a hypothetical "original position" individuals establish an acceptable equity criterion. Others still, propose more restrictive criteria such as the argument that a two-step approach which distinguishes equity from efficiency decisions is logically untenable or that only individual preferences, in contrast to third-party values, should count in judging the equity of a particular distribution.

In this chapter four fairly obvious requirements that an acceptable formulation of equity in the field of health should meet are put forward: (i) a conception should be *easily comprehensible* so that it allows the widest interdisciplinary discourse and deduction of *clear policy solutions*; (ii) it should be *specific and rigorous*, in order that concepts are not left so vague that they generate misunderstanding in application; (iii) the formulation should be readily susceptible to *empirical verification*; and, (iv) the definition should be *intuitively and widely acceptable* for the problem at hand.

The final criterion can obviously be given various interpretations. I shall take it to mean that an equity formulation should not disaccord with the concerns revealed by *health-equity policy statements* in Portugal (see Chapter 2).[2] Roughly stated, some key ideas defended in those statements are as follows. There is, first of all, a distinct concern for *distribution*, separate from other explicit or implicit targets like efficiency, maximization of survival or consumer choice. Second, equity appears to require *equalization* of specific parameters across socio-economic groups rather than a basic minimum of provision. Finally, the *processes* of health production and health care delivery are an important element in reaching decisions on whether equity is being achieved.

Besides the criteria above, other themes that have been raised in the literature are used here as a means of highlighting the differences and implications of proposed formulations. For instance, one might want to know whether a formulation bases itself

---

[2] Given that portuguese health-equity objectives are remarkably similar to those of other countries in W. Europe (see the contributions in Van Doorslaer *et al*, 1993), this is not as restrictive as it seems. Consequently, many of the implications of the analysis are applicable in a wider context.

on aggregation of individual preferences or if there is a remittance to external judgements? What variable is chosen as the metric for judging the equity of a particular distribution: health, health care or utility? Or if any priority is afforded to equity decisions over those which concern efficiency: that is, whether the two objectives are accomplished in a single step or separately? Though some authors have taken particular aspects of these questions to constitute *a priori* assessment criteria it should be noted that here they serve no other purpose than that of contrasting different equity conceptions.

3.3 A critique of traditional approaches to conceptualizing equity

There are six well-established conceptualizations of equity in the literature that may be considered relevant for the health inequality debate: egality, entitlement, the decent minimum, utilitarianism, Rawlsian maximin and envy-free allocations. Virtually all emanate from the discipline of political philosophy, but they share the characteristic of having drawn significant interest from economists. Their status in the literature as specifications of what equity generally entails is not in doubt, but it will be argued that, when applied to the health sphere, they reveal particular shortcomings. I begin with what constitutes the most direct philosophical foundation for the type of policy objectives identified earlier for Portugal: the theory of egality.

3.3.1 Egality

Egality is sometimes taken to mean equalizing individual net benefits (*eg.* health status) or, once it is admitted that some attributes cannot be physically distributed, equalizing individual opportunities for such benefits. In an influential discussion, the philosopher Ronald Dworkin (1981) has distinguished between the two key notions of equality of welfare and equality of resources, arguing that any ethically supportable egalitarianism must call for equalizing the resources available to people, not their welfare or utility. Equality of welfare holds that: " a distributional scheme treats people as equals when it distributes or transfers resources among them until no further transfer would leave them more equal in welfare." Equality of resources, in contrast: " treats

them as equal when it distributes or transfers so that no further transfer would leave their share of the total resources more equal." (Dworkin, 1981, pp. 185-186). Important as they are, these distinctions only raise further questions. For instance, does equality of welfare in the health field require equality of health or attainment of equal levels of utility? Does equality of resources require simply equality of access (or opportunity of access) or does it require the use of resources in equal quantities? Should the definitions be applied in relation to State provided health care or across all resources, public and private?

Lengthy discussions of the distinctions possible within this approach and the competing policy objectives which they imply have, of course, been a feature of recent contributions to the health and social policy literature. Mooney (1983) and Le Grand (1982) proposed a number of interpretations which might be used as guides to health policy.[3] Their definitions may be classified under three separate headings. Those which are formulated in terms of their impact on supposedly homogeneous populations, without regard to differences in health status or need for health care (*eg.* equality of public expenditures per capita); those which relate equity to people's need for care (*eg.* equality of treatment for equal need); and those which focus on the outcome of health care activities (*eg.* equal distribution of health itself).

Though this early work was useful in highlighting the varied interpretations that might be given to the objective of achieving equality in the health domain it failed to relate policy definitions to their economic or philosophical base. Significantly both Mooney (1987) and Le Grand (1987) later argued that the definitions suffer from a number of analytical and practical problems and may in some cases conflict with commonly held views of what is just and fair. Indeed, one cannot help but feel that the lack of rigorous and consistent health-related analyses, within the traditional egality perspective, is due to it remaining too elusive a concept as a principle of distribution.

Arguably, much of the applied work which implicitly draws on the egalitarian view is too permissive to be useful for policy recommendations. The Black Report

---

[3] See also, in this context, the more general discussion by O'Higgins (1987).

(DHSS, 1980), for example, concentrates overwhelmingly on inequalities in *health* without clarifying any logical or policy basis for health equality being pursued. The reason seems to be that egality as an equity formulation simply lacks the specifity required, allowing researchers to imbue their own, or the information system's, values into the monitoring process. Furthermore, the concept fails to establish coherently why equalization of any type should in fact be accomplished. This vagueness and lack of development have been sorely felt by defenders of NHS-type arrangements in the inequality in health debate, which is all the more disheartening since, as will be shown later, an approach is available that can resolve many of the pitfalls which a referral to egality can engender. It does so by analysing thoughtfully the transmission process from resources to outcomes, which in the final analysis is the principal oversight in the egality account.

3.3.2 Distribution according to entitlement

Perhaps the best known rejection of equality has been provided by the libertarian philosopher Robert Nozick (1974) in his theory of distribution according to 'entitlement'. Its core position is that one is entitled to what one possesses provided it was acquired justly: that is, through earnings, through inheritance or through redistribution by government of holdings acquired illegally. It is, therefore, a procedural theory: whether or not a specific distribution is considered equitable depends entirely on the path used to reach it. While this is arguably a desirable characteristic of an equity formulation, its implications for distribution in the health domain are manifestly out of step with the concerns found in portuguese policy statements.

It would seem that both health *and* health care are suitable metrics for assessing equity within the entitlement framework. The only important empirical question, in this context, would be to determine the *manner by which* these two commodities were acquired. It is not difficult, however, to think of examples where its application would lead to outcomes widely regarded as inequitable. Consider the case of health, specifically where a child is born with a congenital deformity. Since the condition was inherited a strict application of 'entitlement' holds it to be fair. Nozick does refer to the

possibility that in the case of "catastrophic moral horrors" entitlement rights might be compromised, but it is not at all clear how his theory would accomodate such waiving of rights, in the absence of formulation of other competing bases of moral judgements. So generally it may be said that the theory attaches no weight to the unfortunate: it is essentially a matter of fate that some are born in a healthy condition and others are plagued with chronic medical problems.

With regard to health care Nozick's approach is similarly restrictive. It suggests that no one citizen has a right to health care unless it has been acquired through the market. Attempts at redistributing resources, even if they were aimed at providing incentives for those who use health services less efficiently (*eg.* the less educated and the poor) would in themselves be considered an injustice. Nor is there recognition of sentiments of caring or generosity by the well with regard to the unwell, often given practical expression in the subsidization of health care services (Culyer, 1980). Finally, it fails to consider either the role of possessions which are received as social goods or the pervasion of externalities and consumer ignorance in the health care market. It matters not, therefore, that the conception is clear and specific, since its pursuit would, in all likelihood, lead to a distribution highly unfavourable to the poor and the sick.

3.3.3 The 'decent-minimum'

Given the extreme consequences of Nozick's principle of distribution for social or health policy other Libertarians have suggested a role for some sort of safety net, *viz.* a standard below which individuals should not be allowed to fall. Such an approach - often designated the decent minimum - should, if applied rigorously, simply pertain to final outcomes of a process (*ie.* health itself). Invariably, however, it is specified as the provision of a minimum standard of *health care*, and points towards a configuration of services strongly weighted towards the private sector, with the State providing a limited and minimal level of care for the poor. Given that the Portuguese health care system aims to achieve equity through universal coverage of citizens and general coverage of benefits, it is doubtful whether the present approach could be considered suitable for that particular system.

It could be argued that by circumscribing the demands of equity to the provision of a 'decent basic minimum' the procedure is strong in terms of practical applicability. This is somewhat misleading since the key to its operationality is that it requires a value judgement as to what constitutes the decent or social minimum. Perhaps in recognition of the problems involved, its proponents have been reluctant to define exactly what it is. In the health field only Enthoven's (1980) discussion of a Consumer Choice Health Plan comes close to doing so. He suggests a list of "basic health services" which Health Maintenance Organizations (HMO's) should provide. But it is far from clear that at the end we are left with a clear idea of what constitutes a decent minimum, since what are accepted as "basic services" may differ across time and contexts and because he provides no sound reason as to why certain types of care should be left on or off the list. Ultimately, the distinction Enthoven makes between high and low option plans suggests that we can choose the 'decent minimum' by reference to average costs for actuarial categories. This seems a somewhat unjust principle for allocating health care.

Loewy's (1987) specification of the decent minimum is different in that he points to criteria rather than supplying a list of items. Although recognizing that specifying a decent minimum is conditional on historical and cultural contexts he is willing to suggest certain minimal conditions which "all would hold to be self evident" (*ibid*). Certain 'things' should be provided: namely those necessary to save life; ameliorate suffering; restore function of vital parts; and prevent future problems by public health, immunization and sanitation. In effect, this view brings us no closer to a rigorous specification of the decent minimum, although there is more than a suggestion that it is approximated by the traditional medical model and its views on the ethics of health care. But as Kennedy (1983) argues a concern for the community values of equity or justice surely implies the abandonment of such a model.

All this goes to suggest that the definition of an acceptable minimum standard is a complicated exercise. But suppose for the sake of argument that one were to be found: what would be the implications for applying the concept? It would seem that two measurements are required: the number or proportion of individuals not achieving the standard and the total quantity of the good required to raise all those below to the level of the accepted minimum. Posed in this fashion the problem appears as simply

a supply-side phenomenom: redistributing health care resources from individuals above the minimum to those below. But if the supply of health care is less than perfectly inelastic it may be more productive to influence other inputs (*eg.* education or income) into the health production function, thereby increasing the demand for health care of those individuals below the minimum. By providing a simple uni-dimensional view of the equity question, the decent minimum approach overlooks this possibility. It appears then that a rule of allocation along these lines offers little scope for operationality, acceptability or precision.

3.3.4 Utilitarianism

The goals of utilitarianism are commonly summarized as "serving the greatest good for the greatest number." In economic terms this implies a decision rule where resources are allocated so as to maximize aggregate utility. Quite why utilitarianism should be seen by so many as a theory of equitable distribution is difficult to perceive. Possibly it has to do with the well established result that an egalitarian distribution will result under classical utilitarian principles when there exist identical preferences (Culyer, 1980). But there is no logical connection between greater equality and greater equity. There is, however, a stronger argument, brought out in Sen's (1973, p. 16) well known comment that "maximizing the sum of individual utilities is supremely unconcerned with the interpersonal distribution of that sum." The activities or individuals to which resources are allocated at the margin depends simply on comparisons of utility. Thus, if a rich individual responds better to a given course of treatment than a poor one, the utilitarian decision rule requires that more resources be attributed to him. The resulting distribution may well be efficient but it is unlikely that it will conform to most people's conception of equity.

Though it could be argued that such a reaction is simply the super-imposition of an alternative (non-explicit) moral judgement, it is based on most solid grounds. There are essentially three principles which underpin utilitarianism (Sen, 1987a). They are: *welfarism*, which implies that extra-utility information (such as individuals needs, capacity for mobility and so on) is either irrelevant or only indirectly relevant as a causal influence on utilities; *sum-ranking*, which asserts that the goodness of a collection

of utilities is simply their sum, thus eliminating the possibility of concern over inequalities in their distribution; and *consequentialism*, whereby all choice variables are judged simply in terms of the goodness of their respective consequences. In the health domain only the latter is defensible - in the sense that it excludes the acceptance of health care activities that do not promote the ultimate goal of good health. The other principles are simply too restrictive as basis for forming equity judgements.

There are also various technical problems associated with utilitarianism, all inevitably linked to the impossibility of measurement and interpersonal comparisons of utility. Indeed, the identification of a just utilitarian distribution depends upon such a wealth of empirical facts which are so difficult to obtain that it seems unproductive to attempt to apply it to health and health care. Furthermore, these factors are not directly deducible from the principle itself, which further complicates the exercise. It seems, then, that an appropriate conceptualization of health-equity concerns must be sought elsewhere.

3.3.5 Rawlsian maximin

Another prominent philosophical discussion of social justice, which has attracted the attention of economists, is John Rawls' (1971) theory of maximin. It makes justice an uncompromising aim in suggesting that social policy, rather than maximizing net benefit in society, should seek to maximize the position of the least well-off. Rawls considers a set of goods whose production and distribution, he suggests, should not be left to individuals themselves. These 'primary social goods' include basic liberties; freedom of movement and choice of occupations against a background of fair opportunities; powers and prerogatives of office; income and wealth; and the social bases of self-respect. Rawls then hypothesises an 'original position' where all individuals operate under a 'veil of ignorance.' In such a context rational men would be risk averse and choose as a preferred arrangement a situation where the worse off have their position maximized. What drives them to such a choice is not a concern for the least advantaged but a fear that they themselves might turn out to be, once the veil of ignorance is uncovered, the worst-off citizens in society.

According to Le Grand (1987), the application of Rawlsian maximin to the health field requires that inequalities in either health or health care be justified only if they operate to the benefit of the least advantaged. He criticizes such a rule as a guide to health policy on two grounds. First, because it raises a number of theoretical and practical difficulties. For instance, are the least advantaged to be defined in terms of their overall consumption of primary goods or in terms of health or health care? Furthermore, is it realistic to suppose that we can readily distinguish those inequalities that benefit the least well-off from those that do not? One could add that the principle implicitly suggests that an equitable distribution would be that where all individuals have the health status of the sickest person. The second objection, has a libertarian strain. It is that maximin would lead to redistribution to those whose poorer health, inadequate consumption of health care or actual poverty were the result of their own decisions. Arguably, however, Le Grand's direct application of the Rawlsian principle to the health field is too ambitious, for neither health or health care were designated as primary social goods by Rawls himself. Indeed, including either health or health care would imply trade-offs with other primary social goods such as income and wealth and inevitably interpersonal comparisons of utility which Rawls is keen to avoid (Arrow, 1974).

Daniels (1981) has suggested that the most promising strategy for extending maximin theory to the health domain is to include health care services among the background institutions involved in providing for fair equality of opportunity. This is justified in the sense that health care is necessary for normal species functioning. However, such an approach merely has the effect of collapsing the definition of equity into one of equality of opportunity of access to health care for equal need.[4] Therefore, although this interpretation appears in tune with the concerns of Portuguese health policy it has the unfortunate effect of making the theoretical structure redundant. We are left with no more than a simple interpretation of equity, which is problematic in terms of specificity, and no idea as to how the concept may be applied in positive analysis.

---

[4] Daniels himself readily admits that his account "does not pressupose the acceptability of Rawls' theory" (Daniels, 1981).

3.3.5 Envy-free allocations

The theoretical and practical problems associated with utilitarianism and maximin have led to a number of economic discussions which seek to provide a more rigorous grounding for equity concerns. The dominant approach concentrates on defining the essential characteristic of an equitable distribution. This is best described, it is suggested, by the criterion of non-envy: that is, where a person's relative advantage is judged by the standard of whether he or she would have preferred to have had the commodity bundle enjoyed by another person (Varian, 1974; Pazner and Schmeidler, 1978; Baumol, 1986). To illustrate the concept, in the context of a simple exchange economy, consider any allocation x_i^g (g=1 to n; i=1 to m) of n goods to each of m individuals. Suppose these individuals have preferences represented by the ordinal utility function $U_i(x_i)$ (i=1 to m) of each individual i's own consumption vector x_i. Then individual i is said to *envy* j if $U_i(x_j) > U_i(x_i)$. An equitable distribution is defined formally as that where $U_i(x_i) \geq U_i(x_j)$ for all pairs of individuals i and j.

Defenders of the approach have argued that it provides an easily comprehensible and specific formulation of equity; that it avoids the arbitrariness of external moral viewpoints by judging the desirability of a distribution exclusively in terms of the preferences of individuals affected by it; and that it lends itself well to the standard constructions of the economist (*ie.* indifference maps and utility theory). This much is true. The theory has indeed been used with considerable formal elegance to establish the conditions under which allocations are simultaneously equitable and Pareto-optimal (so-called "fair" allocations[5]), and hence provides a theoretical insight into the trade-off between efficiency and equity.

However, despite its attraction to economists it is doubtful whether the non-envy approach could be suitably applied as a guide to equitable health policy. On the one hand, it is well-established in the technical literature that the pursuit of non-envy can lead to some peculiar and unpalatable results (Feldman, 1987). When agents are more or less symmetrical the concept seems to work quite well; yet if one or more agents

---

[5] Baumol (1986) has more recently introduced the term "superfair" to describe simultaneously equitable and Pareto-optimal allocations.

happen to be, say chronically sick, there is no opportunity for exogenous compensation within the framework. Similarly, it could not account for a case where a kidney patient's demand for dialysis takes precedence, by general agreement, to a tennis player's demand for rackets; or less trivially, to an occasional headache sufferer's demand for analgesics. What is missing then is a view of what others might regard as equitable and not simply oneself. By concentrating exclusively on individual preferences the concept overlooks that generally, when making a judgement concerning the justice of a situation, one would wish to allow for differences in tastes, needs and so on. In this more common situation the appropriate comparison in determining what is inequitable becomes: whether $U_j\ (x_j) > U_i\ (x_i)$; rather than, as the non-envy account suggests, if $U_i\ (x_j) > U_i\ (x_i)$.

Should one insist on overlooking these most obvious drawbacks and proceed to apply the concept to health policy it seems that only health care should be considered as a metric, given that health itself is indivisible. But even then the approach is uninformative since it fails to provide a more or less complete ranking of alternative states, which clearly appears necessary in the health field. It gives only an answer as to what constitutes a fair distribution; should no such feasible allocation be found (as seems to be the case in many situations in the technical literature, particularly when production is introduced) one is left with no suggestions as to how decisions should be taken. Finally, it is doubtful whether a formulation founded on the idea of "envy" could gather wide acceptance as a measure of equity in health care. This may seem a strange comment when much of economics is built around the deadly sin of "greed", but arguably, a concern for health-equity derives precisely from the desire not to allow self improvement to override community interests (Mooney, 1986). Certainly this is the picture which emerges from health policy statements in Portugal and other european countries and thus *non-envy* must also be rejected as a suitable normative framework for considering the attainment of health-equity objectives.

3.4 Approaches from health economics

Given the assessment criteria formerly laid out, each of the five traditional conceptualizations of equity discussed above have been found wanting in some respect.

Recently, alternative health-specific equity concepts have also appeared in the literature. I shall concentrate on the two approaches which have attracted the greater attention by researchers in the field: one, a formulation advanced by Julian Le Grand which is centred on the role of choice in determining inequalities, and another associated with the York school of health economists, which if adopted holds considerable implications for the way equity is interpreted and monitored (the health maximization account).

3.4.1 Equity as choice

Le Grand's guiding principle is stated as follows: "if an individual's ill health results from factors beyond his or her control then the situation is inequitable; if it results from factors within his or her control then it is equitable." (Le Grand, 1987a).[6] Formally, the representative individual is said to be faced by a choice set which he seeks to maximize subject to constraints. These are defined as the factors beyond individual control and obviously limit the range of possibilities over which a person can make choices. In this context, an equitable situation is that which is the *outcome of individuals choosing over equal choice sets*. What matters here, then, is not the end result but the history of a specific situation. In this sense, equity as choice hearkens back to Nozick's (1974) entitlement formulation.

Figure 3.1 (reproduced from Le Grand, 1987a) illustrates the general argument. An individual's health status (h) is plotted against the quantity of a health-harming activity (q), such as smoking, drinking, or working in a stressful environment. It is assumed that a trade-off is possible between h and q. Le Grand considers two individuals, A and B, whose choice sets between h and q are identical and portrayed by the frontier RT. Both derive utility from the health harming activity and from health itself, but individual A derives greater pleasure from q relative to h, when compared to individual B. These assumptions are incorporated in the position of the indifference curves U_a and U_b. A's equilibrium point (determined where U_a is tangent to RT) implies a lower level of health (h_a) than individual B's (h_b). According to Le Grand's conception this situation is not inequitable since both A and B have made informed

---

[6] See also Le Grand (1984, 1991a) for further discussion of the approach.

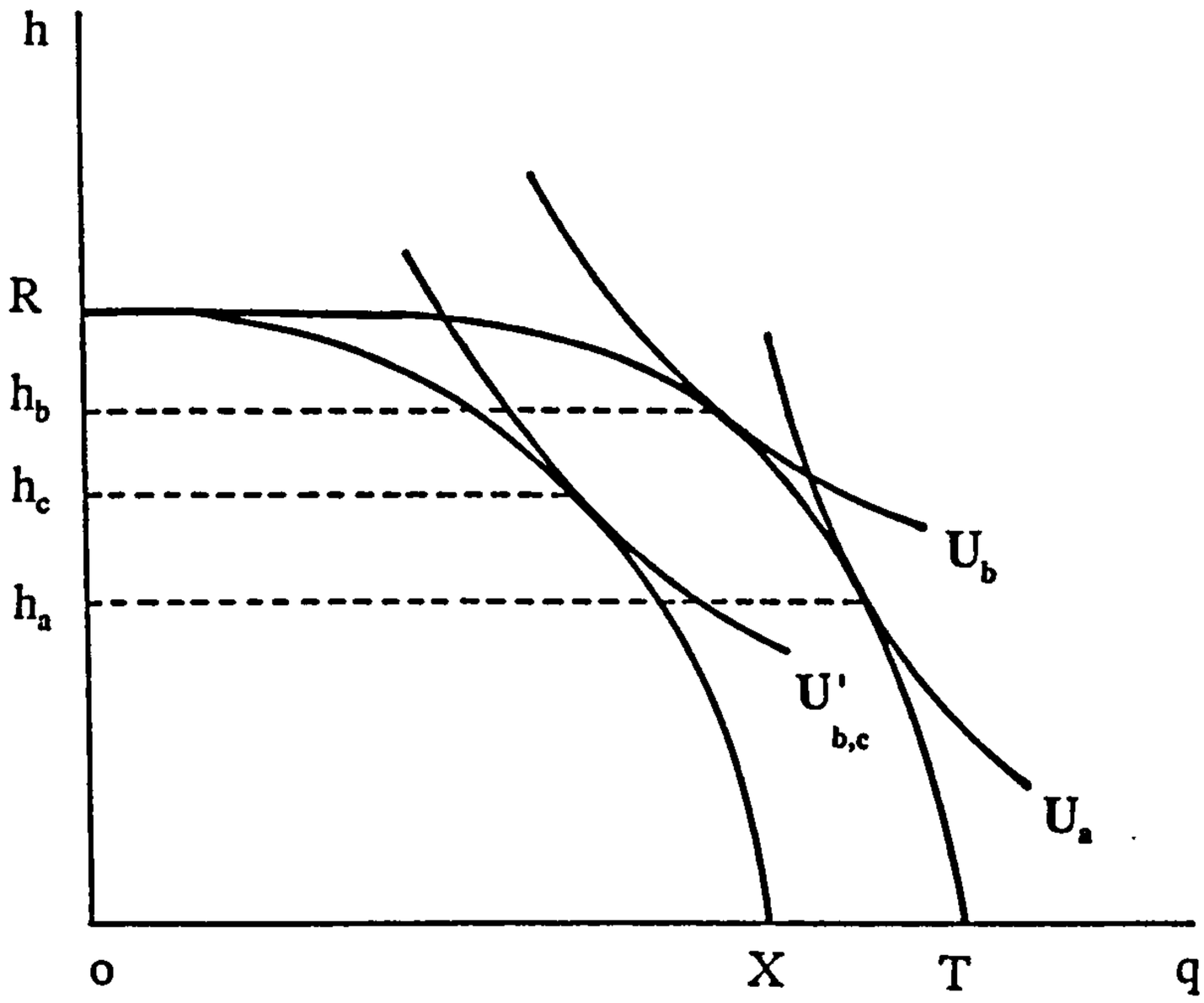


Figure 3.1: Le Grand's equality of choice sets

decisions, exercised over the same range of choices and based on their own particular preferences.

A third individual, C, is now introduced into the analysis. His preference ordering is the same as B's, but given that he faces a different choice frontier (RX), his equilibrium health level turns out to be below B's at h_c. The shape of RX, according to Le Grand, incorporates the assumption that individual C is poorer and less able to withstand the effects of the health-harming activity. The differences in health between B and C are not held to be equitable, since they arise from different feasible choice sets rather than from dissimilar preferences. Therefore, distributions are only equitable if they are the outcome of individuals making choices under equal constraints.

Le Grand's notion of equity has obvious counterparts in the general economics literature, particularly conceptions which stress the importance of claims over commodity bundles and resources (*eg.* Archibald and Donaldson, 1979). The question here, however, is whether it is a suitable formulation of equity concerns in the field of

health? It has some obvious advantages. On the one hand, by paying due respect to a distribution's history it serves as a useful reminder that information on end-states may not provide a sufficient basis for making equity judgements: it is equally as important to know how a particular distribution came about, whether it be health or health care.

Equity as choice also reveals promise for application in positive analysis. In Grossman's (1972) model of the demand for health it has a ready made framework of individuals exercising choices regarding health investment and consumption decisions within constraints. Equity under that model could be interpreted as equalizing the present cost of health investment for all individuals. Intuitively such an approach appears remarkably similar to equalizing the constraints people face. Furthermore, though Le Grand does not directly point to it, there seems to be no reason why the concept should not be used to study inequality in health care rather than inequality in health. Once again our interest would seem to be better directed at the process of health care consumption (*eg.* time spent in a waiting-room, out-of-pocket payments, etc.) rather than its output (*eg.* utilization rates or overall expenditures). If the former are differentially burdensome it can reasonably be argued that choices are not being exercised under equal constraints. Here too there are examples in the literature with regard to empirical analysis [*eg.* Sloan and Bentkover's (1979) equality in process variables approach].

Unfortunately, the equity as choice account is also open to a number of criticisms. It is far from clear, for example, that Le Grand has established, as he contends, a definition of equity which commands wide agreement in society. Some might argue that in the field of health, where uncertainty and consumer ignorance prevail, individuals are simply not in a position to make informed decisions. This problem is particularly acute in the case of medical care, while addiction to health-harming activities, as Le Grand accepts, poses related difficulties. In short, the assumptions of autonomous preferences, complete certainty and perfect information appear rather extreme in the health context. While on the one hand, this is not a crucial argument since relaxing the assumptions in positive analysis could make the account more relevant[7], inevitable problems remain

---

[7] The introduction of uncertainty into the Grossman model as in Dardanoni and Wagstaff (1987) is a case in point.

in deciding what precisely is to be ascribed to choice and what is not. So far the arguments put forward have done little to dispel the fear that such a task is likely to be accomplished in a less than rigorous manner.

Consider, for example, Le Grand's (1987a) discussion of policy implications. One cannot help but be surprised how from an individualist framework arise fairly conventional egalitarian arguments. Although this is, of course, not inherently impossible there is a problem in so far as the latter are not intuitively derivable from the theoretical construct but depend at various stages on the introduction of further value judgements. It begins with the ruling out of equity as choice as a guide to allocation of treatment, because health professionals are judged not to be able to undertake such decisions. Rather, it is suggested that the criterion should only be applied to decisions on individual or community financing of treatment. It is shown that in this case applying equity as choice would yield the development of a perfectly competitive insurance market as the optimal policy. Confronted by the extremeness of this implication, which would leave the poor and the risk-averse uninsured, Le Grand suggests a role for exogenous compensation and opts for a pragmatic solution where a government agency levies a uniform charge on all individuals. None of these steps are logically derivable from the account. They are simply the result of further value judgements being introduced, because the probable outcomes are viewed as inequitable! This problem arises, of course, because the concepts of choice and constraints have been vaguely defined. Thus it is possible to transform what is apparently a precise conception into one where at every stage new value judgements are introduced if outcomes appear unfair. The dividing line between variables over which individuals can exercise choice and those which constitute constraints must, therefore, be the subject of careful definition in future.

Another problem with the approach is brought out in the diagrammatic exposition. A rigorous application of the concept would seem to point to choices being exercised simply over commodity bundles or resources (Archibald and Donaldson, 1979). Le Grand, however implies a trade-off between an activity (which in the example is no more than consumption of the good tobacco) and an individual's health status. Supposedly this is justified in so far as 'smoking' and 'health' are fundamental

commodities in the household production theory sense (Becker, 1965; Lancaster, 1966). But this may be a slightly over-ambitious interpretation with confusing policy implications. In health production individuals do not directly trade-off health-harming or health-producing activities with health itself. They 'choose' their health levels by trading-off the activities against each other: either investing or disinvesting in their health stock; in the former case through consumption of health care or education, in the latter through consumption of insalubrious lifestyle or nutrition commodities. The final product - healthy days - is the result of trade-offs between these commodities. Therefore, although it is undoubtedly an advance on traditional formulations when applied to the field of health, some groundwork would appear necessary if Le Grand's concept is to serve as an appropriate guide to positive analysis.

3.4.2 Maximization of health

The other important principle to have emerged in the health economics literature is conveniently summarized by the phrase that a distribution is equitable *if and only if it serves to maximize the health of the community*. From a somewhat circumspect beginning the criterion has arguably become the principal focus of debate on the normative aspects of equitable health policy in the UK.[8] It is fair to say that, relative to other specifications, its principal distinguishing feature is that it affords primacy to efficiency; distributional questions are only important when judged in terms of their contribution to that goal.[9] This has exposed the maximization thesis to the obvious criticism that its proper domain is the assessment of *total good* rather than the *fairness* of a situation (Broome, 1988). Dismissing the account on these grounds seems, however, altogether premature since a case can be made for interpreting health maximization as the logical corollary of a particular type of distributional concern.

Rather than appealing directly to a general principle of justice or defining what

---

[8] Indeed, a recently published volume (Bell and Mendus, 1988) is almost entirely dedicated to the question of whether health maximization is a suitable equity rule.

[9] Culyer (1988) asserts categorically that "equality matters ... only when it serves the cause of efficiency."

constitutes an equitable distribution, health economists have traditionally looked at what motivates individual concerns for fairness in health. One group of writers have formalized such an approach through what is commonly termed the 'caring externality' [Lindsay (1969); Culyer (1971, 1980)]. Individuals are held to be concerned not only with the bundle of goods and services they are to receive but also with that to be had by others. In this sense, generosity, sympathy or caring are explicitly incorporated into the analysis through the mechanism of specifically interdependent utility functions developed by Hochman and Rogers (1969). This is in stark contrast to the non-envy approach which has dominated economic discussions of equity. There, individuals consider the consumption bundles of others merely for the effects of comparison. Culyer (1971) postulates that it is the quantity of suffering rather than its distribution which forms the basis of the externality relation. This suggests that a preference for increased consumption of health care should be an argument in the utility functions of the well-off, rather than equality of health care consumption as in Lindsay (1969), for it is actual under-consumption by the needy (implicitly associated with greater suffering) which imposes an external disutility on others.

Although this research was originally conceived as an explanation of widespread support for public financing of health care, indirectly, it holds important implications for the type of equity which should guide health policy. In particular, there are three key insights suggested by the approach: health status as the focus of concern; an absolute rather than relational objective; and a role for exogenous compensation. Strangely few writers have picked up on the implications of these aspects for equitable health policy. One reason why this might be so is that the account belongs to a completely separate literature. Equity is about fairness or justice, it refers to what people are due as of right; compassion refers to a desire to provide help to others regardless of their 'due'. This argument, however, misses the point that the two are simply alternative forms of justifying redistribution. One may wish to appeal to a general principle (as in the approaches discussed beforehand) or rely on the exercise of charitable compassion as foundation. The crucial point is that the compassion/caring approach itself suggests criteria for redistribution.

Some confusion appears to have developed in the literature regarding the actual

definition of equity which derives from the caring externality approach. Mooney (1987) suggests that it is 'equality of utilization for equal need.' Culyer (1976) himself, argued that it "leans towards an egalitarian notion of 'equally available care for equally sick' or 'communism in health'." None of these conceptions accomodates adequately what is perhaps the most important insight of Culyer's (1971) analysis: the health status of the 'needy' as the source of concern. Individuals care for others not because they are poor or materially deprived but because they are sick. Although in practice it may well be the rich who feel compassionate towards the poor there is no intrinsic reason why compassion and hence generosity should not be addressed to the rich sick person. This urges that health status be made the focus of distribution rather than income, welfare, commodities or primary social goods. Yet health is not a tradeable commodity, so how is one to develop a consistent definition of equity? It seems that only two routes are logically sustainable. On the one hand, considering a distribution as equitable if consumption of commodities which affect health is optimized and on the other, viewing any arrangement which maximizes health in the community as inherently equitable. In both cases the emphasis is on maximization/optimization rather than distribution; on absoluteness rather than relativity.

The second approach has been taken up by economists at York University. Drawing on the development of the Quality-Adjusted-Life-Year (QALY) measure of health, it has been suggested that general maximization of that metric is the most suitable means for achieving distributional goals. Hence, health maximization is treated implicitly as an equity principle in itself (Williams, 1988; Culyer, 1988, 1989, 1990). It should be noted that these authors have not sought support for the rule in the caring externality framework, even though it appears to provide a case for pursuing community health maximization. Instead, justification has been offered on essentially instrumentalist grounds: namely, that health services exist to promote health and hence, given scarcity of resources, one should strive to maximize the benefits accruing to the community (*ie.* health itself). The incorporation of specifically distributional concerns is then achieved by attaching weights to outcome data. It seems, therefore, that an appraisal of the account in terms of its suitability for informing the health inequality debate rests on the appropriateness of three related ideas: the QALY metric itself, which is the preferred outcome measure, the belief in health maximization as an

embodiment of equity concerns and the notion of attaching distributional weights to outcome measures.

QALY maximization remains a controversial topic among health professionals (Smith, 1987), social scientists (Harris, 1987; Carr-Hill, 1988) and even among economists (West, 1986; Broome, 1988; Loomes, 1988). Besides furnishing the literature with a seemingly endless string of titles which are puns on the term "QALY", these works have raised some apposite objections to the measure, particularly those that base their critiques on the importance of risk and uncertainty or the unrepresentativeness of existing QALY valuations. This is not the place, however, in which to detail these criticisms; merely to highlight that, were the health maximization view of equity (in its present form) to be adopted, considerable effort would have to be expended in making the QALY metric more sensitive and robust as a measure of health status. In any case the equity account does not pressupose acceptance of QALY's as a measure of health outcome. That its proponents have pointed to the measure as a suitable way forward is purely coincidental; what is really at stake is the idea of health *maximization* as an embodiment of equity concerns.

Opponents of the maximization view have consistently argued that it ignores the *distributional* concerns of public health care systems such as the NHS. Maximizing the sum of individual health states, after all, tells us nothing about the interpersonal distribution of that sum.[10] Culyer (1990) has faced up to this criticism with a largely consequentialist rationalization. He urges the reader to:

> "...allow that the sickest in society are by and large those for whom the marginal product of health care in terms of QALY's is highest, that these are also the poorest, and that when (ceteris paribus) health service per capita rises, the marginal product in terms of health falls... [Then]... it evidently follows that efforts to equalize the geographical distribution of resources, to channel more of them to the sick and more of them to the poor, might be seen not as distributional policies to be justified by equity arguments but efficient policies justified by

[10] The philosophical objection to this sum-ranking approach inherent in QALY maximization has been made most forcibly by Broome (1988), Lockwood (1988) and Harris (1988), in the Bell and Mendus (1988) volume.

health maximization." (p. 55)

This counter-argument is intellectually appealing, not least for the fact that it points to an avenue for incorporating equity and efficiency objectives within a single policy procedure. It depends, however, on essentially empirical judgements which if proved wrong undermine the whole case. In particular there is not much evidence available that shows the marginal product of health care in terms of QALY's to be greater for the poor. Intuitively it may be quite the opposite. Consider the not implausible case of Anthony (who is relatively rich, well educated and well nourished) and Brenda (poor and relatively ignorant of efficient health production methods). Both suffer from the same ailment and both undergo the same treatment. Yet because of his personal and environmental characteristics Anthony is able to better respond to treatment and thus gains a greater number of QALY's. Should health policy then redistribute resources to individuals like him? Clearly few would agree with such a principle, since it implies that Anthony somehow merits the better health improvement because of his socio-economic position. At the most basic level, therefore, health maximization does not seem to embody the equity concerns inherent in the portuguese or most other health care systems.

The reason for unpalatable results arising when simple health maximization is followed is that a unit of 'health' is treated as being of equal value no matter who gets it. Proponents have generally accepted this fact and now suggest that distributive weights are built into outcome measures. If these are correctly applied, in the sense that all the accepted features of distributional equity are incorporated, then there is no technical reason for treating equity independently of efficiency in research or policy formulation. Williams (1988a) has begun to tackle this question by finding out what distributional views are actually held by surveyed individuals. This is a promising line of research, but it is worrying that the utilized survey actively encourages respondents to opt for some type of discrimination as to which groups of people should receive treatment (*eg.* the young, the old, those who have been careful with their health, the deprived, etc.). The justification offered is that scarcity of resources means that discrimination will effectively operate. Despite this, as many as forty per cent of individuals in the pilot-survey opted for none of the discriminations presented - a number far greater than for any one particular type. Such neutrality might well have

been greater had the interviewees been invited to present their own distribution rule. One suspects that these people actually do have some type of rule in mind - for example, that *everyone* should have the *opportunity* to be able to benefit from health care if it will do them some good. Curiously, Williams' analysis is rather dismissive of the non-discrimination argument, opting to ignore the reasons why such a view is adopted by such a large proportion of his sample. This is unfortunate for if the view were widely held the maximizing approach would seem to be on shaky ground, since a great many people are apparently willing to forego efficiency gains in order to assure that individual claims are equally or proportionately satisfied.

This brings us back to criticism of the very idea of health maximization as the prime objective of health systems. Of course, it would take a courageous stretch of the imagination to believe that it should not be an aim or indeed that technically, it is not the best way to proceed. Yet health systems consistently reveal other aims as paramount: the demonstration of caring or the non-exclusion of patient groups, for example. Non-discrimination is certainly an important feature of Portuguese health policy objectives. It may be that the overall aim of medical care is indeed health maximization but that specific provision arrangements (such as an NHS) have adjacent objectives (such as guaranteeing equal access) which take precedence in any eventual trade-off. For the maximizing view to be accepted it must be shown that, generally, suitably weighted QALY maximization will yield results which are harmonious with equity concerns and that unpalatable outcomes are not a significant feature; and ultimately, it must prove its ability to incorporate what are effectively perceived as rights within the outcome measure.

Wagstaff (1991) has argued that equity considerations might be incorporated into the health maximization approach by means of an appropriately specified Social Welfare Function (SWF). Essentially, it is proposed that resource allocation decisions be based on maximizing community health weighted by a parameter describing society's aversion to health inequality. This is a promising area for future research. However, Wagstaff's approach is not free of drawbacks. For example, it assumes that society preoccupies itself with differences in health status *per se*, whereas what generally motivates concern is the fact that such inequalities are systematically related to socio-economic status.

Furthermore, the concern for inequality in Wagstaff's SWF derives from the process of aggregating health levels, whereas most modern societies reveal an explicit and autonomous concern for distribution. Indeed, a key feature of Wagstaff's argument in the same paper is that health maximization and equity are distinct objectives. There is no doubt, however, that Wagstaff's contribution is an important one since it shows a way of integrating equity and efficiency measurement. Therefore, I return to it in Chapter 8, as a point of departure for rationalizing a family of 'health social welfare' indices that embody autonomous aggregative and distributional goals.

More recent work by Culyer (*eg.* Culyer and Wagstaff, 1993; Culyer *et al*, 1992) implicitly acknowledges that health maximization is not in fact an equity criterion and that its proper domain is in the analysis of aggregative (as opposed to distributional) goals. In this work, the ultimate focal variable of equity concern continues to be *health*, with equality in that attribute being defended as the most appropriate criterion. The authors do not dismiss equity objectives in the space of *health care*, but argue that they must be seen as derived from the more ultimate concern. As will be seen, the approach presented in the following section has some similarities with this view, and in terms of manageable empirical analysis has more or less the same implications.

3.5 Capabilities and equity in health

This section argues that, relative to the work reviewed previously, a more promising formulation of equity concerns in the field of health is to be found in Sen's concept of equality of capabilities [Sen (1980, 1985, 1992)]. It provides a clarification of the debate on whether resources or welfare should be the object of equitable policy by examining thoughtfully the transmission process from commodities (resources) to final outcomes (welfare) and arguing that it is the capability people have to carry out human functionings (such as being able to work or to enjoy good health) which matters. The approach has been shown to be useful for the study of poverty issues (Sen, 1983) and more generally for the definition and measurement of the standard of living (Sen, 1987a). Here it is suggested that 'capabilities' is a novel way of understanding health equity objectives, how we should go about attaining them and how progress should be

monitored, while at the same time sharing some common themes with mainstream health economics. It thus warrants much closer attention by economists with an interest in the health inequality debate than has hitherto been the case. Furthermore, the approach may be seen to be a faithful reflection of the equity concern inscribed in Portugal's Constitution, which implies that individuals should have equal opportunity to maximize health potential.

It should be noted that the Sen approach is not a necessary condition for the measurement methodologies adopted later in the thesis. Each of the empirical analyses could quite reasonably be derived from other normative foundations, including the statements inscribed in portuguese legislative and policy statements. However, I hope to show that the capabilities approach does provide a coherent intellectual framework for interpreting the health-equity issue and for identifying the entities of ethical interest. Though other normative bases may point in similar directions, particularly when confronted with deficiencies in available data, the Sen framework is a plausible rationale for the empirical analyses reported further on.

3.5.1 The capabilities approach

Sen's rationalization for focusing equity analysis on capabilities actually derives from weaknesses inherent in the Rawlsian and Utilitarian approaches. The first is said to suffer from goods fetishism: a focus on the goods rather than what they can do for people. To take the example of health care, it is generally acknowledged that people do not demand the good in itself, but rather for what it may contribute to health. It is the opportunities it provides for pursuing a healthy life that matters. Utilitarianism is, of course, concerned with what goods do to people but it uses a measure which overly focuses on mental and emotional reactions to those goods. Arguably, non-utility information is equally as important and this requires that the central focus of analysis be on a much wider range of variables which explain what commodities do for people and how people use them to produce human activities.

The argument can be exemplified through Figure 3.2 which shows the chain from goods to utility. On the left hand side is the world of commodities which has been the

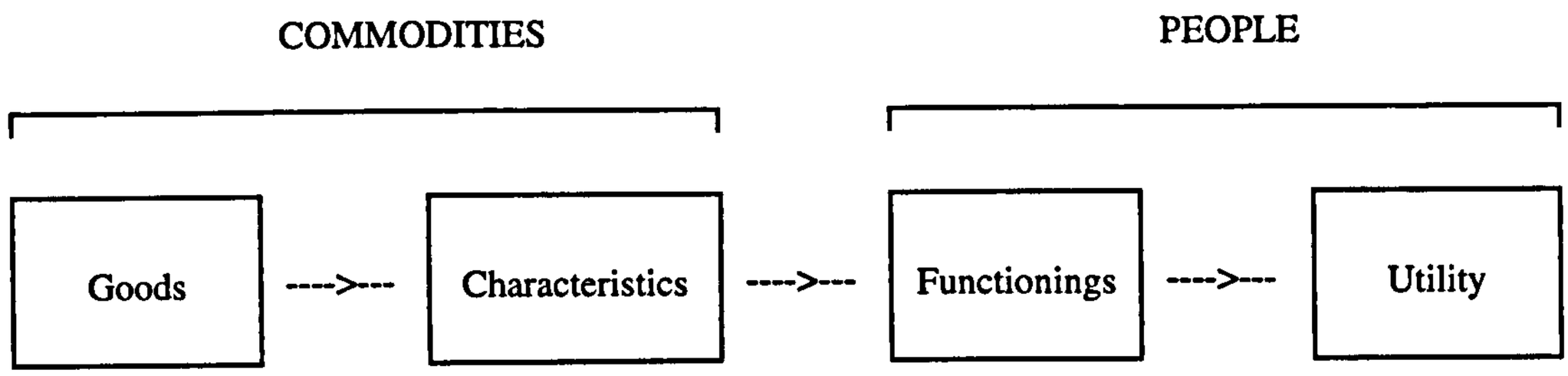


Figure 3.2: The chain from goods to utility
[Source: Sen (1982, p. 30)]

traditional focus of economics when discussing questions of distribution (*eg.* the non-envy approach). These commodities are transformed into more fundamental intermediate products, which Sen in common with Lancaster's (1966) pioneering approach calls characteristics. A focus on characteristics would lead one to interpret the demand for health care as a demand for factors such as clinical efficacy, caring by the GP and so on. Moving to the world of people, how individuals use characteristics of goods to produce human activities is described by Sen as functionings (*eg.* earning one's living, following leisure pursuits, being in good health, etc.).[11]

Most economists would typically regard the link from functionings to utility as unproblematic. Sen disagrees, arguing that although higher levels of utility are associated with better functionings the connection is by no means straightforward. For instance, suppose we were faced with the problem of distributing resources between Ann who despite being physically disabled has an invariably optimistic disposition and Bob, who suffers from no particular ailment, has a high marginal utility of income, but is essentially pessimistic at heart so that in terms of total utility he is actually worse off than Ann. Focusing on utility would lead to a preferential allocation to Bob which does not seem very fair. The reason is, of course, that what most would acknowledge to be Ann's greater needs no-where figure in the analysis. Concentrating on functionings, on the other hand, makes the interpretation of need paramount and allows it to be

---

[11] Culyer (1990) has also defended the reasoning of Sen's approach. He argues that a more useful terminology is to express functionings as being the 'characteristics of people'.

incorporated as non-utility information. There are echoes here of the long standing call by health economists that needs must be seen as instrumental to the accomplishment of a desired end-state (in the example above, being able to move from one place to another) and that the success of health policy should be measured "in terms of changes in individual attributes" (Culyer, 1980). Indeed, Sen's approach may be seen as providing a rationalization for a good deal of work in the health-economics discipline that adopts the so-called "extra-welfare" perspective (Culyer, 1989).

Sen further argues that in terms of equity our interest is less in whether a person is functioning in a certain way and more on whether that individual has the capability to do so. Hence the guiding equity principle being 'equality of capabilities'. A focus on functionings could imply that a person should be continuously using medical services, whereas the crucial idea is that they should be able to, when they choose so to do. This choice element is held to be generally important, although far less so in the universe of *basic* capabilities in which Sen includes such things as abilities to meet one's nutritional requirements, to be clothed and sheltered and, significantly for our purposes, to enjoy good health. In economic terms this suggests that it is the extent of people's opportunity set rather than simply the point in it that happens to be chosen which is important. Thus, there are similarities with Le Grand's equity account, the significant distinction being that whereas Le Grand defines choices over goods, the present approach emphasises what the goods can do for people and what people are able to do with them.[12]

Sen's formulation also has much in common with models of household production which derive from Becker (1965). This work has considered the importance of inputs such as time and environmental constraints in the production of fundamental utility-yielding commodities. In Grossman's (1972) model of the demand for health individuals produce durable health capital which may be accumulated and at the same time may require maintenance through investment in non-genetic human characteristics and the characteristics of goods. The fundamental commodities produced by households

---

[12] It may be noted that Sen's framework clarifies the weakness highlighted in Le Grand's diagrammatic exposition. Since health is best seen as a functioning, commodities or their characteristics are what are required to alter that functioning should its distribution be judged inequitable.

- being able to work, to enjoy life, etc. - coincide for all intents and purposes with Sen's functionings. In household production models fundamental commodities are produced from market goods, environmental inputs and personal characteristics, which are effectively the source of the capability set. What distinguishes the two approaches is that writers in the human capital tradition would not normally consider the link from functionings to utility as problematic. Sen on the other hand, emphasises that any two individuals, or the same individual at different times, may make identical choices when faced with the same capability set and yet may experience quite different utility levels. Therefore our focus should be on the capability set. Muellbauer (1987, p. 47) argues that in empirical analysis this problem is not unsurmountable : "What is important is that the relationships determining the capability set are relatively universal and that the determining variables and the chosen functionings are relatively observable".

3.5.2 Equality of health capability

A clearer impression of what the capabilities framework implies for equity in health can be gathered with the help of some notation and specification. Assume the partial equilibrium space of the particular functioning "good health" and consider

x_i = the vector of health-related commodities possessed by the representative individual i,

$c(*)$ = a function converting a commodity vector into a vector of characteristics of those commodities,

$f_i(*)$ = a production possibility function transforming characteristics into health functioning,

F_i = the set of production functions f_i, any one of which may be chosen by person i.

If i chooses $f_i(*)$, then given x_i the achieved health state is given by h_i,

$$h_i = f_i[c(x_i)]. \tag{3.1}$$

Equation (3.1) provides an indication of the individuals health status. Similarly, a

summation of all h_i's would provide the community health index. I assume for simplicity that all dimensions of health status can be conveniently valued through a scalar measure.

For a given commodity vector $\boldsymbol{x}_i$, feasible functionings are defined by the set $P_i(\boldsymbol{x}_i)$,

$$P_i(\boldsymbol{x}_i) = \{\ h_i \mid h_i = f_i[c(\boldsymbol{x}_i)],\ \text{for some } f_i(*) \in F_i\ \} \tag{3.2}$$

If the person's choice of commodity vectors is restricted to the set X_i, then his/her functionings are given by the set $Q_i(X_i)$,

$$Q_i(X_i) = \{\ h_i \mid h_i = f_i[c(\boldsymbol{x}_i)],\ \text{for some } f_i(*) \in F_i \text{ and for some } \boldsymbol{x}_i \in X_i\ \} \tag{3.3}$$

Equation (3.3) represents the freedom or opportunity that a person has in choosing the functioning 'good health', given his/her personal features, F_i , and command over commodities, X_i. Therefore, Q_i may be interpreted as the capability of person i to generate the desired function. If one is interested in the equity of a particular distribution the task for positive analysis becomes the identification of set Q_i compared with Q_j, the capability set of person j. If $Q_i = Q_j$ then a situation is considered equitable.

Figure 3.3 presents a simplified diagrammatic framework of the arguments above which provides further insight into the question of equity in the domain of health.[13] In quadrant I, h and g are, respectively, the levels of health functioning and of another functioning.[14] The frontier OAA portrays the *capability set*, which shows the various levels of feasible functioning available to the individual. Assuming that he does the best that he can for himself, and given own preferences depicted by the functioning indifference curve (IC), point h^* will be chosen as the equilibrium level of health. This

---

[13] The framework has obvious similarities with Wagstaff's (1986a) exposition of the Grossman model. As noted above, Sen's approach is not unlike that of household production models.

[14] The functioning measured by g may be viewed as a composite of all valued beings and doings other than health (*eg.* being well nourished, being well educated, etc.).

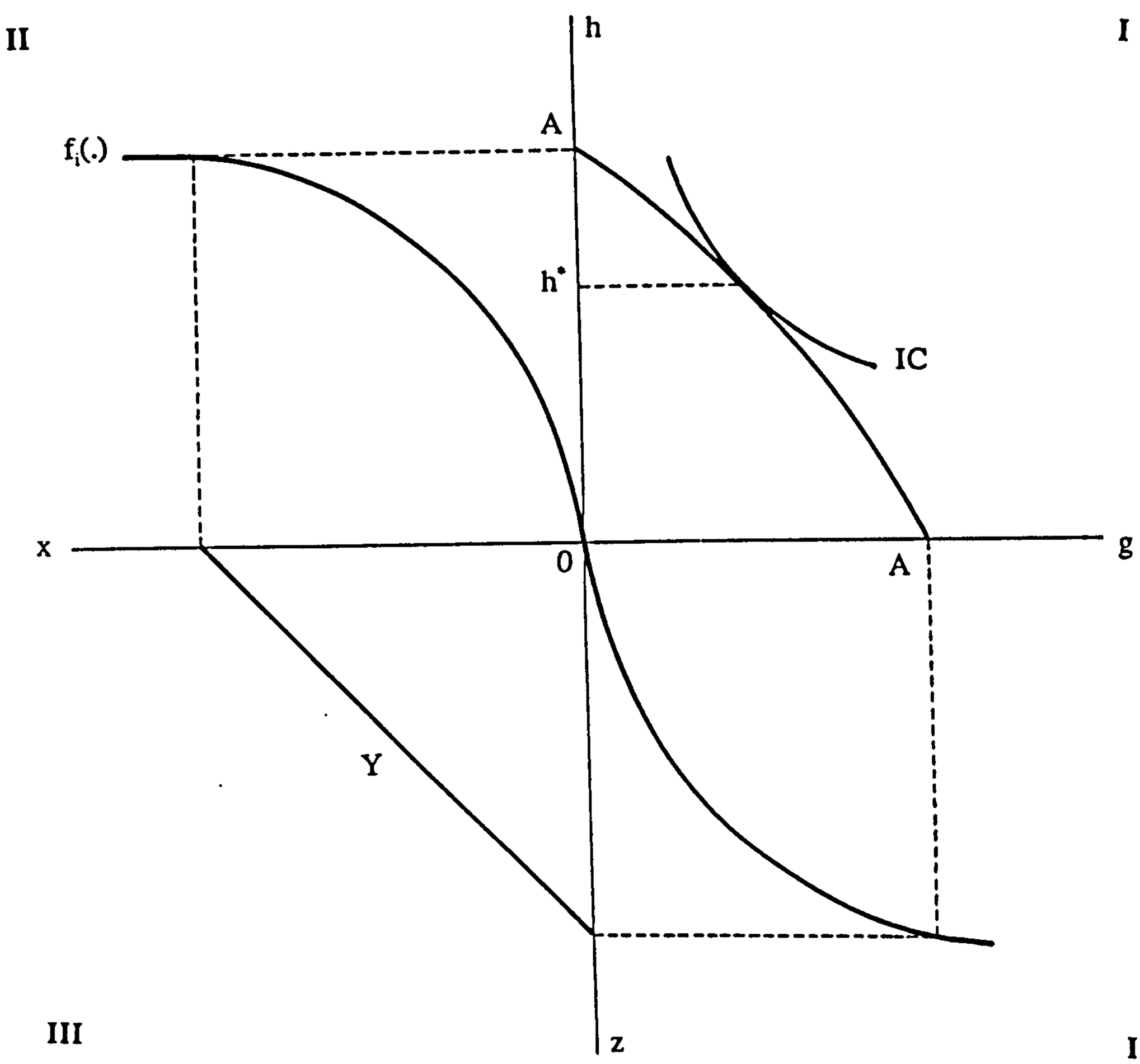


Figure 3.3: A simple diagrammatic framework of the capabilities approach

particular value, and indeed all feasible levels of health status, are however, dependent on the constraints shown in quadrants II to IV. Quadrant II depicts a production relationship between health investment goods (x), such as medical care, food, heating or education and the level of health functioning (h). The curve f_i *(.)*, is the transformation function described above (avoiding for simplicity the intermediate category of characteristics of goods). It shows how much health can be obtained from a given quantity of health related goods inputs for a given set of personal characteristics (*eg.* age, sex, family size, etc.). The range of points along the f_i *(.)* function that can be chosen will depend on the budget constraint (Y) depicted in quadrant III. This shows the various attainable combinations of health related goods (x) and other goods (z) input to the production of the composite functioning (g); and may be viewed as representing the extent of choice over commodity bundles that is available to the individual. Finally, quadrant IV shows the production relationship for the composite functioning, drawn to reflect decreasing marginal returns of goods input, as in the case of f_i *(.)*.

This diagrammatic framework permits the establishment of *ceteris paribus* predictions with regard to equity in the domain of health. In Figure 3.4 a further individual, B, is now introduced. She is identical in every respect to the one considered previously, except for the fact of being older. It is conjectured that this makes her less efficient in producing health functioning from a given level of commodity ownership. The health functioning production function is now given by f_j *(.)*, which will mean that the capability set is restricted to OBA. Therefore, B's opportunity for achieving a desired level of health is necessarily reduced when compared to that of the first individual. Given identical indifference maps, the best that she can do for herself is an equilibrium level of health status below h^*.

Consider next the case of a third individual, C, whose only difference in relation to A is that his income is lower. The possible combinations of x and z commodities that he can attain are given by the constraint Y'. Given the f_i *(.)* function, the range of feasible functionings is reduced to the set OCC. Naturally, identical indifference maps will once again imply an equilibrium level of health status below h*. Thus, the extent of health capability (and the actual level of functioning) is reduced by virtue of differences in commodity ownership.

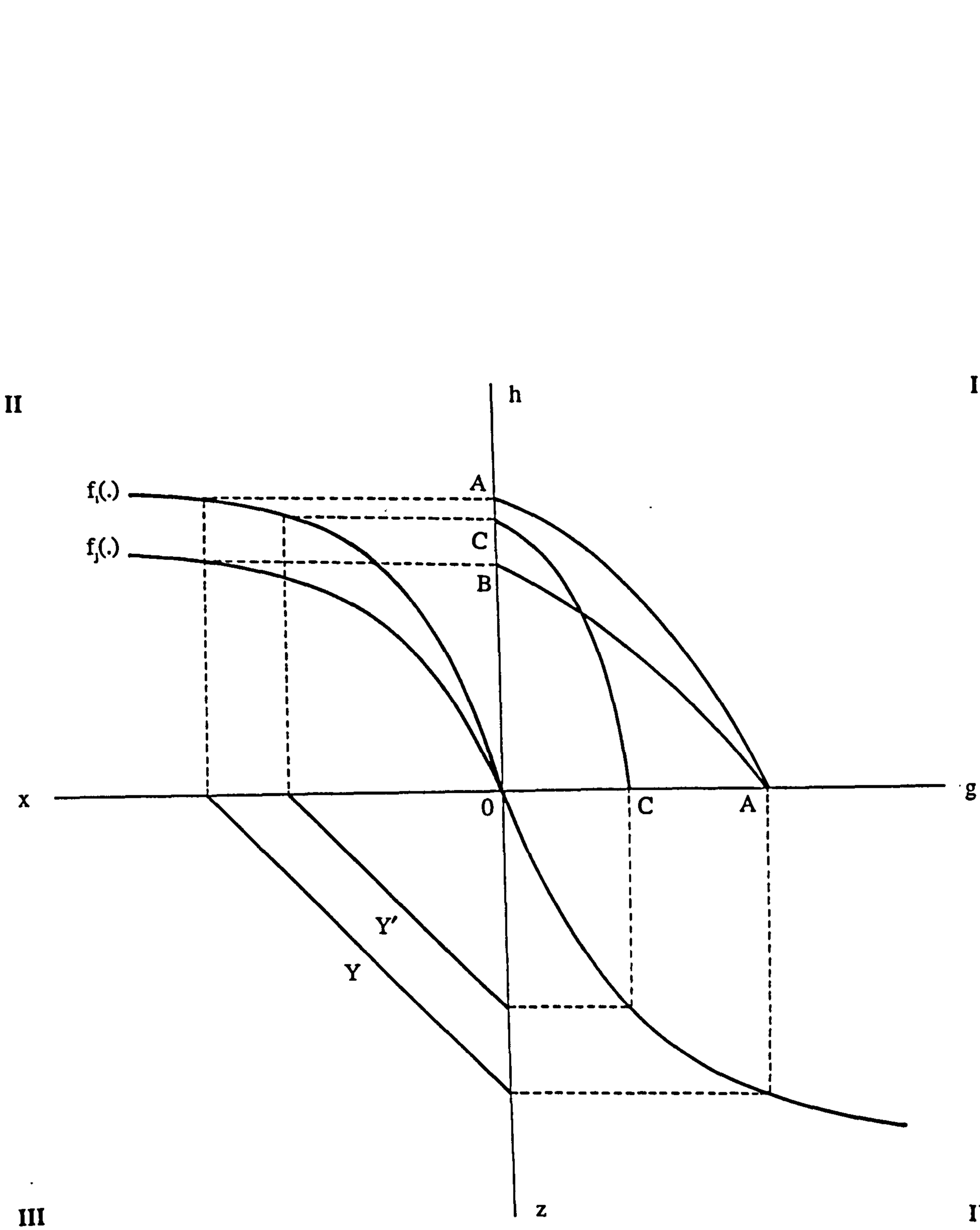


Figure 3.4: The effect of changes in determining variables on the extent of health capability and level of health functioning

There are two points worth noting about this analysis with implications for what is to follow. First, capability is defined in terms of the *same* focal variables as functionings. This means that measurement of health capability might, as a second-best solution, concentrate on chosen or observed functionings. There is some information loss but the evaluative space is the same. Moreover, if freedom (or opportunity) in the domain of health is judged as having only *instrumental* importance for a person's well-being and no *intrinsic* relevance (as seems reasonable), then the value of the chosen functioning provides a suitable basis for examining inequality in health capability.

The second point is that differences in health capability stem from a variety of sources, some of which may be of more ethical interest than others. For instance, the contraction in health capability brought about by person B's age may not be viewed as important as C's reduced opportunity to acquire the means for promoting health functioning. Therefore, judgements on the capability enjoyed by different persons need to be complemented with empirical information from evaluative spaces other than that of functionings. One needs information, for example, on the extent of choice in the space of commodities, on the disposition of health related goods and on the elements that determine the production functions.

3.6 Measurement implications of the capabilities approach

Although the capabilities approach has attracted a good deal of discussion concerning its underlying ethical structure[15], the implications it holds for measurement have so far remained largely untapped. In particular, there have been no studies that concentrate specifically on the capability for healthy living nor any attempts to operationalize the distinctive equity arguments which the approach entails.[16] Ideally,

---

[15] See, for example, Basu (1987), Williams (1987), Arneson (1989), Cohen (1990), Daniels (1990) and Sugden (1993).

[16] Researchers have instead confined application to measurement of the overall living standard. This has been viewed as strictly comprising direct evaluation of the capability set by means of the value of functioning vectors (what Sen calls "elementary evaluation"). For examples see Slottje (1991), Schokkaert and Van Ootegen (1990) and Kakwani (1993). The capabilities approach also seems to be the guiding theoretical structure behind the UN's attempts to measure 'human development' as a

assessment of equality in health capability should be carried out through a measure that describes both the value of the chosen *element* in a given set and the *freedom* of choice which the set allows. I believe that in empirical analysis this task may prove overly ambitious given the lack of suitable information for all variables in the sets F_i, F_j, X_i, X_j, identified in equation (3.3). However, one may reasonably opt for a partial analysis that concentrates on identifying particular aspects of the capability set. For the valuation to have content it need not necessarily be complete. It would make a great deal more sense to accept a partial analysis than to insist on logical completeness and be left with complete lack of information. As Sen (1992) argues the important point is "to keep the *underlying motivations clearly in view* and to see practical compromises as the best we can do under the circumstances [of data limitations]" (p. 135, *emphasis added*).

One potentially useful strategy that may provide considerable insights into inequality in the capability for healthy living is to separate actual functioning levels from the disposition of health related goods that permit their attainment, and to relate the relevant variables to key factors determining opportunity in health functioning space. This approach, to be followed in the thesis, is not in fact dissimilar to previous work that sets out to measure socio-economic inequalities in health and health care. The difference is that here it is derived from a global ethical theory rather than being guided by a variety of seemingly unrelated egalitarian objectives. As will become clear, some important matters are left unresolved but the approach has the advantage of requiring only commonly available data. One essentially needs access to empirical elements on health status (h), use of health related goods (x), command over resources (Y), and the factors that determine the production function f_i (.), all of which are relatively observable.

A possible objection to this approach is that empirical application is based on point comparisons of realized functionings, rather than comparisons of capability sets. The traditional view in economics is that set evaluation requires measurement of opportunities; whereas point estimates will reflect the interaction of opportunities *and* preferences. By implication, the case for Sen's approach as an ethical rationale for

---

multi-dimensional entity comprising longevity, knowledge and decent living standards [Griffin and Knight (1989); UNDP (1990); Desai (1991)].

health-equity empirical analysis would appear weakened. However, the habitual contrasting of opportunity and preference can often be deceptive (see Sen, 1991). In general, it is hard to see how set comparisons can be independent of the preference ordering over the elements of the respective sets. Information on individual preferences must also be taken into account when evaluating the *freedom to choose*, the principal reason being that simple measurement of the extent of choice says nothing of the value individuals attach to the elements included in the set. From this perspective, concentrating on realized functionings would appear to provide useful information to the underlying objective of health capability comparisons. This does not mean that there is no information loss. Measuring the extent of choice in an adequate manner is obviously a key question for health-equity research. However, a partial indicative answer is provided in the thesis (based on readily available data), by virtue of the proposed two-step approach to measurement.

The two tasks suggested for empirical analysis will be referred to as evaluating *achievement* and *advantage*. The measurement of *achievement* concerns the actual level of the functioning 'good health' that a particular person obtains. Measuring *advantage*, on the other hand, involves examining the experience of individuals in the space of commodities that are instrumental to the attainment of the desired functioning. Though the former is clearly the more important attribute from the point of view of Sen's approach, measurement of *advantage* cannot be overlooked. Commodities are what is required to enable people to achieve their health potential. If health care is considered important in this respect, then it is natural - and indeed necessary - to enquire about its distribution.[17]

The analysis will concentrate on inequality related to the extent of choice over commodity bundles. As the diagrammatic framework of 3.5.2 suggests, this is a crucial contributory factor to opportunity in health functioning space. Obviously, choice over commodities is not the only factor affecting the capability set; production possibilities also constrain individual freedom to achieve desired levels of health functioning.

---

[17] Culyer, van Doorslaer and Wagstaff (1992) make basically the same point when they write: "... it is the distribution of health which is ultimately of interest. But the means by which the desired distribution of health is to be achieved is health *care*. Thus one *can* talk about an equitable distribution of health care: it is one that gives rise to an equitable distribution of health" (p. 209).

However, it is also the case that personal characteristics affecting production functions tend to be less amenable to policy intervention. Moreover, as shown in Chapter 2, the notion that health and health care distribution should not depend on command over resources underlines equity concerns in portuguese policy statements. Therefore, it seems useful to measure the degree to which relevant health domain variables are related to economic position. Obviously, one needs to control at the same time for factors that determine the configuration of health production function (*eg.* age and sex). If this is done, a partial, but significant, view of inequality in health capability becomes available.[18]

The primary empirical task in evaluating *achievement* will be to measure the distribution of health in relation to potential command over resources. Further conceptual analysis relative to this issue is carried out in Chapter 5 with actual empirical application being addressed in Chapters 6 and 7. Some empirical innovations are introduced such as the explicit incorporation of varying social judgements concerning the degree of equality preference. However, the overall procedure is the same as that adopted by many studies that measure socio-economic inequalities in health. This means that the results may easily be compared. It is important to note, however, that previous efforts rarely provide an ethical rationale for measuring inequality in the space of health, leaving them open to criticism that equality of health outcomes is not an objective of health systems (*eg.* Mooney *et al*, 1991). The capabilities framework provides this foundation, suggesting that measurement of health inequality associated with command over resources is a partial (but nevertheless important) evaluation of people's opportunity to realize health potential.

The above task does, however, place total emphasis on the question of distribution, disregarding aggregative objectives (*eg.* maximizing the community's health). It is difficult to see how a satisfactory evaluation of *achievement* can be made without contemplating its actual size. Consider, for example, communities A and B, each made

---

[18] One important element in which the analysis is partial is that a somewhat restrictive view is taken of the "health production function". I shall control only for demographic factors ignoring, for example, the role of education in enhancing production possibilities. This limitation is acknowledged in the work reported in Appendix 1 where a comprehensive model of health related behaviour is developed.

up of two persons. Admit also that health states are cardinally measurable and that they are distributed in the following manner: A = {4, 6}, B = {7, 15}. Community B reveals the greater inequality and yet both its citizens are in better health than those in community A. What this example shows is, of course, that even if our overriding concern is equity-motivated, evaluation of *achievement* based solely on distributional assessments is likely to be incomplete. What is needed is a measure that balances aggregative goals with distributional considerations. An attempt at designing such a measure is reported in Chapter 8. The approach is described as *evaluating the social welfare associated with health*, and unsurprisingly, is similar to the economic concept of social welfare. Just as social welfare is a representation of the 'goodness' of the social-state, so health social welfare embodies the 'goodness' of a community's health state, defined over the two major aims of policy, equity and efficiency. The Chapter is primarily aimed at laying the foundations of future research, but empirical illustrations drawing on portuguese data are also presented.

Logically, measurement of *advantage* should consider all goods that are instrumental in realizing health potential (*eg.* housing, nutrition, social services, etc.). The focus of the thesis is, however, on one element - health care. The obvious objection is that other goods in the x-commodity vector might have a relatively higher marginal impact on health achievement. But in public policy, there is no doubt that health care is generally seen as the single-most important instrumental variable affecting health. Thus, policy objectives invariably relate simply to the distribution of health care (recall once again Chapter 2). In this context, it seems appropriate to pay particular attention to this one commodity.

As with achievement, evaluation of *advantage* can also be viewed in an absolute or relative sense, but only the latter need concern us here. The one aspect in which the absolute level could be important is if we were to concentrate on the size of commodity bundles required to attain a given level of functioning. Clearly, even then some element of distributional analysis would be required as opposed to mere evaluation of the overall size. Such an approach does, however, imply the specification of a 'basic minimum' of commodity provision, which, as we have seen, is a highly problematic time and culture contingent concept. It may be that for some health producing commodities, such

as housing, a basic standards approach is appropriate, but for others, such as health care, clearly not. It is fair to expect that people should have as much health care as will do them some good (Culyer, 1991). The implication, in terms of evaluating *advantage*, is that emphasis should be wholly on distributional issues and furthermore, that the entire range of the distribution should be examined. Consequently our interest will be directed at the differences that are observed between people with unequal command over resources as regards their experience of health care goods.

There are two related aspects of interest in evaluating *advantage* in health care commodity space: the use people make of such care and the payments that go to finance it. Asymmetries in health care utilization are seen, under the Sen approach, as contributory factors to gaps in health capability. For any two given individuals presenting identical personal circumstances, the higher level of health care use by one is judged to confer upon him a wider opportunity for choosing a desired level of health functioning. This suggests that comparisons between persons with different income constraints must be calibrated by relevant information on personal characteristics, such as age, sex and initial health status. I shall not be addressing the question of unequal utilization in empirical analysis, but will instead concentrate on conceptual issues related to measurement indices (Chapter 9).[19] The discussion is relevant to previous research which focuses on the problem of inequity in the delivery of care (*eg.* Le Grand, 1978; Wagstaff *et al*, 1991b). It should be stressed, however, that the rationale for concentrating on health care use is distinct from that of previous work, where the general motivation has been to measure the extent of horizontal inequity (*ie.* the degree to which there is unequal treatment for equal need). In the context of the capabilities approach, unequal utilization is important in so far as it implies unequal advantage in attaining potential improvements in health status.

The empirical interest in health care finance is implied by the concern for inequity in health care utilization. It is possible that a person's advantage in terms of the use

---

[19] An application based on some of the indices discussed may be found in Pereira and Pinto (1993), which used what is still the most recent nationally representative portuguese data source. The current context of portuguese health policy suggests that there is less interest in this question than in that of inequity in health care finance. Given contingencies of space, I choose to concentrate on the latter issue.

they make of health care goods is itself determined by the payment mechanism. Therefore, inequity in the distribution of payments should also be examined. A possible objection to this approach is that the focus ought properly to be upon all those variables that affect the demand for care and not simply the method of payment. However, it is also the case that a more aggregated approach that concentrates on the relation between ability to pay and health care payments can provide useful insights into inequities in health capability. The more an individual pays towards health care the less she has available to spend on other health enhancing commodities, thereby limiting the scope for potential health improvements. Consequently, fairness in terms of opportunity for health functioning should include a *desideratum* that payments in the space of x-commodities (of which health care is one) be related to ability to pay. The most obvious means for measuring such an objective is by evaluating the progressivity of different health care financing mechanisms (*cf.* van Doorslaer *et al*, 1993). The methods for addressing this task are discussed in the next chapter with actual empirical application being carried out in Chapter 7.

As I have stressed, the empirical agenda described above can only be seen as a partial evaluation of the capability for healthy living enjoyed by different persons. One aspect in which the proposed analysis is silent is in the *process* whereby goods inputs are transformed into health status levels. Distributional assessments of instrumental commodities provide useful insights into capability, but ultimately one must consider the *impact* of commodity use on actual achievements. This inevitably involves the specification and testing of a model of health behaviour. The similarities between the capabilities approach and models of household production suggest that the Grossman (1972) model might be a useful means of accomplishing this task. I therefore explore this question in Appendix 1 with a view to clearing ground for future empirical work.

3.7 Conclusions

In this chapter I have reviewed normative aspects of the health inequality debate which only very recently have begun to be addressed by economists and other social scientists. This is despite wide agreement that grasping their implications is a

prerequisite for understanding why people are concerned about inequity, how it should be measured, what causes it and how policies may be formulated and monitored. Although I have sought to identify economic conceptions of equity which reflect the objectives of Portugal's health system, the discussion is also relevant to other countries.

Given that no consensus definition of equity has emerged in the literature the most immediate task has been to critically review available conceptions in terms of their suitability for informing the health-equity debate. A set of assessment criteria were previously laid out in this respect. Although all the distribution rules are easily comprehensible there is no doubt that some do not take us very far in terms of clear policy solutions. With regard to specificity and rigour, there is clearly much room for further development at the analytical level. Since ultimately our interest in different equity formulations is based on their potential for empirical analysis it is disheartening to find that, in the health domain, many of the approaches either have ambiguous implications or simply suggest excessive information requirements. Finally, most of the approaches tend to go against the views put forward in the portuguese policy statements. It clearly makes no sense to adopt a particular health-equity formulation if there is no widespread support for its ethical base.

Sen's capabilities framework comes closer to achieving all previously specified standards. This does not mean that important insights cannot be extracted from the other formulations or that these are irrelevant for the equity in health debate. The Le Grand approach, once developed, may come to be seen as analogous to the capabilities approach. The crucial difference between them is that the choice sets are defined over commodities and capabilities, respectively. In empirical analysis, this distinction may be of little importance. Further refinement of the health maximization approach could also yield results of great significance. For the moment, however, it is not at all clear that it can serve as a suitable guide to empirical evaluation of health-equity objectives. It seems to be largely uninformative with regard to the comparison of interpersonal feasibility of attaining rewards and does not focus on the processes of health production and health care delivery which policy objectives emphasise. Thus, although it has the logical advantage of treating efficiency and equity attainment in a single policy step it is silent on particular aspects of the equity problem. Sen's framework has most obvious

advantages in this respect and consequently appears the more fruitful as a basis for positive analysis of specifically equitable objectives.

Admittedly, the Sen approach leaves many issues unresolved. It provides an ethical justification for concentrating measurement of inequity on a particular evaluative space (*ie.* positive freedom to achieve health) but acknowledges that data limitations may effectively restrict the focus of analysis to partial elements of that space (*eg.* what health levels are actually achieved). This limitation means that various avenues for positive analysis can be justified under the ethical base. Indeed the chosen empirical strategy that emerges does not adopt the obvious route of conceiving and comparing unique representations of health capability sets. This may raise the objection that the Sen framework does not meet two of the assessment criteria laid out (*viz.* that a distribution rule should be specific and rigorous and that it should be readily susceptible to empirical documentation). Such an argument is not wholly convincing. The Sen framework when applied to health is quite clear on what constitutes an equitable distribution (equality of health capability); however, the limits imposed by practical calculations mean that only a partial answer may be given in empirical analysis. The important point is that 'capabilities' provides the global ethical rationale that appears to be missing from a good deal of previous research on equity in health and health care. Empirical work in the field is commonly guided by a variety of unrelated egalitarian objectives. The present analysis suggests that the common focus on issues like economic inequalities in health and health care utilization, though partly based on the unavailability of more refined data, may not be as misconstrued as some authors have suggested.

Part 3

METHODOLOGY

Chapter 4

The Measurement of Inequity in Health and Health Care

"What is just, then, is what is proportionate, and what is unjust is what is counter-proportionate."
Aristotle, Nicomachean Ethics

4.1 Introduction

The previous chapter proposed a new specification of equity in the domain of health, based on Sen's capabilities approach. Two separate issues were identified as the correct focus for positive analysis: the actual level of the functioning 'good health' that particular individuals obtain, described as their *achievement*; and the experience of those persons with regard to commodities that are instrumental to attainment of the desired functioning, their *advantage.* The framework entails various specific empirical tasks, two of which are addressed in the present thesis: (i) measurement of inequity in health, and (ii) measurement of inequity in the payment of health care goods. The present chapter provides the conceptual means by which these tasks are to be accomplished. The remaining empirical tasks suggested by the Sen framework are only examined at the conceptual level, and hence their discussion is postponed until Part 5. They are: (i) measurement of the level of social welfare associated with the health distribution (Chapter 8); and (ii) measurement of inequity in the utilization of health care goods (Chapter 9).

The approach to measurement adopted in the present case is based on the application of *cardinal indices of economic inequality* to the problem of inequity in health and health care. Although this is also the route taken by recent economic work on the subject, it is not uncontroversial. One plausible objection is that a focus on cardinal measures carries an implication that complete and unambiguous rankings of

distributions can always be made. This argument suggests that it might be more appropriate to concentrate on ordinal measures (*eg.* the Lorenz dominance criterion) when studying inequity in health and health care. It is well known that such measures are based on a smaller set of underlying assumptions likely to command wider societal agreement. However, the ordinal method also has a drawback: it excludes the possibility of evaluating the *degree* of inequality, a feature that is of fundamental interest for policy-making. Adequate monitoring of inequitable situations requires information which does more than simply rank distributions; it should also show how much more or less inequitable one situation is in relation to another. For this reason, the cardinal approach is accepted in the thesis as providing the most informative route to measurement. However, it is also recognised that its evaluative potential is increased if (a) one is able to choose indices that make explicit possible differences in social judgements (*eg.* a preference for more or less inequality); and (b) the method is complemented by pragmatic comparisons of graphical representations of distributions.

Traditionally, economists have concentrated on the distribution of a single variable, *income*, which is accepted to be a suitable proxy for more ultimate objects of social concern, such as the level of welfare or the opportunity set faced by an individual. Analyses of health domain inequality, on the other hand, are generally agreed to involve measurement of the joint distribution of a health variable and one that describes an individual's command over economic resources, given that it is the systematic relation between the two that is usually at the root of equity concerns. Rather surprisingly, there are few conceptual answers to this type of problem in the literature dealing with economic inequality.[1] This poses the question of whether single-attribute measures developed in that context can be suitably extended to the health inequality debate? Although important recent work suggests that caution should be exercised (*eg.* Coulter *et al*, 1992a), the few studies of multi-dimensional inequality to have emerged (*eg.* Atkinson and Bourguignon, 1982; Jenkins and O'Higgins, 1989) draw on techniques that are essentially extensions of the univariate case. It seems reasonable, therefore, to take

---

[1] The neglect of conceptual analysis of multi-dimensional inequality is exemplified by a recent series of monographic issues of the *Journal of Econometrics*, dealing with inequality measurement (Vol. 42, no. 1, 1989 and Vol. 43, no.'s 1-2, 1990). Of a total of twenty-seven articles, twenty-three are simply dedicated to the univariate distribution of income or wealth.

work on summary measures of income inequality as the starting point for measurement of inequity in health and health care.

The indices chosen for application in empirical analysis are based on the so-called concentration curve approach. The methodology was introduced to health economic research by Wagstaff, van Doorslaer and Paci (1989) under an international project financed by the European Community (the *ECuity study*). The present chapter provides a formal presentation of the measures used by the ECuity group. It uses a continuous random variable framework, which may be seen as useful in two respects: (i) it permits access to mathematical calculus so that normative and statistical properties of the indices can be conveniently determined; and (ii) it helps to distinguish conceptual from data-specific features of the measures suggested. These are aspects in which the ECuity analysis appears incomplete. The method of exposition is also important for establishing the results of Part 5, where further extensions to the concentration curve approach to measuring health domain inequity are reported.

The present chapter also extends the methodological base of the ECuity group by proposing indices that take due account of alternative social judgements concerning the degree of equality preference. This seems to be a crucial element missing from previous analyses of health domain inequity, where researchers have typically used measures that entail what some might regard as unappealing assumptions about the weight to be attached to inequity at different levels of the income distribution (*eg.* the range, Gini coefficients and standard concentration indices). Such measures also implicitly presume a degree of agreement about equity judgements, which is rarely found in practice. By contrast, the 'generalized' measures here introduced can yield numerical estimates of the degree of inequality that are conditional on the choice of a single parameter which reflects different distributional points of view. Thus, the earlier adoption of the capabilities framework does not imply that the positive analysis to be carried out is irrelevant to policy-makers who choose a different equity norm. What it has done is to demarcate the entities of ethical interest; the precise degree of desired distributional equity is still very much open to *explicit* empirical discretion.

The rest of the chapter is organized as follows. In the next section, a selective

review of the more important economic approaches to measurement of inequality in incomes is provided. The purpose of this analysis is to evaluate the relative merits of measures that might be applied to the field of health, while at the same time covering necessary background material to the rest of the thesis. Section 4.3 describes the ECuity approach to measuring health and health care finance inequity, which is based on standard concentration curves of health and health care population shares ranked by income. Section 4.4 extends the scope of the ECuity measures through the mechanism of *generalized concentration indices*. Section 4.5 discusses the empirical procedures used in the thesis to compute summary measures. The main interest of this analysis is in the methods for estimating the 'generalized' indices, since these have not been applied elsewhere in the literature. However, the section also presents results on the bias involved in computing measures from grouped data. These are useful for interpreting later empirical results and have implications for other studies. The final section summarizes and concludes the analysis.

4.2 Income-based indices of inequality[2]

Broadly speaking, cardinal indices of income or income-related inequality can be divided into three main groups: (i) *normative* measures that are derived directly from specific social welfare functions; (ii) measures that are founded on *axiomatic* consideration of the desirable properties which an inequality index ought to have ; and (iii) measures that are based on the slopes and ordinates of *Lorenz* and *concentration* curves. The applications in this thesis are drawn from the third approach. It is useful, nevertheless, to begin by establishing the reasons why so-called normative and axiomatic measures are not applied. In the first case, this is done with reference to the widely used Atkinson family of indices, which, paradoxically, can also be seen as a precursor in the literature to the parametric statistical measures to be employed. The discussion is also relevant to the analysis of Chapter 8.

---

[2] The discussion of income-based inequality indices presented below is necessarily selective. For greater detail the reader is referred to the numerous reviews of the income inequality literature which are available. Recent examples are Lambert (1989), Chakravarty (1990) and Jenkins (1991).

Atkinson (1970) criticized conventional measures of inequality, such as the coefficient of variation and the Gini index, on the grounds that these are simply statistical devices that measure the relative dispersion of a frequency distribution without reference to a normative notion of social welfare. He also showed that the implicit weights which these measures attach to transfers at different levels of the income distribution are unlikely to command wide support. As an alternative, Atkinson proposed a family of indices that is derived directly from a social welfare function (SWF) that incorporates what are judged to be desirable social values. He assumed that the SWF can be defined in the additive separable utilitarian tradition of Dalton (1920).[3] Given such a formulation, Atkinson introduced the concept of the *equally distributed equivalent income* y_e, which is the amount of income, if received by every individual, would result in the same level of social welfare as that given by the observed income distribution. Atkinson's index of inequality is then defined as:

$$A = 1 - \frac{y_e}{\mu_y} \tag{4.1}$$

which can be interpreted as the share of mean income (μ_y) that is "wasted" as a consequence of inequality. Given that the index ranges from zero to one, if A = 0.2 it is implied that only 80% of the present total income is required to achieve the same level of social welfare.[4] However, for the index to be mean-independent it is necessary to place a further restriction of constant elasticity on the U(·) functions. In these circumstances, the Atkinson family of indices becomes:

---

[3] Specifically, the SWF is given by

$$W = \sum_{i=1}^{n} U(y_i)$$

where $U(y_i)$ is the utility derived by an individual with income y_i. It incorporates the following assumptions: (i) *symmetry* - everyone's incomes are evaluated using the same U(·) function; (ii) *non-decreasing welfare* - an increase in one person's income must improve society's welfare or at least leave it unchanged; (iii) *additive separability* - implying that changes in social welfare arising from changes in one individual's income are entirely unaffected by those of others; and (iv) *strict concavity* - the welfare weight always decreases as income increases, implying a preference for equality. Subsequent research has shown that similar results can be obtained by relaxing assumptions (iii) and (iv) [See, *eg.* Dasgupta *et al*, 1973)].

[4] Alternatively, a maximum of 20% of current total income would need to be sacrificed to achieve equality.

$$A(\epsilon) = 1 - \left[\frac{1}{n} \sum_{i=1}^{n} \left(\frac{y_i}{\mu_y} \right)^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} \qquad \epsilon \neq 1, \quad \epsilon \geq 0$$

$$= 1 - \prod_{i=1}^{n} \left(\frac{y_i}{\mu_y} \right)^{\frac{1}{n}} \qquad \epsilon = 1 \qquad (4.2)$$

where ε is a parameter measuring the degree of inequality aversion. This parameter can be given any value in the range zero to infinity, with the two limiting cases respectively implying indifference to inequality and the Rawlsian criterion. Hence, as ε rises, greater emphasis is attached to transfers at the lower end of the distribution relative to those at the top.

In terms of policy prescription, there is much appeal in tracing out different degrees of inequality aversion by varying a single parameter. Atkinson's method has, therefore become particularly influential, setting the research agenda for the normative approach to inequality measurement and also challenging those that adopt different methods to develop measures that reflect alternative distributional judgements. However, it has also come in for a certain amount of criticism (*eg.* Meade, 1976; Sen, 1978). It is argued that what enables the set of derived indices to be greatly reduced is simply the arbitrary form of the SWF. Important issues, such as the extra-welfare characteristics of individuals or their concern for the position of others in the income distribution, are neglected by the additive separable utilitarian SWF. Equally damaging is the argument that Atkinson's approach does not measure inequality in any sense. It is actually a measure of inefficiency or utility-loss arising from a less-than-optimal distribution of the available income. A reduction in ε, for example, decreases measured inequality even though the gap in individual utilities might be increasing, which is a result of those on higher incomes producing utility relatively more efficiently.

The upshot of these arguments for the present thesis is most obvious. Adopting Atkinson-type indices would contradict two fundamental arguments of the normative analysis in Chapter 3: that extra-welfare information must be brought to bear on social decisions and that equity analysis should involve actual measurement of the relative differences in the variable of interest. Consequently, an approach which involves direct

consideration of the relevance and practicality of particular indices seems better attuned to our needs.[5]

The second approach to developing indices of income inequality consists of specifying at the outset a number of desirable properties that an index ought to have, and then to use these to characterize the index. Five axioms, in particular, have gathered wide assent. If inequality is given by a continuous and twice differentiable function $I = \psi(\mathbf{y})$, where $\mathbf{y} = y_1, ..., y_n$, these properties can be stated as follows:

(i) *Population homogeneity (replication).* $\psi(\mathbf{y}, \mathbf{y}, ..., \mathbf{y}; r\ n) = \psi(\mathbf{y}; n)$ for any arbitrary positive integer *r*.

(ii) *Scale-independence.* $\psi(\lambda y_1, ..., \lambda y_n) = \psi(y_1, ..., y_n)$ for any arbitrary positive scalar λ.

(iii) *Symmetry (anonymity).* $\psi(\mathbf{y}) = \psi(\sigma\mathbf{y})$ for all permutations σ.

(iv) *Principle of Transfers.* ψ is Schur-convex.

(v) *Additive decomposability.* There exists Ω such that

$$\psi(y_1, ..., y_n) = \Omega[\ \psi(y^1_{(1)}, ..., y^1_{(n)}), ..., \psi(y^G_{(1)}, ..., y^G_{(n)}); n_1, ..., n_G; \mu_{y1}, ..., \mu_{yG}\ ]$$

for any arbitrary partition $g = (1, ...,G)$ of $\{1, ...,n)$.

The first three axioms respectively imply that the inequality index I is entirely unaffected by replications of the population, proportionate changes in the scale of incomes and permutations of income recipients. The fourth, which was initially proposed by Dalton (1920), implies that a progressive transfer of income (*ie.* from a richer to a poorer person) must lessen the degree of inequality.[6] The final axiom requires that the index be capable of being expressed as the weighted sum of the inequality values calculated for specified population subgroups, plus a term based on mean incomes and group size.

---

[5] Obviously, this should not be interpreted as disregard for normative issues. These have been extensively discussed in Chapter 3, leading to identification of the entities and relationships of ethical interest, and will be further analyzed through the application of parametric inequality measures, and through the development, in Chapter 8, of an index of health social welfare.

[6] This axiom is sometimes strengthened with the requirement that the inequality reduction resulting from a progressive transfer be greater the lower the income of the recipient (*transfer sensitivity*).

Cowell (1980) and Shorrocks (1980) have shown that the entire class of indices satisfying these properties can be written in a very simple fashion, namely the *generalized entropy* (GE) family:

$$I_c(y_i) = \frac{1}{n}\frac{1}{c(c-1)}\sum_{i=1}^{n}\left[\left(\frac{y_i}{\mu_y}\right)^c - 1\right] \qquad c \neq 0,\ 1$$

$$I_1(y_i) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{\mu_y}\right)\log\left(\frac{y_i}{\mu_y}\right) \qquad c = 1 \qquad (4.3)$$

$$I_0(y_i) = \frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\mu_y}{y_i}\right) \qquad c = 0$$

This class includes many well-known inequality measures. I_1 and I_0 correspond, respectively, to Theil's (1967) entropy measure and the mean logarithmic deviation, also suggested by the same author. When $c = 2$, I_c is equivalent to half the square of the coefficient of variation. Furthermore, for every member of the Atkinson family of indices there is a corresponding, ordinally equivalent member of the GE family (*viz.* Atkinson indices from ε upwards correspond to GE measures from c=1 downwards). It appears, therefore, that the parameter c can be interpreted as a measure of the degree of inequality aversion, in very much the same way as Atkinson's ε. As c decreases (to negative values) the inequality index becomes more sensitive to transfers at the lower end of the distribution and less weight is attached to transfers at the top.

The GE family of indices has another important property for empirical work: additive decomposability. If there is an interest in the relationship between variables that contribute to income differences and actual inequality then an obvious approach is to decompose the inequality measure by population subgroups (*eg.* age, occupation, region). In such circumstances, the form given by (4.3) is particularly important, because GE indices are the only inequality measures for which a *ceteris paribus* increase in the inequality within a given group always leads to an increase in inequality overall. Without this property it is impossible to attribute total inequality to its component inter- and intra-group inequalities in an unambiguous way (Cowell, 1988). Moreover, the decompositions are achieved with relatively little information, namely the mean incomes of the subgroups, population shares and total subgroup inequality values.

GE indices appear to have some important properties for examination of health domain inequity. The reasons for their not being applied in this study are largely pragmatic and related to the advantages of using the concentration based measures discussed below. First, concentration indices have associated graphical representations which facilitate their interpretation by non-specialists (*eg.* policy makers); whereas "entropy measures are for specialists only and difficult to conceive and to interpret for the many other people interested in redistributive effects" [Pfähler (1987, p. 16)]. Second, concentration indices have an established mechanism for handling income-related health domain inequality, which is the main focus of the study. Third, concentration measures can also be extended to incorporate different degrees of inequality aversion by varying a single parameter. Fourthly, use of GE measures would make comparisons with previous work somewhat uncertain. Given that research by health economists has largely drawn on the concentration curve approach and that I wish to measure the impact of alternative empirical specifications on the degree of measured inequity, it makes sense to focus on indices drawn from that approach. Finally, it may also be noted that the individual welfare functions implied by GE measures are somewhat unappealing. In particular, it is assumed that a person's well-being is independent of the opportunities for choice of others [see Sen (1973, p. 36) and Kakwani (1986, p. 18)]. However, a crucial aspect of equity judgements is the comparison made by specific individuals (or an ethical decision maker on their behalf) of their own position in relation to others'. The implied social welfare functions of concentration measures appear to incorporate precisely such a notion [see, *eg.*, Sen (1974), Pyatt (1980), and the work reported in Chapter 8].

I turn therefore to the third approach to examining economic inequality: the class of Lorenz and concentration measures. The Lorenz curve - which maps the cumulative share of total income accruing to each cumulative share of the population, when incomes are ordered from poorest to richest - is the most widely used graphical representation of income distributions. A convenient formal definition is available by assuming that income Y is a non-negative random variable with probability density function $f(y)$, cumulative distribution function $F(y)$ and mean μ_y. In this case the Lorenz curve is given by:

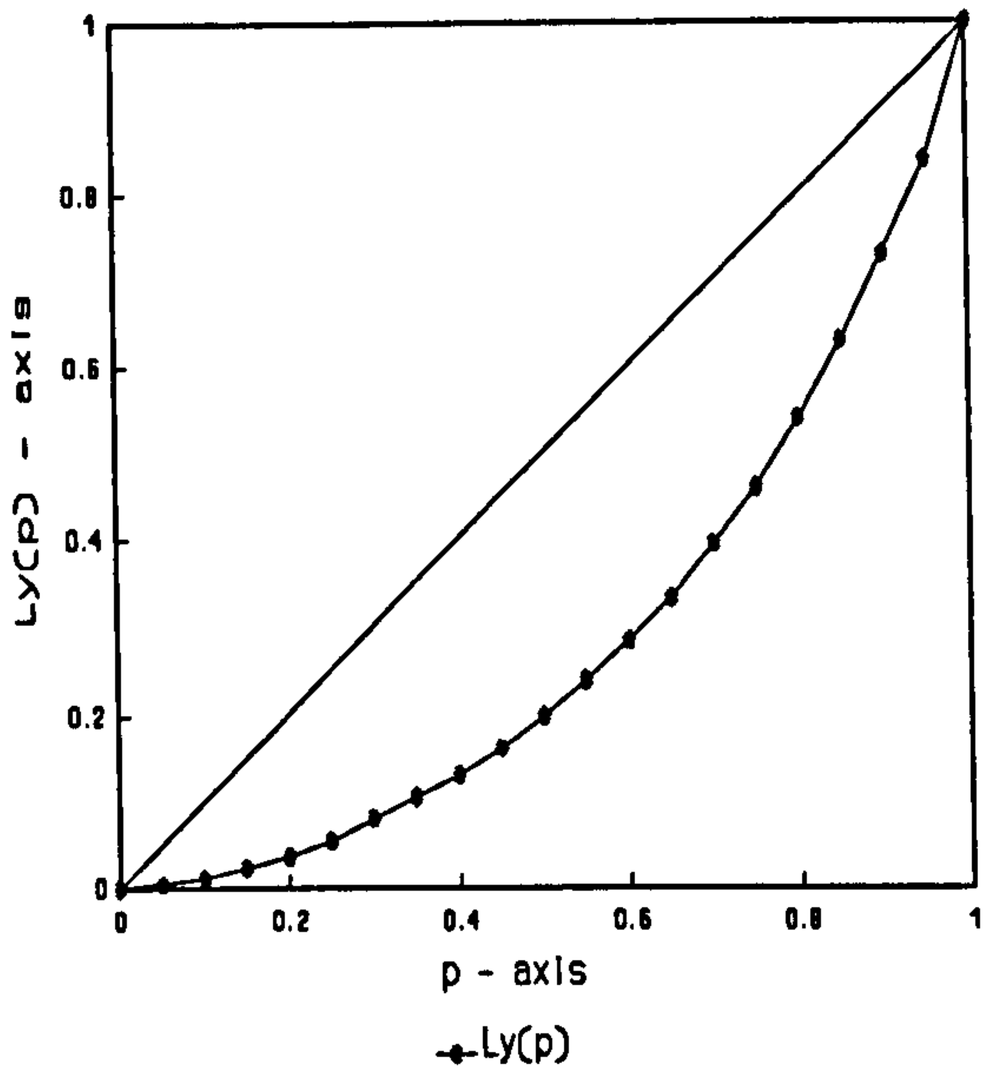


Figure 4.1: The Lorenz curve

$$L_y(p) = \int_0^y \frac{y\, f(y)\, dy}{\mu_y} \qquad 0 < p < 1 \tag{4.4}$$

where $p = F(y) = \int_0^y f(Y)dY$, interpreted as the proportion of income units having an income less than or equal to y; and $L_y(p) = F_1(y) = 1/\mu_y \int_0^y Yf(Y)dY$, interpreted as the proportional share of total income held by units with income less than or equal to y.

In Figure 4.1, the straight line $L_y(p) = p$ defines the points at which each unit receives exactly the same income. In the case of perfect inequality of incomes, the Lorenz curve is ⌋ - shaped, implying that all income is received by a single income unit. Various summary measures of inequality can be derived from the Lorenz curve, the best known being the *Gini coefficient*, which is equal to twice the area between $L_y(p)$ and the diagonal, or equivalently one minus twice the area under $L_y(p)$.

A generalization of the Lorenz method to deal with variables that are related to income was suggested by Mahalanobis (1960) and later developed by Kakwani (1977a). It essentially requires postulating a functional relationship between Y and a further random variable X, *eg.* $X = g(Y)$, and mapping out values of $F(y)$ and $F_1[g(y)]$, so

that a *concentration curve* for variable X is generated. This curve describes the cumulative share of variable X possessed by each cumulative share of the population, ranked in ascending order of incomes.[7] It is then possible to define a *concentration index* for x = g(y) as:

$$C_x = 1 - 2 \int_0^1 G_x(p)\, dp \qquad (4.5)$$

where G_x(p) is the concentration curve for variable X.

This generalization underlies many of the most widely used statistics of tax progressivity (*eg.* Kakwani, 1977b; Suits, 1977) and horizontal inequity (*eg.* Atkinson, 1980; Plotnick, 1981). It is also the basis for the approach to health and health care inequality measurement developed by Wagstaff, van Doorslaer and Paci (presented in the next section and applied in the empirical chapters). If it is assumed that problems associated with the presence of confounding variables, such as age and sex, can be satisfactorily resolved through careful empirical specification, then the concentration curve approach has some obvious merits. It provides a clear expositional device that is easily understood by policy makers and opens the way to establishing summary measures that take due account of the socioeconomic dimension to health domain inequity. Moreover, in contrast to the Atkinson method, it involves actual measurement of relative differences in the variables of interest.

It is well-known, however, that the family of standard concentration indices implicitly entails an assumption about the weight to be attached to inequity at different income levels which some might find undesirable. This is because the sensitivity of such measures to a hypothetical transfer of the variable X depends on the rank orderings of the persons at which the transfer takes place, rather than on their actual income levels. A greater weight is given to transfers of X around the mode of the income distribution rather than at its tails. An interesting mechanism to counter this problem has been provided in the income inequality literature through a parametric variant of the Gini coefficient - the *extended* or *generalized* Gini (Kakwani, 1980a; Donaldson and

---

[7] Note from this argument that the Lorenz curve of income y can be interpreted as a special case of the concentration curve of the function g(y), when g(y) = y.

Weymark, 1980). This family of measures incorporates a distributional judgement parameter which has most of the properties of Atkinson's ε. It can also be expressed as a function of the Lorenz curve which suggests that the result can be readily extended to other concentration indices (Yitzhaki, 1983; Lambert, 1988). In Section 4.4 such an extension is carried out in relation to health and health care concentration measures, thereby providing health equity analysts with a range of indices that reflect alternative distributional points of view.

4.3 The concentration approach applied to health

This section provides a formal presentation of the *ECuity* indices that are applied in later empirical analysis.[8] The review builds on the continuous random variable framework used to describe the Lorenz curve. Attention is first directed at analysis of inequity in health status, which serves as a building block to examination of inequity in health care financing.

4.3.1 Health Status

Allow that income Y is a continuous random variable with p.d.f. $f(Y)$, distribution function F(Y), and c.d.f. F(y). Assume further that morbidity M is also a random variable, related to Y by means of a continuous invertible function $M = g(Y)$, so that its first derivative and $E[g(Y)] = \mu_m$ (the mean level of morbidity) exist, and $g(Y) \geq 0$ for all $Y \geq 0$. In a manner analogous to the Lorenz curve, the relationship between $p = F(y)$ and $G_m(p) = F_1[g(y)]$, interpreted as the proportional share of total morbidity experienced by persons having an income less than or equal to y, yields the *illness concentration curve*. The curve is represented in a unit square (Figure 4.2).[9]

---

[8] The methodology adopted in the ECuity study is described in a series of publications by the project leaders: *eg.* Wagstaff *et al* (1989), Wagstaff *et al* (1991a,b), Wagstaff *et al* (1992), van Doorslaer *et al* (1992), van Doorslaer *et al* (1993: Chapters 3 and 4).

[9] For ease of exposition, in this chapter I take M to represent *morbidity* and thus speak of *illness* concentration curves and indices. In Chapter 6, the empirical analysis considers *mortality* as an indicator of health. In order to make the present discussion relevant to that context, one simply has to substitute the terms 'morbidity' and 'illness' with 'mortality'.

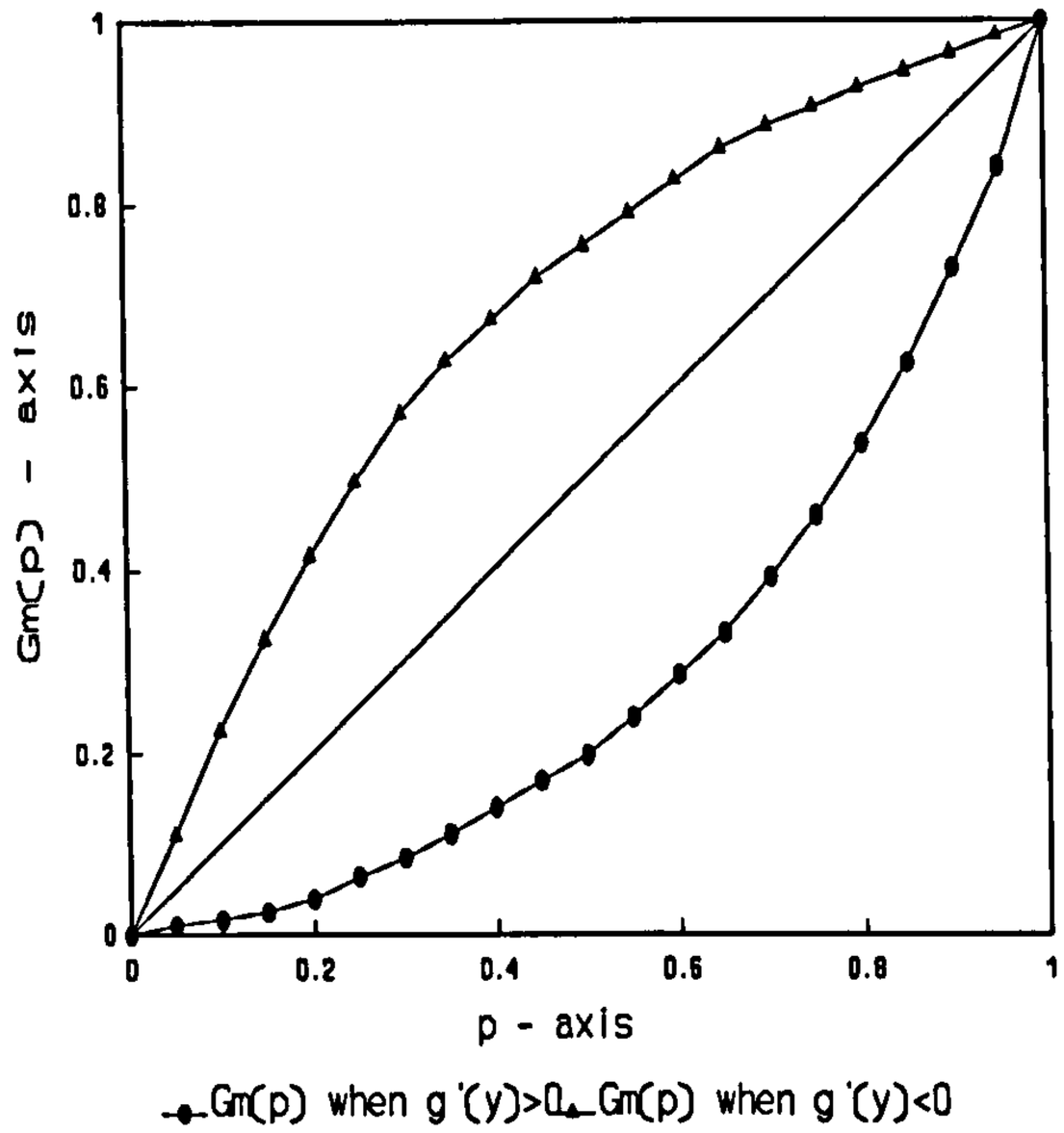


Figure 4.2: Illness concentration curves

The actual shape of the illness concentration curve depends on the sign of $g'(y)$, the first derivative of morbidity with respect to income. If $g'(y) > 0$ for all $y \geq 0$, the curve is convex to the p=F(y)-axis, implying that $F_1[g(y)] < F(y)$ for all y. In this case,the concentration curve falls below the diagonal, signifying that higher income groups experience a greater than proportional share of morbidity. If $g'(y) < 0$ for all $y \geq 0$, the concentration curve is concave to the F(y)-axis and lies above the diagonal, with $F_1[g(y)] > F(y)$. In this case, the poor unambiguously experience a higher burden of total illness. Obviously, when $g'(y) = 0$ at all levels of income, the curve coincides with the diagonal, signifying that morbidity is distributed in exact proportion to income. These results are easily obtained if nonzero values of y and $g(y)$ are excluded. The slope of the illness concentration curve is then given by:

$$\frac{dF_1[g(y)]}{dF(y)} = \frac{g(y)}{E[g(y)]} > 0 \qquad (4.6)$$

implying that the curve is monotonic increasing. To establish curvature take the second derivative:

$$\frac{d^2F_1[g(y)]}{dF^2(y)} = \frac{g'(y)}{E[g(y)]} \frac{1}{f(y)} \qquad (4.7)$$

which is either positive or negative depending on the sign of $g'(y)$ [see Kakwani (1980b, pp. 157-158)]. It is of course quite possible that the illness concentration curve crosses the diagonal (*eg.* if $g'(y)$ is positive at low levels of income but negative at higher levels), in which case no unambiguous conclusions can be made regarding the association of income and morbidity.[10]

Further extending the analogy to the Lorenz curve makes it possible to define the *illness concentration index* as:

$$C_m = 1 - 2 \int_0^y F_1[g(y)]\, dF(y) = 1 - 2 \int_0^1 G_m(p)\, dp \qquad (4.8)$$

where $p = F(y)$, and $G_m(p)$ is the illness concentration curve. This index, which was first suggested by Wagstaff *et al* (1991a), provides a measure of the extent of inequalities that are systematically associated with income position. C_m varies from -1.0, when only the poorest person is ill (in which case $G_m(p)$ is ⌜ - shaped), and +1.0, when the richest person experiences all illness in the community (with $G_m(p)$ being ⌟ - shaped). It takes a value of zero when there is no systematic association between morbidity and the individual's rank in the income distribution.[11]

It should be obvious from the above that in order for C_m to provide an unbiased picture of income related inequalities in health status, it is necessary to remove the effect of confounding intervening variables in the income-health relationship. This issue is addressed by means of direct standardization in the empirical analysis of Chapter 5, where the index is applied to portuguese survey data.

---

[10] See the example shown later in Figure 4.5 when consideration is given to empirical estimation of concentration indices from grouped data.

[11] Given that $G_m(p)$ can cross the diagonal, C_m might register a value of zero even if morbidity shares do not correspond to income shares. This would happen if, for example, the area between $G_m(p)$ and the diagonal lying to the right of the intersection were equal to the corresponding area to its left.

It is worth emphasising that the illness concentration index is not the same as the Gini measure used by Le Grand and associates (*eg.* Le Grand and Rabin, 1986) to measure health inequality. They suggested plotting the cumulative proportions of health against cumulative proportions of the population, ranked by their *health*. This approach has been widely criticized for failing to address the main question which motivated the health inequality debate, namely the systematic association between health variation and differences in economic status [see, *eg.* Carr-Hill (1987); Wagstaff *et al* (1991a)]. In the framework developed above, the illness concentration curve is related to Le Grand's health Gini by

$$C_m = \frac{R[m, r(y)]}{R[m, r(m)]} Z_m \tag{4.9}$$

where R[*a, b*] is the coefficient of correlation between *a* and *b*, Z_m is the Gini index for morbidity, and r(y) and r(m) stand for the rank of y and m, respectively. Thus, the two measures will give the same result only if the ranking of units by morbidity is the same as the ranking by income. If the view is taken that the worrying aspect about inequalities in health is not that they exist but that they are associated with economic position then comparisons based on the univariate distribution of health variables, such as the health Gini, are clearly insufficient. The illness concentration index, on the other hand, takes due account of the income-health relationship.

As with any inequality measure, C_m embodies a particular weighting scheme, which in turn has normative implications for inequality measurement. Given the widespread use of standard concentration measures in economics, these implications are well known [*cf.* Atkinson (1970); Pfähler (1987)]. Greater weight is attached to health inequities occurring around the mode of the income distribution rather than at the sparser lower and upper tails. Furthermore, the measure is not sensitive to income differences as such, but instead to relative rankings of persons given their income positions. There seems to be no particularly compelling reason to believe that such a weighting scheme is desirable. For this reason, the measure is 'generalized' in Section 4.4 to reflect alternative normative viewpoints.

4.3.2 Health Care Finance

The ECuity analysis of health care finance is based on the notion that an equitable health system is one that is financed according to a households' ability to pay. There appears to be some support for this view in Portuguese policy statements (recall Chapter 2), and indeed in those of most other European countries (Van Doorslaer *et al*, 1993). Just as in the analysis of fairness in taxation, the ability to pay principle in health care financing can be interpreted in terms of both *horizontal equity* (households with the same income should make identical payments for health care) and *vertical equity* (households with unequal incomes should provide appropriately dissimilar contributions). The ECuity work focuses on the latter, which appears to be the aspect which most preoccupies observers.

The ECuity researchers make use of the well-known Kakwani (1977b) and Suits (1977) global progressivity indices, which are used in the applied public finance literature to measure the extent to which different people pay different rates of taxation.[12] This approach implies two value judgements of some importance. In order to calculate degrees of progressivity a reference system of redistributional neutrality has to be posited. The implicit assumption made in Van Doorslaer *et al* (1993) is that this is given by *proportionality* of health care payments; which in turn implies, that one accepts a portrayal of the income and health care payments distributions as *relative distances* between equidistant income levels. As an alternative, the distributions could be portrayed in terms of absolute differences, in which case a *per capita* health care tax would constitute the reference system of redistributional neutrality (Pfähler, 1987). The other (related) value judgement implied is that proportionality is the relevant counterfactual when comparing health finance distributions. Aaron (1992) argues that in "analysing health care financing, this standard has no obvious appeal" (p. 468), and seems to suggest that the observed financing distribution of a particular country should be taken as the relevant counterfactual. But this method is scarcely more agreeable.

---

[12] By analogy to the progressivity of taxation, a health care financing system is described as *progressive* when health care payments rise as a proportion of income as income rises; *regressive* when payments fall as income rises; and *proportional* when everyone contributes towards the cost of health care in the same proportion as the income they hold.

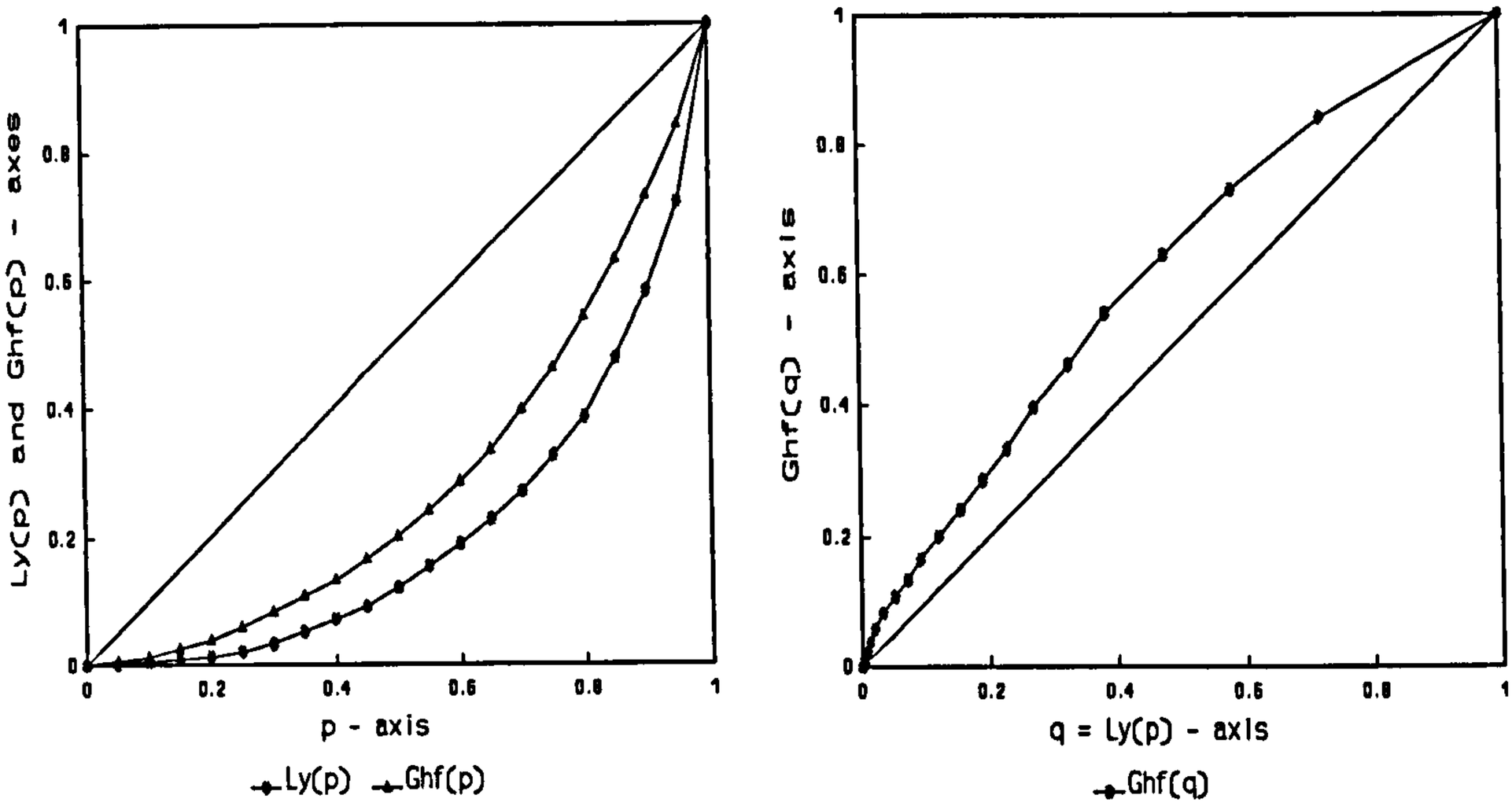


Figure 4.3 (a)-(b): Lorenz and health care payments concentration curves
The Kakwani and Suits progressivity indices

Given the lack of precise policy definitions of what constitutes an equitable distribution of financing, the acceptance of proportionality as the reference distribution seems as good as any other benchmark. The same argument can also be applied to the portrayal of the distributions as relative distances. It is, however, important that these value judgements are recognized when the measures are applied in empirical analysis.

In order to define a Kakwani index of health care finance progressivity, let financing HF be a continuous function of Y, denoted by $hf = g^*(y)$ with $g^*(y) \geq 0$ for all $y \geq 0$. The relationship between $p = F(y)$ and $G_{hf}(p) = F_1[g^*(y)]$ defines the *health care finance concentration curve* [See Figure 4.3(a)], which has analogous properties to those given for $G_m(p)$. The Kakwani index is defined as an area measure between this curve and $L_y(p)$, the Lorenz curve for income defined earlier:

$$\pi_K = 2 \int_0^1 [L_y(p) - G_{hf}(p)]\,dp \qquad (4.10)$$

$G_{hf}(p)$ will lie everywhere above $L_y(p)$ *iff* the income elasticity of health care

payments, $\eta_{g^*}(y) = g^{*\prime}(y) \cdot y / g^*(y)$, is less than unity for all $y \geq 0$, which is equivalent to saying that health care payments are regressive at all levels of income [the case depicted in Figure 4.3 (a)]. Similarly, $G_{hf}(p)$ will lie below (coincide with) $L_y(p)$, *iff* $\eta_{g^*}(y)$ is greater than (equal to) unity for all $y \geq 0$, signifying respectively, progressivity and proportionality of health care payments with respect to income (*cf.* Kakwani, 1977a).

Equation (4.10) may also be expressed as:

$$\pi_K = C_{hf} - Z_y \tag{4.11}$$

where C_{hf} is the concentration coefficient for health care payments and Z_y is the Gini coefficient for income. The bounds of π_K depend on inequality in the income distribution. Maximal regression is given by $-1-Z_y$ (-2.0 if all income is held by one person and all health care payments are made by another). Maximal progression, on the other hand, is given by $1-Z_y$ (+1.0 if income is equally distributed and a single person finances the entire health care system!). Negative, positive and zero values of π_K imply regressivity, progressivity and proportionality, respectively.

The Suits measure is based on *relative* concentration curves, rather than the *standard* curves depicted until now. The difference between the two approaches results from substituting the function $p = F(y)$ on the abcissa by $q = L_y(p) = F_1(y)$, the proportion of total income held by units with $Y \leq y$. Therefore, the *health care finance relative concentration curve*, $G_{hf}(q)$, expresses the relationship between $F_1[g^*(y)]$ and $F_1(y)$, which means that the diagonal provides the reference line of payments proportionality in relation to income [Figure 4.3 (b)]. In some comparisons of vertical inequity in the distribution of health care financing, this graphical representation is more convenient than that for the Kakwani index, since all relevant information can be shown on a single curve for each component of observation.[13]

$G_{hf}(q)$ will lie below the diagonal if $\eta_{g^*}(y) > 1$, signifying that payments are

---

[13] Note, however, that Figure 4.3(a) permits visualization of inequality in the distribution of income as well as non-proportionality in the financing distribution. For some comparisons this may be relevant.

progressive in relation to income, and above the diagonal if $\eta_{g^*}(y) < 1$, signifying that payments are regressive. The Suits index follows logically as an area measure between $G_{hf}(q)$ and the diagonal:

$$\pi_S = 2 \int_0^1 [q - G_{hf}(q)]\,dq \tag{4.12}$$

If the health care financing system is regressive in relation to income, as in Figure 4.3(b), π_S is negative. Maximal regressivity is given by $\pi_S = -1.0$ (implying that all health care payments are made by the poorest person). If, by contrast, the system is progressive, π_S is positive, with maximal progressivity given by $\pi_S = 1.0$ (suggesting that the richest person bears the burden of all health care financing). In common with the Kakwani index, a value of $\pi_S = 0$ is implied by proportionality of payments in relation to income.[14]

Given that, nowadays, the Kakwani and Suits measures are the most widely used indices of global progressivity their descriptive and normative properties are well known.[15] Both measures satisfy the important statistical properties of population replication, scale independence, anonymity and transfers (encountered earlier in the discussion of GE measures). They can also be straightforwardly decomposed by finance source. For instance, if there are T sources of health care finance (*eg.* taxation, private insurance, etc.) the Kakwani index π_K can be expressed as a weighted average of the $T \cdot \pi_{Kt}$ indices for each finance source. Specifically, $\pi_K = \sum_1^T \omega_t \pi_{Kt}$, with weights ω_t being the proportions of each source of finance in total revenue.[16] This decomposition also makes clear that, under the ECuity approach, the vertical equity characteristics of a health care financing system depend on the proportion of total revenues raised from each source and the degree of deviations from proportionality of each of these sources.

---

[14] The definition of the Suits index in the unit interval {-1, +1} is another convenient property in its favour since it facilitates comparisons. Note, however, that the Kakwani index can also be normalized to same interval as $\{-1 = \pi_K/(1+Z_y),\ \pi_K/(1-Z_y) = +1\}$.

[15] Pfähler (1987) provides a detailed discussion of this issue.

[16] The analogous decomposition for the Suits index is given by: $\pi_S = \sum_1^T \omega_t \pi_{St}$. It is also possible to decompose these and other concentration measures by population subgroup (*cf. eg.*, Pyatt, 1976; Silber, 1989). However, the decomposition is not particularly convenient. Given that information on individual rankings in the subgroups and in the total population is used, it is possible for inequality or progressivity within one group to increase but overall inequality/progressivity to decrease.

Despite these points in common, the Kakwani and Suits measures may provide inconsistent rankings whenever the distribution of income is not fixed (which is invariably the case in time and cross-national comparisons). Furthermore, if the health care payments concentration curve crosses the relevant benchmark curve, π_K and π_S may not agree on whether a financing system is overall progressive or regressive (Formby *et al*, 1991). These inconsistencies arise because of differences in the two measures' weighting schemes. The Kakwani index assigns weights according to household rank in the income distribution, implying that maximum significance is given to changes in health care financing ocurring around the mode of the income distribution. The Suits measure, on the other hand, assigns weights according to household rank *and* size of income, implying that maximum significance is given to changes in health care financing ocurring at income levels higher than the mode. However, in one respect the weighting schemes are similar: both are positive and monotonically increasing in income. This implies that they strongly react to payment redistributions from low to high income people (Kakwani, 1986; Pfähler, 1987).

These normative properties suggest two immediate implications for empirical analysis. One is that the computation of indices must be accompanied by careful examination of the relevant concentration curves, checking in particular for crossings. Therefore, in the empirical analysis of Chapter 7, graphical representations of the distributions are also presented. The other implication is that the descriptive power of the analysis would be improved if the measures were extended to accomodate alternative weighting schemes. This leads to consideration, in the following section, of 'generalized' progressivity indices which have negative monotonically increasing weighting schemes.

There are various other progressivity indices which could be drawn on to measure vertical inequity in the distribution of health care payments. In particular, to the Kakwani and Suits indices there correspond specific measures of *redistributive effect*, namely the Reynolds and Smolensky (1977) and Pfähler (1983) indices. These indices measure the redistribution of income which takes place as a result of health care financing rather than deviations of the payments distribution from proportionality. Although the measures have been applied in some ECuity study publications [*eg.* Van

Doorslaer *et al* (1991), Lachaud and Rochaix (1992) and Pereira and Pinto (1992)], there seems to be no great benefit in using them as well. This is not so much because there is "very little evidence that equity objectives in the health field are couched in terms of income redistribution" (Van Doorslaer *et al*, 1993, p. 15). In the partial equilibrium framework adhered to by the ECuity researchers, income redistribution effects are necessarily implied by progressivity of health care payments. The main reason is that the weighting schemes of the Reynolds/Smolensky and Pfähler measures are identical to those of the Kakwani and Suits indices.[17]

4.4 Generalized concentration measures

All the measures considered up to this point embody *specific* normative judgements about the importance attached to inequity at different positions in the income distribution. The adoption of any particular index for empirical research means that the analyst effectively accepts its implied weighting scheme. Yet there is no reason to believe that others will also share the normative judgements implicit in the measure. It is advisable, therefore, that checks are made on the robustness of empirically observed inequality rankings to different distributional judgements. The parametric variant of the Gini coefficient mentioned earlier provides a useful tool in this respect. By changing a single parameter it is possible to make the measure more sensitive to particular points of the income distribution. The work of Yitzhaki (1983) and Lambert (1988, 1989) suggests that the result can be readily extended to other concentration indices. The objective of this section is to propose generalized versions of the indices discussed in 4.3. These will then be used in the empirical chapters to verify the robustness of the ECuity measures to alternative distributional judgements.[18]

---

[17] On a practical note, computation of redistributive impact measures is also more cumbersome, since it is necessary to re-rank the distribution of post-health care payments income if micro-data are available, or apply a correction formula (Jenkins, 1988a). None of the authors cited above rearrange their distributions or apply the formula accordingly, and therefore their Reynolds/Smolensky and Pfähler measures are effectively *underestimates* of the true degree of implied income redistribution.

[18] As far as I am aware Le Grand (1987) is the only study to have used a parametric index to measure health domain inequity. He computed the Atkinson index of health inequality (proxied by age-at-death) for 32 developed countries, using two values of the inequality aversion parameter, 0.75 and 1.25. However, this index, like the health Gini discussed in 4.3.1, does not measure ability-to-pay related inequities.

Consider first a generalization of the illness concentration index, C_m. Making use of the fact that the Lorenz curve for income can be interpreted as a special case of the concentration curve for the function $g(y)$ when $g(y) = y$, and drawing on Kakwani's (1980a, p. 444) and Yitzhaki's (1983, p. 620) formulations of the extended Gini, it follows that a family of *generalized illness concentration indices* can be expressed as a weighted integration of the area below the illness concentration curve:

$$C_m(\delta) = 1 - \delta(\delta - 1) \int_0 (1 - p)^{\delta-2} \, G_m(p) \, dp \qquad (4.13)$$

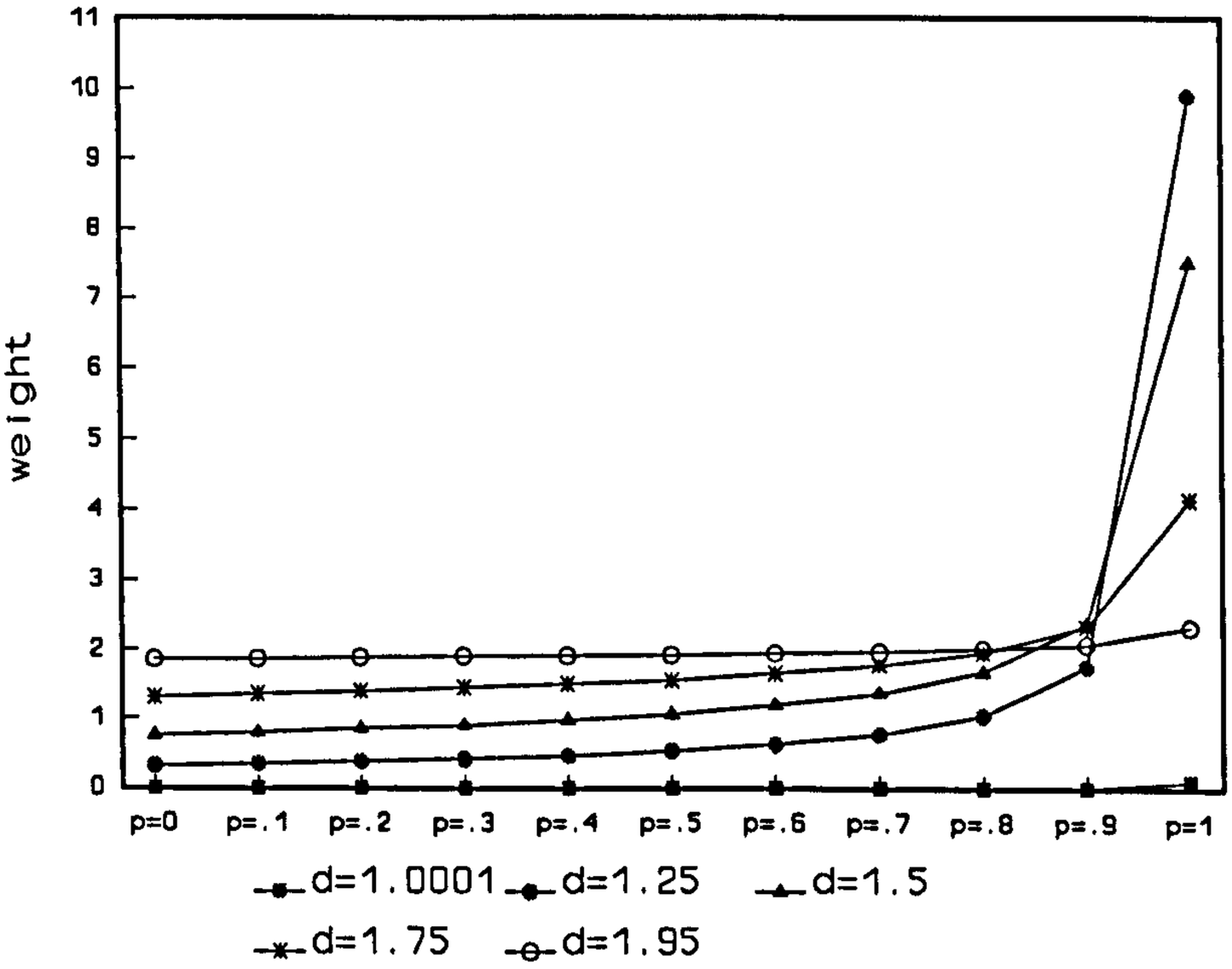
This equation defines various coefficients, one for each value of the distributional judgement parameter $\delta > 1$. As $\delta \rightarrow 1$, $C_m(\delta)$ approaches zero for all distributions, implying indifference to income related health inequalities. At $\delta = 2$, $C_m(\delta)$ is equivalent to the standard illness concentration index, C_m. In the limiting case $\delta \rightarrow \infty$, the implicit ethics of $C_m(\delta)$ approximate the maximin rule (defined over incomes).[19] Each of the $C_m(\delta)$ measures has the same sign properties as C_m, varying from -1.0 to +1.0, with negative values signifying pro-rich health inequality and positive values, pro-poor health inequality.

The effect of varying δ can be seen from Figure 4.4(a-b), which shows the relationship between the weighting scheme [$w = \delta(\delta\text{-}1)(1\text{-}p)^{\delta-2}$] in (4.13) and rank in the income distribution. At values of $1 < \delta < 2$, the weights increase with income rank, signifying that $C_m(\delta)$ attaches greater weight to health inequities at the top of the income distribution (Figure 4.4(a)). At δ=2 (the standard concentration index), weights and rank are independent. Subsequent increases in δ leads the weights to decrease with income rank (Figure 4.4(b)). Therefore, higher values of δ mean that greater weight is attached to health inequities at the lower end of the income distribution and lesser to those ocurring at the upper end. Thus, the generalized illness concentration index puts at the disposal of the analyst a parameter which reflects different judgements about how the concentration curve values are aggregated.

---

[19] It should be apparent that δ has a similar effect on measured inequality to Atkinson's inequality aversion parameter. The behaviour of $C_m(\delta)$ at the extremes $\delta \rightarrow 1$ and $\delta \rightarrow \infty$ resembles that of $A(\varepsilon)$ at the extremes $\varepsilon \rightarrow 0$ and $\varepsilon \rightarrow \infty$ of inequality aversion.

(a)



(b)

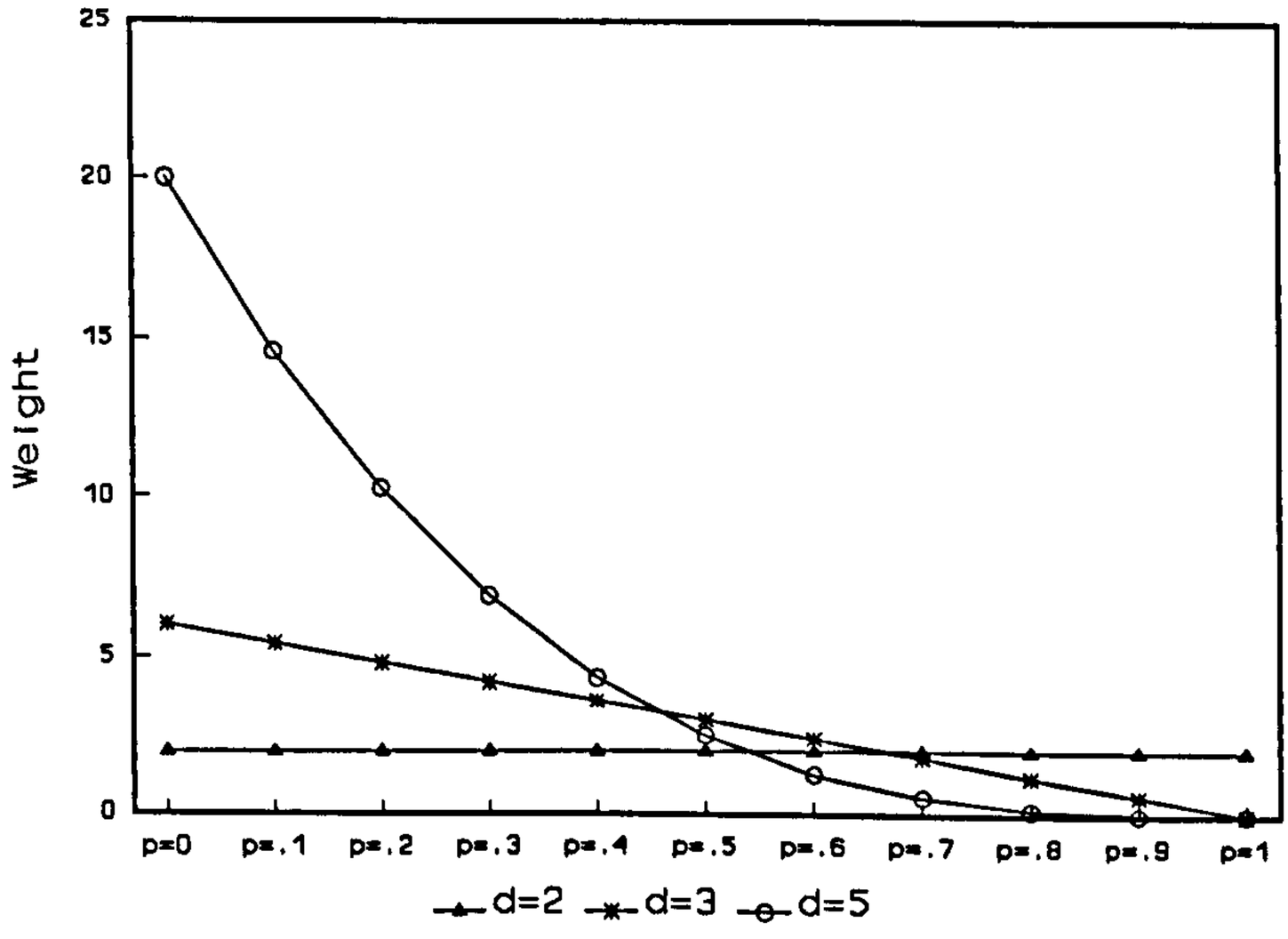


Figure 4.4 (a-b): Weights and income rank using the generalized illness concentration index

The analysis presented above can be readily extended to the concentration measures of vertical inequity in financing of health care. Thus, the family of *generalized Kakwani indices* is given by:

$$\pi_K(\delta) = \delta(\delta-1)\int_0^1 (1-p)^{\delta-2}\,[L_y(p) - G_{hf}(p)]\,dp \qquad (4.14)$$
$$= C_{hf}(\delta) - Z_y(\delta)$$

where $C_{hf}(\delta)$ and $Z_y(\delta)$ are generalized concentration coefficients for health care payments and generalized Ginis for income, respectively. Similarly, the family of *generalized Suits indices* is given by:

$$\pi_S(\delta) = 1 - \delta(\delta-1)\int_0^1 [1-L_y(p)]^{\delta-2}\,[L_y(p) - G_{hf}(p)]\,dL_y(p) \qquad (4.15)$$
$$= \delta(\delta-1)\int_0^1 (1-q)^{\delta-2}\,[q - G_{hf}(q)]\,dq$$

The sign properties of $\pi_K(\delta)$ and $\pi_S(\delta)$ are identical to those of the related standard progressivity measures. When $\delta = 2$, $\pi_K(\delta)$ and $\pi_S(\delta)$ give the standard Kakwani and Suits indices, respectively. Values of $\delta > 2$ yield progressivity estimates that register concern with the health care financing position of poorer members of society. In the case of the Kakwani measures, greater weight is given to the poor on the basis of their *number*; whereas for the Suits indices, this weight is assigned on the basis of number *and* size of income.

4.5 Empirical estimation of concentration indices

The indices presented above are calculated in Part 4 of the thesis by either the *covariance* or *linear approximation* methods. The first of these is particularly useful because of its simplicity and accuracy. Several authors have shown that concentration measures can be rewritten as a linear transformation of the covariance between an income unit's rank in the income distribution and its own value of the variable under analysis [Kakwani (1980, p. 173), Lerman and Yitzhaki (1984), Shalit (1985) and Jenkins (1988c)]. If micro-data are available, this result opens the way for accurate

estimates without the need to assume a specific frequency function.

The covariance formula for the illness concentration index is as follows:

$$C_m = \frac{2\ cov\ [m, F(y)]}{\mu_m} \qquad (4.16)$$

where cov [·] denotes covariance, m = g(y) is the morbidity level of an individual with income y, F(y) is the population share of individuals whose income is no greater than y, and μ_m is the mean level of morbidity. Thus, the procedure for calculating the index is straightforward. First, the income rank *r(y)* for each observation $i = 1, ..., n$ is obtained. Next the covariance between *r(y)* and m is calculated. Since *r(y)/n* terms are the empirical representation of F(y), the covariance must be divided by *n*. Finally, the resulting value is multiplied by 2 and divided by μ_m to obtain C_m.

Taking into account (4.11) and (4.16), the Kakwani index of health care finance progressivity is computed as :

$$\pi_K = \frac{2\ cov\ [hf, F(y)]}{\mu_{hf}} - \frac{2\ cov\ [y, F(y)]}{\mu_y} \qquad (4.17)$$

with the terms having analogous definitions to those given above (*eg.* hf = g^*(y) is the level of health care payments of an income unit with income y). For the Suits index, given that it is based on relative concentration curves, the formula is slightly different (Jenkins, 1988c). The term F(y) in the first covariance must be substituted with $F_1(y)$, defined earlier.

With regard to the generalized indices, Lerman and Yitzhaki (1984) have derived the covariance formula for the Gini. Since the Lorenz curve for income y can be seen as a special case of the concentration curve of a function g(y), when g(y) = y, their result can be readily applied to measures based on concentration curves. Thus, the generalized illness concentration index can be computed as:

$$C_m(\delta) = \frac{-\delta\ cov\ \{m, [1 - F(y)]^{\delta-1}\}}{\mu_m} \qquad (4.18)$$

where δ is the distributional judgement parameter. The corresponding formulas for the

generalized Kakwani and Suits indices follow trivially from (4.17).

When data are only available in grouped form (as is the case in Chapter 6 and in the standardized distributions of Chapter 5), there is little to be gained in applying the covariance procedure. I have chosen in these cases to use the traditional linear approximation method, derived on the assumption that the concentration curve is piecewise linear (Morgan, 1962).[20] One such formula, used here to compute the illness concentration coefficient is:

$$C_m = \sum_{j=1}^{J} (p_j\, m_{j+1} - p_{j+1}\, m_j) \qquad (4.19)$$

where p_j is the cumulative proportion of income ranked units up to the *j*th class and m_j is the cumulative proportion of total morbidity experienced by the same units. The standard progressivity indices for grouped data are given by analogous formulas taking into account equations (4.11) and (4.12).

Lerman and Yitzhaki (1989) have provided a covariance formula for computing the generalized Gini from weighted samples. I use this method to compute the generalized illness concentration index from grouped data in Chapters 5 and 6. The formula is the same as that shown in (4.18) with the covariance component being given by:

$$cov\ \{\cdot\} = \sum_{k=1}^{n} w_k(m_k - \mu_m)\ [(1-E_k)^{\delta-1} - \sum_{k=1}^{n} w_k(1-E_k)^{\delta-1}] \qquad (4.20)$$

where w_k is the proportion of the population in income group *k*, m_k the mean level of illness in group *k*, and E_k a mid-interval estimator of F(y), *ie.* $E_k = \Sigma\ w_{k+1}+w_k\ /\ 2$.

---

[20] The results arrived at by this procedure are identical to those obtained when the covariance method is applied to grouped data. The latter can be applied in either one of two forms: (i) by using a variant of the covariance formula, *viz.*

$$C_m = \left[\frac{(n^2-1)}{6n}\right] \cdot \frac{\hat{\beta}}{\mu_m}$$

where $\beta\hat{}$ [the slope coefficient obtained by regressing *m* on *r(y)*] is estimated through weighted least squares (Wagstaff *et al*, 1991a), or (ii) by weighting the covariance directly (Lerman and Yitzhaki, 1989). In both cases the weights are provided by the relative size of the income groups.

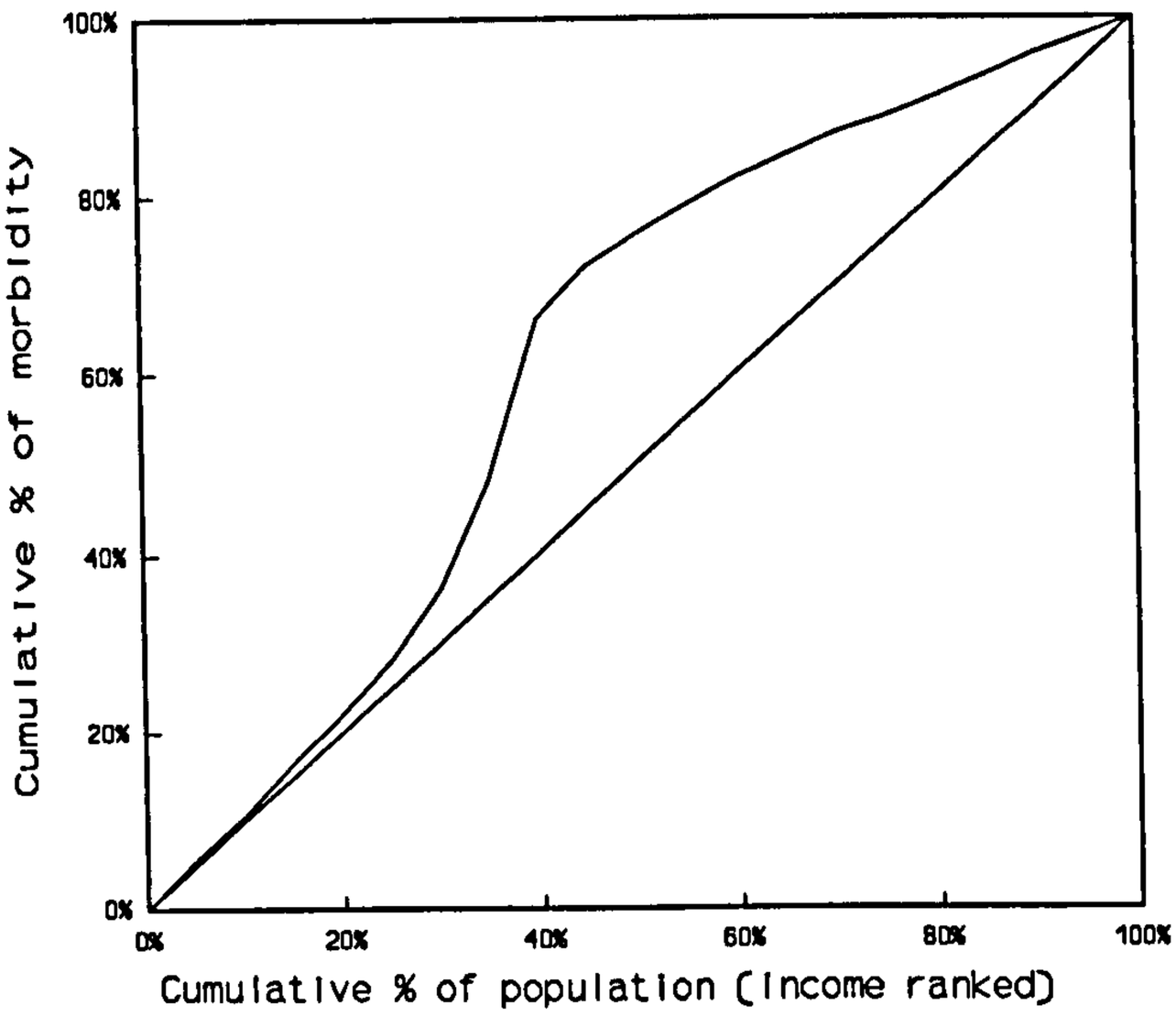


Figure 4.5: Hypothetical concentration curve for morbidity

It is well-known that the linear approximation method leads to biased estimates of concentration indices, since it assumes that inequality is zero within income classes.[21] What is less appreciated is that the direction of the bias cannot be ascertained *a priori*. For the Gini index, linear approximation of the Lorenz curve must inevitably lead to downward bias since the curve is convex to the diagonal (Gastwirth, 1972). However, concentration curves may not be convex (or concave) over the whole range of incomes given that the variable whose shares are represented on the ordinate is not the same as the ranking variable. Hence, a configuration of the type shown in Figure 4.5 is conceptually possible. In this example, income-indexed health inequality is negligible at low levels of income but rises substantially after the third decile, leading to a kink in the illness concentration curve. If the linear approximation method is used with a small number of partitions (say quintiles), the resulting C_m index will *over-*

---

[21] The weighted covariance procedure described above and the weighted least squares estimator used by Wagstaff *et al* (1991a) also suffer from the same problem. It is possible to improve the accuracy of grouped data estimates by using non-linear approximations (*eg.* Kakwani and Podder, 1976). However, such methods have high time costs which make them unattractive when the number of calculations are large.

estimate the true value. An obvious response to this problem is to verify actual bias by means of sensitivity analysis.

Tables 4.1 and 4.2 report the results of estimating standard measures from micro- and grouped data. The data sets are those used in Chapters 5 and 7, where the variables are defined. For present purposes the precise definitions are not important. In the morbidity analysis (Table 4.1), linear approximation leads to over-estimates of the true degree of inequality suggesting that the concentration curves may be similar to the one shown in Figure 4.5. The health care finance analysis shows that the bias is in the opposite direction.

Of greater importance, however, is the *degree* of bias when the number of income groups is varied. The results show unsurprisingly that increasing the number of partitions improves the accuracy of inequality estimates. Segmenting the data by vingtiles leads to a negligible degree of bias (invariably less than one per cent). Conversely, if quintiles are taken as the basis for linear approximation the resulting estimates are inaccurate by between 5 and 12 percent. This may be judged as acceptable. However, the differences are not so small when compared to observed cross-national and time variations (see, *eg.*, van Doorslaer *et al*, 1993). Therefore, two conclusions may be drawn. First, if methods that take no account of inequality within income classes are used, high levels of aggregation should be avoided. Obviously, in some cases the researcher has no choice but to accept published tabulations. However, there are numerous examples in the literature where micro-data are inappropriately segmented into small partitions for the purpose of demographic standardization (*eg.* Pereira and Pinto, 1993). In the next chapter standardization is carried out using vingtile classes, a method which the above results suggest produces an insignificant level of bias. The second conclusion is that due caution should be exercised when comparing results generated through different levels of aggregation. This is particularly relevant given that health domain inequity comparisons are likely to become increasingly available in the future. As will be seen, it also has implications for the empirical analysis reported in the following chapters.

Table 4.1: Comparison of empirical measurements of health inequality (concentration indices) by level of aggregation

National Health Survey, 1987

Variable	Income groups of equal size			
	Micro data	Vingtiles	Deciles	Quintiles
Equivalent income	0.3456	0.3441	0.3400	0.3247
% bias relative to micro-data		0.4%	1.6%	6.0%
Bed days	-0.2032	-0.2038	-0.2118	-0.2253
% bias relative to micro-data		0.3%	4.2%	10.9%
Limiting chronic	-0.1244	-0.1253	-0.1324	-0.1397
% bias relative to micro-data		0.7%	6.4%	12.3%

Table 4.2: Comparison of empirical measurements of health care finance progressivity by level of aggregation

Family Budget Survey, 1989/90

Variable	Income groups of equal size			
	Micro data	Vingtiles	Deciles	Quintiles
Concentration indices				
Equivalent income	0.3514	0.3491	0.3438	0.3267
% bias relative to micro-data		0.7%	2.2%	7.0%
Direct payments	0.1656	0.1649	0.1627	0.1540
% bias relative to micro-data		0.4%	1.8%	7.0%
Total payments	0.3231	0.3212	0.3164	0.2998
% bias relative to micro-data		0.6%	2.1%	7.2%
Kakwani indices				
Direct payments	-0.1858	-0.1842	-0.1811	-0.1727
% bias relative to micro-data		0.9%	2.5%	7.1%
Total payments	-0.0283	-0.0279	-0.0274	-0.0269
% bias relative to micro-data		1.4%	3.2%	5.0%

4.6 Concluding comments

This chapter has presented six summary measures of health domain inequity, three of which are new to health economic research. Appendix 4 provides a summary of all indices discussed (as well as others to be proposed in Part 5). Like all inequality measures used previously by health economists, the new indices draw on analogous measures in the income inequality and public finance literature. They incorporate a key element missing in previous work: flexibility in the method of aggregating inequality at different points of the income distribution. By varying a single parameter the measures become more sensitive to the position of specific income groups; and may therefore be seen as incorporating alternative social judgements concerning the degree of equality preference.

The indices chosen for application in empirical analysis are all based on the slopes and ordinates of Lorenz and concentration curves. This choice rules out other approaches to distributional measurement, such as the GE family of indices which might conceivably provide a versatile route to evaluating the structure and time-trend of health domain inequity. The view taken here is that the approach introduced by Wagstaff and van Doorslaer has obvious merits for examining key questions raised by Sen's ethical structure, and that there is much to be gained in refining and extending their initial effort. The present chapter has provided a formal exposition of the approach that clarifies implicit normative and statistical properties of the indices and is useful in establishing further results such as those to be presented in Part 5. However, the main work required of studies that adopt the Wagstaff-van Doorslaer methodology is in the area of empirical application. The analysis in section 5 suggests that alternative methods for estimating indices can lead to errors in the interpretation of results from comparative studies. The empirical chapters report further measurements aimed at establishing the impact of methodological choices on the degree of measured inequity.

Part 4

EMPIRICAL ANALYSIS

Chapter 5

Measuring Achievement I: Inequity in Health

"If the fundamental imperative is maximum feasible equality of opportunities to 'flourish', and health is a necessary condition for 'flourishing', then greater equality of health itself dominates the other [distributional] principles."
Culyer and Wagstaff (1992, p. 17).

5.1 Introduction

The next three chapters provide empirical estimates of the degree of health domain inequity in Portugal. Each chapter examines a particular aspect of the problem: namely, inequity in the population's health (the present chapter), inequity in infant mortality (Chapter 6), and inequity in family payments towards the health service (Chapter 7).

The justification for focusing on these issues was spelt out in Chapter 3. Recall that in spite of the capabilities approach being adopted as the underlying ethical theory, the empirical analysis does not seek to conceive or to compare unique scalar measures of health capability sets. Instead, the study adopts an indicative pluralistic approach, characterized by separate measurement of inequity in functioning levels (*achievement*) and in the use and payment of commodities which permit their attainment (*advantage*). This approach maintains the key elements of Sen's framework but is actually more informative than the alternative of comparing a global index of the capability set. It is possible, for example, to re-interpret the empirical chapters as attempts to evaluate the degree to which specific Portuguese health policy objectives are being met (*cf.* Chapter 2). Similarly, each of the chapters can viewed as an extension to previous work on particular aspects of health domain inequity (*ie.* inequalities in health and progressivity

of health care payments), which means that the results may easily be compared with those of other studies.

There are three key options underlying the empirical chapters. First, the interpretation given to the term *inequity*. It refers to inequality in the distribution of achievement and advantage attributes that is systematically associated with inequality in the extent of choice over commodity bundles (measured by available income resources). This definition derives from the fact that commodities, or their characteristics, are what an individual requires to achieve functioning capabilities. The second option follows directly from the first. The basic tools of analysis are the concentration indices and associated graphical representations discussed in Chapter 4, which are a suitable means for measuring the type of inequities that command our attention. Finally, there is an underlying preoccupation throughout the empirical analysis to measure the impact of specific methodological choices on the results. These concern, for example, the issues of equivalization of income variables, age and sex standardization of the distributions and, of course, the parameterisation of indices to reflect different distributional points of view. That these and similar issues are discussed in greater depth than in previous work is perhaps the main contribution of the analysis from the disciplinary point of view. Equally important, however, is the unearthing of the health equity profile of the Portuguese health system, so far relatively understudied.

The present chapter measures the extent of inequity in ill-health, drawing on morbidity data from the 1987 National Health Survey. In Section 5.2, the survey is described together with the income and morbidity variables used in the analysis. The section is also essential background to the empirical illustrations provided in Chapter 8. Section 5.3 presents the results, setting apart two issues for further analysis: standardization of the distributions by demographic variables and examination of inequity within age groups. Where appropriate, the results are contrasted with British evidence, mainly because comparative Portuguese research is not available. Section 5.4 provides evidence on an important issue not previously addressed in the literature: the impact of equivalence scale relativities on the degree of measured inequity. The final section (5.5) concludes the chapter.

5.2 Data and variable definitions

5.2.1 The National Health Survey

The data source for this chapter is the *Inquérito Nacional de Saúde (INS)*, or *National Health Survey*. This is a cross-sectional household interview survey used primarily by the health planning authorities to obtain national and regional estimates of the incidence and prevalence of illness, the utilization of services and other health related phenomena (*eg.* direct expenditures, waiting times, smoking habits and patient satisfaction). Despite being the only nationally representative data source which allows such information to be systematically linked to the socio-economic characteristics of individuals, the INS micro-data base has scarcely been used by health economists and other academic researchers.

The survey has been carried out regularly since 1983 and is very closely based on the U.S. National Health Interview Survey.[1] It covers the non-institutionalized civilian population and employs household sample selection with the individual as the unit of analysis. The sampling plan follows a three-stage probability design which yields representative estimates at the national level and for Portugal's five regions: North, Centre, Lisbon and Tagus Valley, Alentejo and Algarve.

Data are collected continuously throughout a calendar year by means of face-to-face interviews based on a questionnaire. This consists of two basic parts: (i) a 'core' set of health, socio-economic and demographic items, and (ii) one or more sets of 'supplemental' health items, which unlike the former are not repeated each year. The 'core' questions include: basic demographic characteristics of household members; disability days occurring in the two-week period prior to the interview; health care expenditures during the same period; doctor visits and other types of service utilization during a three-month reference period; and the acute, chronic and symptomatic conditions responsible for the illness and service use reported.

---

[1] In the initial phases the INS drew on technical support from the U.S. National Centre for Health Statistics. More recently, it is carried out jointly by the Ministry of Health the National Statistical Institute (INE).

Like most large scale health interview surveys, the INS's validity and reliability have been insufficiently researched. There is, however, some evidence suggesting that the survey provides acceptable estimates of key health-related parameters of the Portuguese population. Of the ten national surveys used in the ECuity study, for example, the INS had by far the largest sample size [van Doorslaer *et al* (1993, p. 52)], which generally implies lower sampling error. The questionnaire design appears to guarantee a low degree of observational bias (Falcão and Silva, 1983). The response rate is usually at least 80% (roughly the same as the U.K. GHS), with the majority of non-responses due to eligible respondents not being found at home after repeated calls (Ministério da Saúde, 1987). Validation studies on comparable data sources in other countries have shown national health interview surveys to provide "significant agreement with medical records, doctors statements and/or medical examinations" [Kars-Marshall *et al* (1988, p. 228)]. With regard to reliability of the data, the INS carries out follow-up interviews on an 8% random sample of interviewed households in order to check for reproducibility. Relatively large variations have been detected in this process [Ferreira and Lemos (1987, p. 119)], but they are no greater than those found in comparable surveys [Kars-Marshall *et al* (1988, p. 229)].

The present study uses the micro-data base for the 1987 INS, which is actually the only year that the survey has covered the whole of mainland Portugal.[2] The sample size is 41 585 individuals, of which 37 002 (89%) were used in the analysis following deletion of cases with missing values for the income variable. This loss of information does not appear crucial. Non-parametric tests (Mann-Whitney and Kolmogorov-Smirnov) were carried out for two income proxies (SEG and educational level of the family head) and for the morbidity and demographic variables used in the analysis. The null hypothesis of the two samples having identical distributions was not rejected (at the 1% level) for the income proxies, sex and most of the morbidity variables. Therefore, the representativeness of the results does not appear compromised by the level of non-response to the income question.

---

[2] In other years, the INS was circumscribed to specific geographic regions, making the term *national* in the survey's title largely a misnomer. A further nationwide survey is planned for 1995.

5.2.2 Measurement of income position

The chapter adopts the individual's position in the *income* distribution as a measure of his extent of choice over commodity bundles. In common with most health surveys, the INS only obtains categorical income data. Families are asked to indicate on a card showing ten monthly gross income groupings that which best represents their own. Each family member is then attributed his or her family income category.[3] In order that these data serve as an adequate representation of the variable of interest and allow reliable computation of inequality measures, it is necessary to carry out three adjustments: interpolation of the grouped data, transformation of the gross values into income net of taxes and homogenization of differences in family characteristics. The importance of these adjustments cannot be overestimated. In the final analysis, the accuracy of health inequity measures is likely to depend as much on the appropriateness and detail of the variable used to segment the population as it is on the adequacy of health status indicators.

Interpolation of the categorical income data increases the range of estimation techniques that may be applied, and is in anyway indispensable if adjustments for family characteristics or taxation are to be made. The data were interpolated by assuming that all observations in a given interval *i* are concentrated at the midpoint of the boundaries, *ie.* $x_i = \frac{1}{2}[d_i + d_{i+1}]$. Sensitivity analysis using Frank Cowell's *INEQ* package showed that alternative methods of interpolation - piecewise Paretian, linear and histogram - had a negligible impact on final results. The top interval, which is open-ended in the INS questionnaire, was assumed closed by choosing a value consistent with the gross income distribution of the 1989/90 Family Budget Survey (FBS).[4]

---

[3] A similar procedure is also used to generate yearly income values. I have used the monthly observations since they appear to be more reliable. The yearly variable has a greater number of missing values (6583) compared to the monthly counterpart (4583), and more importantly, dispersion of annual income is actually greater, which is contrary to the logical expectation of the longer time period averaging out short-term fluctuations [see Atkinson (1983, pp. 41-46)]. The reason for these discrepancies may be simply that the notion of "annual income" is hardly ever used in Portugal (even among higher-paid professionals), thereby leading to incomplete and innacurate responses.

[4] The FBS, which is similar to Income and Expenditure Surveys in other countries, is discussed in Chapter 7. It uses the same sampling frame as the 1987 INS, based on the 1981 Census.

A further problem with the INS data is that the respondent is asked to consider gross income (*ie.* factor income plus transfers). It is generally accepted that disposable income is a more suitable proxy of the family's potential command over economic resources [Atkinson (1983), Jenkins (1991)]. Therefore, the interpolated gross incomes were adjusted by subtracting the expected value of income and property taxes and social insurance contributions, thereby providing a measure of the *net cash income* available to families. This was carried out by means of a cell-based approach which matched the INS observations with those of the 1989/90 FBS. The procedure involved, first, computing direct tax rates for 60 family groups (cells) defined by their gross income ranking and family characteristics. That is:

$$t_{kl} = \frac{\sum_j T_{kl}}{\sum_j Y_{kl}} \qquad (5.1)$$

where Y_{kl} and T_{kl} are respectively, gross income and total direct tax liabilities of families belonging to the *k*th gross income percentile range (corresponding to those observed for the INS income variable, after aggregating the data file by family) and *l*th family type (single person aged under 65; single person aged over 64; couple without children; couple with children; single adult with children; other families). The computed tax rates (t_{kl}) were then used to correct the INS gross income values through the formula:

$$N_{jkl} = Y_j (1 - t_{kl}) \qquad (5.2)$$

where N_{jkl} is net cash income of the *j*th family belonging to the matched *k*th percentile range and *l*th family type. An important implication of this procedure is that we assume that effective tax rates were the same in 1987 as in 1989/90.[5]

The third adjustment - homogenization of differences in family characteristics - is required because of the mismatch in units of analysis: the income variable is recorded on a *per family* basis, but the empirical interest (by virtue of the ethical concern for the functioning 'health') is with the *individual*. The simplest solution to this

---

[5] There was a reform of the income tax system in 1989 (discussed in Chapter 7) which is a potential source of bias for these adjustments. It should be noted, that contrary to the approach to be adopted in Chapter 7, here I have used only taxes levied under the "old" income tax system. Therefore, the degree of bias is unlikely to be so great as to significantly affect the computed indices.

problem is to divide the reported income by the number of family members, but this is unlikely to provide an accurate measure of each person's choice opportunities given that, for many items of consumption, economies of scale are likely to operate. The preferred procedure is to use an *equivalence scale* (*ie.* an index that measures the relative income required by families of different composition to maintain the same level of commodity choice opportunities). There is, however, no agreement in the literature as to which of the many scales available is the most appropriate. Moreover, there is growing evidence that the results of distributional comparisons are sensitive to the choice of scale [Buhmann *et al* (1988), Coulter *et al* (1992a, 1992b), Atkinson *et al* (1993)]. Since this is also likely to be true of health inequity analyses that use income to rank the population, section 5.4 checks the robustness of the results to different equivalent scale relativities. There are, however, many situations for which sensitivity analysis of income assumptions is impractical (*eg.* where parameterized inequality indices are calculated for a range of health and health care indicators, the number of calculations can grow out of all proportion). Consequently, there is a need to establish a plausible base-line equivalence scale for the empirical analysis.

Given the thesis' objectives, it is logical to turn to Portuguese experience to decide upon the reference scale. The only study where equivalence factors were derived econometrically from Portuguese expenditure data is Santos (1984), who produced two scales based on the Engel and Prais-Houthakker models of household behaviour. These scales have, however, scarcely been used in subsequent research, since they draw on data from the early seventies when family expenditure patterns were markedly different from what they are now, and because the range of goods considered (food and clothing) was rather limited. Furthermore, as Coulter *et al* (1992a) argue, in a comprehensive review of equivalence scale methodologies, "budget data scales do not have a monopoly of virtue with respect to (other scales explicitly based on) normative judgements" (p. 95). Indeed, they suffer from a number of well-known conceptual and empirical weaknesses.[6]

---

[6] For example: (i) the identification problems raised by Pollak and Wales (1979) imply that econometric analysis of expenditure data cannot provide a single 'correct' equivalence scale; (ii) the reliance on consumer behaviour effectively makes consumer sovereignty the sole basis of welfare judgements, which can lead to ethically unappealing outcomes (Fisher, 1987); (iii) budget data studies

Distributional studies involving Portuguese data have tended to rely on normative scales prescribed by international organizations. The National Statistical Institute, for example, uses the ILO scale. Virtually all income inequality research has, however, drawn on the scale recommended in the OECD *List of Social Indicators* (*eg.* Teekens, 1990; Rodrigues, 1993; Ferreira, 1993; Costa, 1994). This is a very simple construct which assigns a weight of 1.0 for the first adult, 0.7 for other adults and 0.5 for children aged less than 14. Although there are no strong conceptual reasons to accept the OECD scale for the present study, it does have the advantage of maintaining comparability with previous Portuguese research. The choice of any other scale could only be justified if it was clearly more appropriate. Since this is not the case, the OECD's suggestion is used as the reference scale throughout the empirical analysis. It is worth stressing, however, that the sensitivity analysis of section 5.4 shows that the scale does not produce results which are greatly at odds with those obtained through the Santos or ILO relativities.

The income values N_{jkl} were equivalized as: $E_j = N_{jkl} / n^*$, where E_j is equivalent income and n^* is the number of equivalent adults in the family. Equal weighting was assumed with each family member being attributed the value E_j. Therefore, the resulting income distribution is the *individual distribution of equivalent net income.* This variable was used to rank the population when computing the health concentration indices.

5.2.3 Measurement of health

Measurement of health status is probably the most debated issue in health services research. McDowell and Newell (1987) pinpoint the problem rather well:

> "There will probably always be a debate over how best to measure health, and one reason for the debate lies in the complexity and abstract nature of health itself. Like attitude or motivation, health cannot be measured directly ... [T]here is no

---

concentrate solely on private goods and services, but time allocations, leisure and non-cash consumption are also relevant to family well-being; and (iv) the econometric analyses invariably include a number of *normative* choices (*eg.* characteristics used to distinguish family types) whose empirical robustness is hardly ever addressed.

standard scale for health; rather, its measurement relies on assembling a number of *indicators*, each of which more or less adequately represents an element of the overall concept" (p. 12).

The indicators used in this chapter are measures of *self-reported morbidity*. They focus on particular physiological deficits or consequences of ill-health, thereby providing *partial* measurements of true underlying health status. There is no attempt to construct a global health index given that the INS data does not lend itself well to such a task. Instead, the heterogeneity of the INS morbidity measures is used to draw out aspects of critical importance to health-equity evaluations. Among these are the severity of illness experienced by individuals and cause-specific disease prevalence, both of which have been insufficiently addressed in previous economic research.

The choice of indicators is based on Blaxter's (1989) useful classification of morbidity measures commonly found in health surveys. She distinguishes between three categories according to underlying conceptual model: (i) a *social-interactional* or *functional* model, in which ill-health is defined in terms of incapacity to perform 'normal' tasks or roles; (ii) a *subjective* model, in which ill-health is defined in terms of an individual's perception; and (iii) a *medical* model, defining ill-health in terms of pathology or deviation from physiological norms. Blaxter's schema actually reflects the distinction that is often made between sickness, illness and disease.

Definitions and summary statistics of the indicators used in the analysis are shown in Table 5.1. There are six variables conforming to the social-functional model. The first three measure the number of days, in a 2-week reference period, that an individual experienced a particular disfunction: namely, confinement to bed, absence from work or school and restricted activity. The remainder are dummy variables indicating the prevalence of the same disfunctions (*eg. bedridden* attributes a value 1 to all persons who experienced at least one bed day during the reference period). I follow the convention in the literature in describing these six indicators as measures of *disability*.

The subjective health model is not well represented in the INS. The typical survey question whereby respondents are asked to assess their general health as "good, fair or

Table 5.1: Morbidity indicators used in the analysis: description and summary statistics

Morbidity indicator	Description	Mean	S.D.	Min.	Max.
Social-functional model					
Bed days	Number of bed days	0.425	2.084	0	14
Off-work days	Number of days absent from work or school for health reasons	0.420	1.877	0	14
Restricted activity days	Number of days with restricted activity	1.142	3.238	0	14
Bedridden	At least one bed day	0.061	0.239	0	1
Off-work	At least one day absent from work or school for health reasons	0.079	0.270	0	1
Restricted activity	At least one day with restricted activity	0.161	0.386	0	1
Subjective model					
Illness	No disability, but reported illness or not feeling well	0.137	0.344	0	1
Medical model					
Limiting chronic	Disability due to longstanding condition	0.116	0.320	0	1
Acute	Disability or illness due to acute pathology	0.063	0.243	0	1
Chronic	Disability or illness due to chronic pathology	0.122	0.327	0	1

(cont.)

Table 5.1: Morbidity indicators used in the analysis: description and summary statistics *(cont.)*

Health Variable	Description	Mean	S.D.	Min.	Max.
Medical model (cont.)					
Circulatory	Disability or illness of the circulatory system	0.036	0.187	0	1
Respiratory	Disability or illness of the respiratory system	0.051	0.221	0	1
Digestive	Disability or illness of the digestive system	0.031	0.172	0	1
Muscoskeletal	Disability or illness of the muscoskeletal system	0.057	0.232	0	1
Injury and poisoning	Disability or illness due to injury or poisoning	0.014	0.117	0	1
Acute respiratory	Disability or illness due to acute pathology of the respiratory system	0.038	0.190	0	1
Chronic respiratory	Disability or illness due to chronic pathology of the respiratory system	0.011	0.103	0	1
High blood pressure	If suffers from high blood pressure	0.176	0.381	0	1

Notes: 1. The means of the two category variables indicate the proportion of persons with positive reporting of the respective condition.
2. All variables have a two week reference period except for the high blood pressure indicator.

not good" was not included in the 1987 survey. Persons who reported no disability days were, however, asked if they felt "ill" or "unwell". The responses to this question make up the *illness* indicator. Given that persons reporting functional disability are excluded, it is likely that, on average, this indicator reflects less serious conditions.

Eleven measures were chosen as representative of the medical model. The first ten are based on supplementary information to the disability and self-assessment questions. For example, *limiting chronic* indicates persons whose disability in the reference period is the result of a medical condition present for at least three months. The INS also enquires about the precise pathology which caused disability or illness.[7] The variables listed from *acute* to *chronic respiratory* were constructed from this information. Diseases are distinguished on the basis of (i) acute and chronic pathologies, and (ii) broad ICD categories. In the latter case, the five most prevalent disease groups identified in Chapter 2 (Figure 2.1) were chosen. Disability or illness of the respiratory system is also further disaggregated into acute and chronic pathologies, given that in this case the implications for the health of individuals are particularly clear. Acute respiratory illness (*eg.* colds, influenza, upper respiratory infection) seldom leads to serious or prolonged health problems, whereas chronic manifestations (*eg.* emphysema, asthma, chronic bronchitis) have a profound impact on the quality of life of sufferers. The final 'medical' indicator - prevalence of high blood pressure - is distinct in the sense that it refers to a predisposing condition and also because it was generated by the interviewer providing the respondent with a check-list of diseases.

It is important to recognize the advantages and limitations of these variables as instruments for measuring health inequity. Four basic criteria that one might expect such indicators to meet are: (i) relevance; (ii) validity; (iii) independence of the decision to report from cultural and economic bias; and (iv) ability to distinguish between different degrees of ill-health. In one sense, all the indicators are relevant. They are

---

[7] Individuals are asked to report either the disease or the four main symptoms experienced. This information is then codified by the interviewer according to an adaptation of the ICD-9 classification, specially geared to illness. The information is later verified by survey staff at the central level, who also codify the pathology as acute, chronic or symptomatic. The validity of the data thus generated is discussed below.

partial measures of underlying health *according to* the particular model under which they are classified. Since all three models are legitimate approaches to defining health, it makes sense to compare and contrast the results obtained for each category. The point is often made, however, that from the perspective of social policy, disability measures are the most suitable (*eg.* Jazairi, 1976; Culyer, 1978). The important questions for the policy-maker are: who is incapacitated by ill-health and what is the time duration of their inability to participate in day-to-day activities. Illness which has no social consequences is of less interest. Furthermore, it has been shown that, of all conventional morbidity indicators, disability days are the most important predictor of medical care use (Newman, 1975).

With regard to validity, there are problems with the 'medical model' variables. Given that individuals' awareness of pathology is often imperfect, it is not at all clear that a population-based survey can properly ascertain the prevalence of diseases. Indeed, where interview data have been compared with medical records, the extent of non-correspondence has sometimes been shown to be relatively large (NCHS, 1967; Kirscht, 1971). However, this does not mean that self-reports are necessarily invalid. Most of the conditions declared will be medically diagnosed. Some will not, but it could be argued that a true record of morbidity ought to include those conditions which respondents suffer from, can identify in broad medical terms, but never consult about. A more serious problem is that responses have been shown to depend on the method of eliciting information. Asking persons to name their condition leads to underestimates of the prevalence of diseases (NCHS, 1967), whereas providing check-lists tends to stimulate reporting (Blaxter, 1989).[8]

The third criterion is in many ways the most important. In measuring inequity, we need to be reasonably certain that there is no selection bias, related to income position, in the decision to report illness. Many of the chosen variables raise doubts in

---

[8] Blaxter (p. 207) provides the example of the General Household Survey. Before 1977, about a quarter to a third of the population declared 'chronic illness', but when in the 1977 edition a check-list of chronic illnesses was introduced, 56% of men and 70% of women identified themselves as suffering from one or more of the conditions listed.

this respect. Better-educated individuals (who also tend to be richer) are likely to give more informative answers to the pathology questions.[9] Similarly, self-perceived health may be subject to distortion because of cultural norms related to socioeconomic status (d'Houtaud and Field, 1984). Undoubtedly too, the work absence indicator (and the restricted activity measure, to a lesser extent) may show a degree of bias related to the nature of work or social security arrangements. However, such doubts would in effect rule out most morbidity measures available in large scale sample surveys. If we wish to monitor inequity in a country's health, the best that can be done is to use the indicators while at the same time remaining alert to possible biases.

Finally, it is reasonable to expect that a measure provide some indication of the severity of illness, given that there is no simple cut-off point between those who are healthy and those who are not. The sickness days measures are useful in this respect, since they indicate the time duration of incapacity. In general, longer periods of disability will be associated with greater severity. The categorical variables do not permit such inferences to be made, when observed in isolation. However, by comparing the degree of inequity in, say, acute and chronic disease, it is possible to shed some light on the variation of illness severity across income groups. By the same token, comparisons of inequity in ICD disease categories provide a means of ascertaining whether particular income groups carry a heavier burden of specific diseases.

In sum, all the morbidity indicators used in this chapter have deficiencies of one kind or another; but they also have positive features which render them useful in health-inequity analysis. On balance, it would seem that the disability indicators are the most acceptable: they are clearly defined, less open to subjective variation, permit inferences to be made about the severity of illness and represent a view of health which has greater relevance for social policy.

---

[9] The correlation between equivalent income and years of schooling among adults in the INS sample is r = 0.57.

5.3 Results

5.3.1 Actual and standardized distributions of morbidity

Table 5.2 presents the estimated concentration indices for the various morbidity indicators.[10] There are two sets of calculations: one for the observed distributions and another for the age-sex standardized counterparts. Consider first the unstandardized results. Without exception, these indicate a negative relationship between equivalent income and self-reported morbidity. However, the degree of measured inequity varies considerably according to type of illness. The largest absolute values are for the disability indicators and the smallest for 'medical model' variables mainly representing acute conditions. The concentration curves for seven variables actually cross the 45° line. Nevertheless, chi-squared tests on cross-tabulations of morbidity with equivalent income quintiles indicate that the relationship between income and illness is statistically significant at the 0.001 level for all but the two acute illness variables (*acute* and *acute respiratory*). This tends to suggest that income related inequity is inexistent for short-term illnesses but prevalent when more serious conditions are considered.

However, there is a possibility that these findings are biased due to the confounding influences of intervening demographic variables (*ie.* the determinants of the health production function specified in the diagrammatic framework of the Sen approach). Given that age and gender are likely to be systematically associated with both income and morbidity, the observed relationship may simply be a reflection of distinct age and sex structures at different points of the income distribution. The common procedure for dealing with this problem is direct standardization for age and sex.[11] This involves calculating the number of persons (or number of sick days) in a

---

[10] The indices were computed from the micro-data base using the covariance formula shown in equation (4.16). Note that the method is valid since the adjustments made to the income variable produced a relatively large number of distinct observations (over 500). In common with the rest of the literature, confidence limits for the index values were not calculated. Frequency distributions of morbidity by equivalent income quintile are shown in Appendix 4 (Tables A4.1 and A4.2), which reports further results pertaining to the analysis of this chapter.

[11] The technique of direct standardization is commonly identified with epidemiologists, but it has been widely used by economists when measuring socio-economic inequity in health and health care (*eg.* Le Grand, 1978; O'Donnell and Propper, 1991a).

Table 5.2: Concentration indices for morbidity
Actual and age-sex standardized distributions.

Indicator	Concentration Indices	
	Actual	Standardized
Bed days	-0.203	-0.192
Off-work days	-0.146	-0.126
Restricted activity	-0.159	-0.108
Bedridden	-0.169	-0.160
Off-work	-0.114	-0.080
Restricted activity	-0.120	-0.082
Illness	-0.059 *	-0.010 *
Limiting chronic	-0.124	-0.124
Acute	-0.023 *	-0.017 *
Chronic	-0.094	-0.071
Circulatory	-0.112 *	-0.074
Respiratory	-0.021 *	-0.014 *
Digestive	-0.042	-0.035
Muscoskeletal	-0.130	-0.109
Injury and Poisoning	-0.092	-0.088
Acute respiratory	-0.004 *	0.004 *
Chronic respiratory	-0.058 *	-0.055
High blood pressure	-0.072 *	-0.039

* Concentration curve crosses the diagonal

given income group on the assumption that the group has the same sex-age distribution as the whole sample. Naturally, this procedure does not address the issue of causality, and would be inappropriate if the aim of the research were to explain inequalities in health. However, it is a valid technique for measuring income-indexed health inequity, since it removes collinearity between income and demographic factors. The formula used to calculate the standardized values was

$$m_y^* = \sum_j \sum_k \frac{m_{yjk}}{n_{yjk}} \cdot \frac{n_y n_{jk}}{N} \qquad (5.3)$$

where m_{yjk} denotes the number of individuals reporting morbidity (number of days sick for the disability days indicators) in the *y*th income group, *j*th age group and *k*th sex; n_{yjk}, the number of individuals in the *yjk*th income-age sex group; n_y, the number of individuals in the *y*th income group; n_{jk}, the number of individuals in the *jk*th age-sex group; and N, the sample size.[12]

Given that the process of standardization involves grouping the data by income class, the estimated indices for the standardized distribution are less exact than those computed from micro data. In order to minimize such bias, standardization was carried out using 20 income groups of roughly equal size. Concentration indices were then computed from the resulting vingtile frequencies by means of the linear approximation method (equation 4.19). The results of the simulation exercise carried out in Chapter 4 suggest that this approach produces an insignificant degree of bias.[13] Therefore, the indices presented in Table 5.2 permit reliable comparisons of the effect of demographic standardization.

For each type of morbidity the effect of standardization is to reduce the degree of measured inequity. This is consistent with findings in other countries [*eg.* Christiansen (1990) for Denmark or O'Donnell and Propper (1991a) for the UK], and reflects the fact that the elderly, who are generally more susceptible to illness, are disproportionately located in lower income groups. The reduction in the degree of pro-rich inequity varies between 5 and 25 per cent for most of the morbidity indicators. However, there is a much larger variation for the 'subjective' *illness* measure, with the absolute value of the concentration index declining from 0.059 to 0.010. One possible explanation for this effect is that subjective evaluations of health status are highly dependent on the

---

[12] Five age groups were used for age-standardization (0-17 years, 18-34 years, 35-44 years, 45-64 years, and 65+ years).

[13] Recall from Table 4.1 that linear approximation estimation based on unstandardized income vingtile frequencies resulted in a less than 1 per cent bias relative to the exact micro-data measures.

interaction between demographic characteristics and socio-economic position. Thus, there may be a significant degree of selection bias in the decision to report illness. There is some evidence, for example, that middle-aged poor individuals hold a particularly pessimistic view of their own health (*eg.* Blaxter, 1990). If the conjecture is correct, then the rather large degrees of health inequity identified for self-assessed health in van Doorslaer *et al* (1993) - where age-sex standardization was not carried out - may be somewhat overstated.[14]

The standardized indices continue to show varying degrees of inequity depending on the type of morbidity under consideration. The largest absolute value (-0.192) is recorded for the *bed days* indicator, which describes the most serious loss of social function. The relative frequencies reported in Table A4.2 indicate that the poorest income quintile accounts for 28% of all bed days reported compared with 11% for the richest. It is noticeable that the disability *days* indicators reveal a higher degree of pro-rich inequity *vis-à-vis* the prevalence measures based on the same questions. This is indicative of greater severity or slower recovery periods in lower income individuals.[15]

Further evidence of greater severity among poorer income groups is provided by the *illness* and 'medical' model variables. Recall that *illness* describes persons who felt unwell in the two-week reference period, but suffered no functional incapacity. This indicator shows no systematic relationship with income, whereas the disability prevalence measures all reveal a significant pro-rich bias. Note also the differences between the computed indices for *limiting chronic, chronic* and *acute* disease. There is scarcely any income related inequality for acute conditions but this rises significantly once chronic diseases are considered (particularly if they limit the individual's activity). These results are consistent with British evidence. For example, Blaxter (1990), who analysed data from the *Health and Lifestyles Survey*, found that lower income groups

---

[14] The estimates of health inequality computed by van Doorslaer *et al* are shown in the Appendix as Table A4.6. I do not compare them with those estimated here since they were not age-sex standardized. A comparison would be open to the objection that observed cross-national differences were simply due to different demographic compositions of the income groups in the various countries.

[15] It also casts doubt on the suitability of broad-based prevalence measures as a means of identifying persons with 'equal' levels of health status when measuring horizontal inequity in the delivery of care (*cf.* Chapter 9).

were significantly more likely to suffer from multiple and more serious conditions. Research based on the *General Household Survey* also points in the same direction. Propper and Upward (1992) computed concentration indices for acute and limiting chronic illness in the order of -0.047 and -0.122, respectively.

The pathology indicators reveal that health inequity is by no means uniform across disease categories. There are strong negative associations with income for muscoskeletal disease and injury and poisoning but the relationship is rather weak for respiratory disease. The latter is in stark contrast to the steep socio-economic gradients commonly found for mortality differentials in the same disease category (*cf. eg.* DHSS, 1980). Although there are no comparable mortality data for Portugal, the discrepancy may be a sign that inequities in cause-specific death rates are a bad proxy for analogous inequities in illness. Note, however, that there is a large difference between the *acute-* and *chronic respiratory* disease indices and that the concentration curve for chronic conditions does not cross the diagonal. This suggests that there is no income related inequality for the first case; while in the second, the poor support a relatively high burden of illness. Finally, the *high blood pressure* variable also points to a pro-rich bias, although of a reduced degree.

Figure 5.1 shows the concentration curves for (standardized) *limiting chronic* and *acute* illness. The two configurations are in fact typical of the curves for other variables, which for reasons of space are not shown.[16] The acute illness curve crosses the diagonal twice and does not deviate from it a great deal at all points of the income distribution. This is also true of the *illness, respiratory* and *acute respiratory* curves, confirming the view that there is no systematic relationship of these variables with individuals' command over resources. The curves for all other morbidity indicators lie everywhere above the 45° line. They are also, like the *limiting chronic* curve shown, generally concave to the population-axis, suggesting that functional incapacity, chronic illness and most disease categories, decrease monotonically with income. Therefore, we can safely conclude that pro-rich inequity exists for these types of illness.

---

[16] See, however, the frequency distributions in Table A4.2.

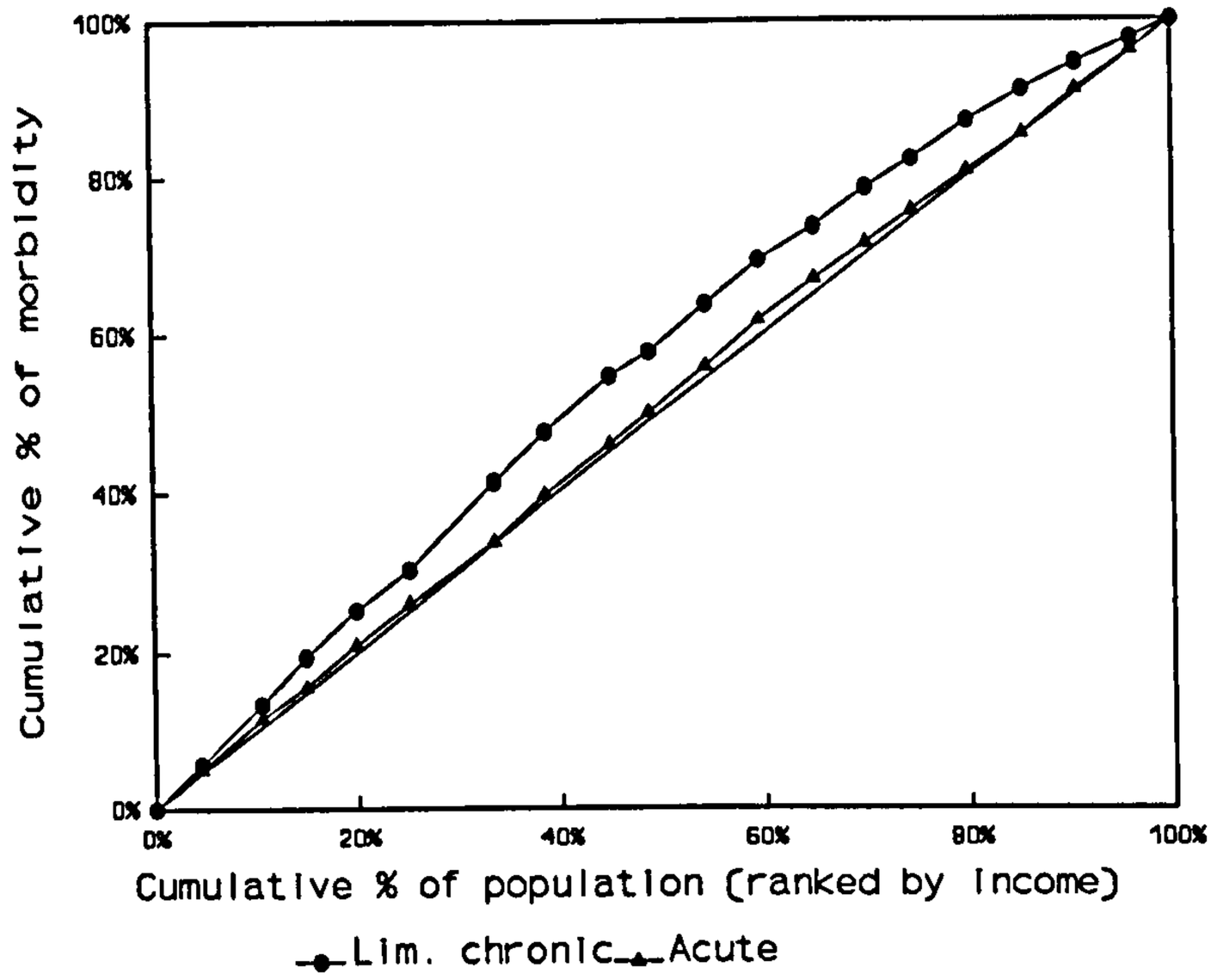


Figure 5.1: Concentration curves for two morbidity variables
Age-sex standardized

The results are also robust to different distributional perspectives. Table 5.3 presents generalized concentration indices for six of the variables previously examined.[17] Recall that values of $\delta \rightarrow 1$ imply indifference to income related inequality; $\delta = 1.5$, that greater weight is attached to the top end of the income distribution; $\delta = 2$, a weighting scheme identical to the standard index (*ie.* more emphasis on modal incomes); and $\delta = 3$, greater weight to persons with lower incomes. Unsurprisingly, the degree of measured inequity rises as δ is increased. When there is a strong concern for the position of the poor, this has the effect of accentuating the absolute differences between inequity levels for each type of morbidity. However, the ranking of these levels does not change at $\delta > 1$. Therefore, unless one is completely disinterested in the relationship between economic inequality and health, the conclusions reached by the standard concentration analysis remain valid.

---

[17] The unstandardized and standardized indices were computed by means of equations (4.18) and (4.20), respectively. The age-sex standardized vingtile distributions were used in the second case. Inequality indifferent indices were computed with δ set at 1.001. The slight differences in the δ=2 indices in relation to those reported in Table 5.2, are due to rounding error. Generalized equivalent income Gini's were also computed. The following values were found: 0.001 at δ=1; 0.229 at δ=1.5; 0.346 at δ=2; and 0.467 at δ=3.

Table 5.3: Generalized concentration indices for selected morbidity indicators
Values of δ = 1.0, 1.5, 2.0 and 3.0

Value of δ	Bed days	Off work days	Restr. act. days	Limiting chronic	Acute	Chronic
Unstandardized						
1.0	0.000	-0.000	-0.000	0.000	0.000	0.000
1.5	-0.126	-0.093	-0.099	-0.079	-0.014	-0.063
2.0	-0.203	-0.146	-0.160	-0.125	-0.023	-0.094
3.0	-0.274	-0.197	-0.218	-0.166	-0.029	-0.106
Standardized						
1.0	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
1.5	-0.114	-0.079	-0.065	-0.077	-0.010	-0.047
2.0	-0.193	-0.127	-0.108	-0.121	-0.017	-0.069
3.0	-0.290	-0.180	-0.162	-0.176	-0.026	-0.092

5.3.2 Distribution of morbidity by age group

The standardization procedure used above has obvious merits but also some drawbacks. One is that it obscures information that may be useful to the policy maker. In particular, since the morbidity experience of different age groups is conflated into a single number, it is no longer possible to examine inequity at different stages of the age cycle. It is important that this be done for two reasons: (i) because illness at different ages has different causes, and (ii) because opposing socio-economic trends in different phases of the life cycle may cancel each other out. Thus, although standardization is a useful device for measuring the overall level of health inequity, a detailed analysis requires that the global estimates be disaggregated by relevant age groups.

The approach taken here is to compare the degree of income related inequality within age groups. The INS data were divided into five sub-samples, respectively

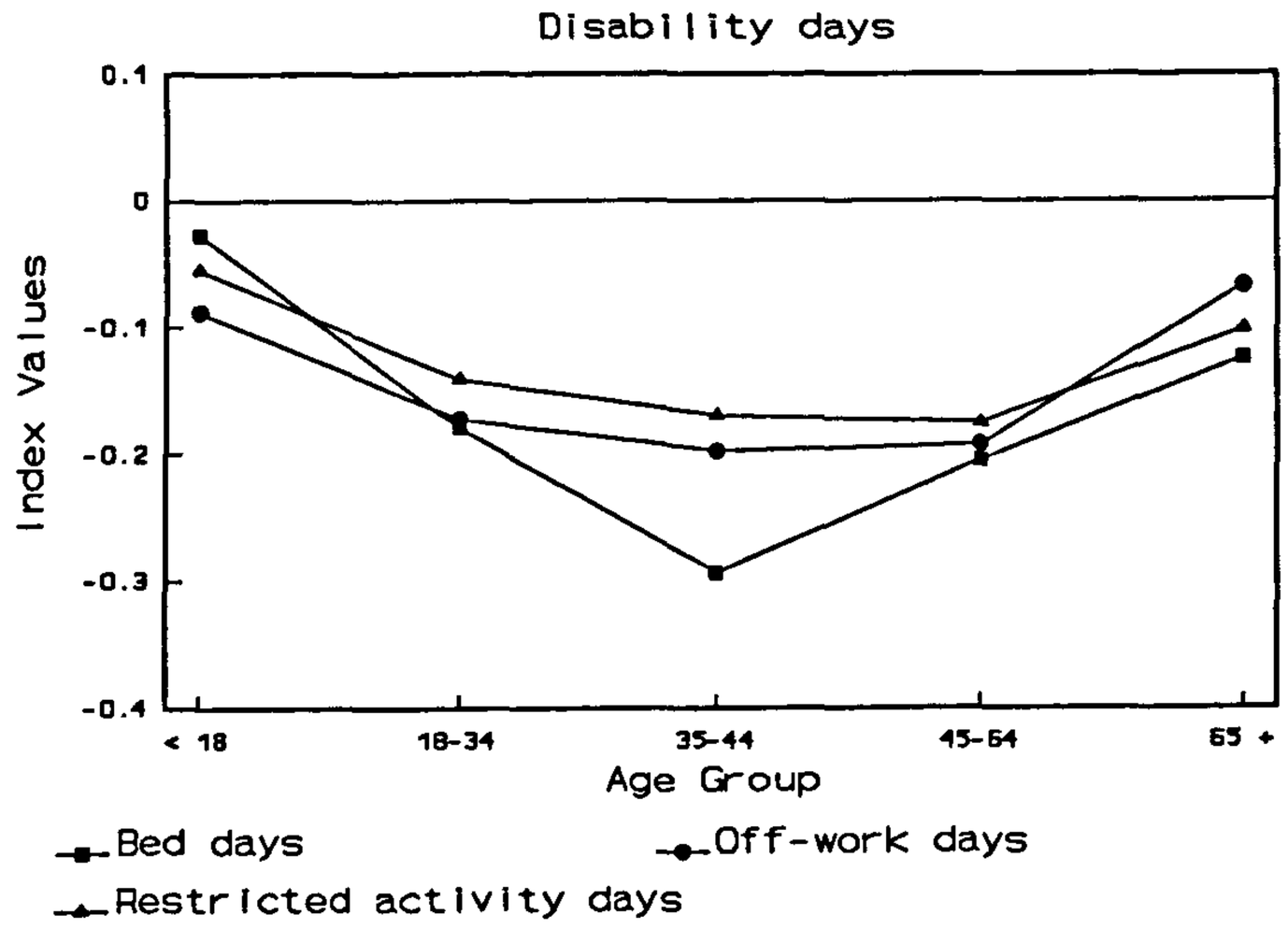
comprising individuals aged under 18, 18-34, 35-44, 45-64 and 65 years or more. Age-specific concentration indices were then computed for twelve of the morbidity indicators examined previously.[18] Note that this method does not indicate the contribution of age to overall inequality. For that purpose, the appropriate measures would be the generalized entropy indices discussed in Chapter 4, which would provide an unambiguous decomposition into within- and between-group components. The present method seeks only to provide a succinct quantitative description of health inequity at different stages in the life cycle.

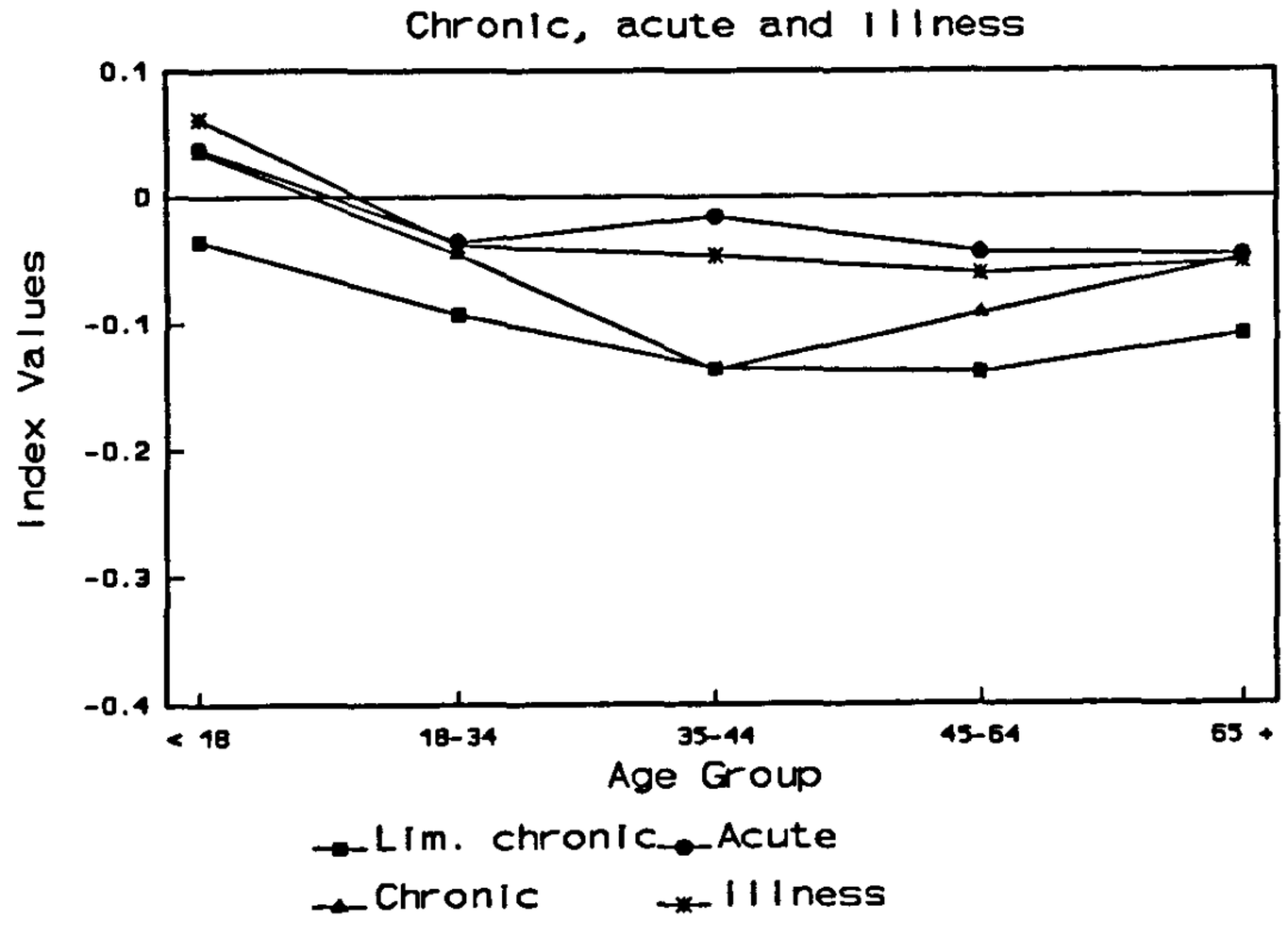
The main results are shown in Figure 5.2, which plots the computed index values for different types of morbidity. The actual values are shown in the Appendix as Table A4.3. Chi-squared tests were also run on the data to verify the significance of the association between morbidity and income at different ages. These are important because of the likelihood of small numbers reporting particular types of illness in some age groups. The results are shown in Table A4.4. Note that the distributions were not sex-standardized as this was found to have a negligible impact on the degree of measured inequity.

There are two basic patterns to the results. One, where pro-rich inequity is non-existent in childhood and adolescence, particularly high in middle age and lower, but still significant, in persons aged over 64; and another, where the relationship between income and morbidity is weak across all (or most) age groups. The disability days indicators provide the clearest evidence for the first pattern; but similar U-shaped age structures are found for the 'medical' model variables: *limiting chronic, chronic, circulatory, muscoskeletal* and *injury and poisoning*. Note from Table A4.4 that none of these indicators show a significant association with income for the under 18 age group. Thus, although varying levels of the concentration index are computed (ranging between -0.088 and +0.035) the results are not strong enough to suggest that either rich

---

[18] The excluded indicators are the prevalence of incapacity measures, acute and chronic respiratory disease and high blood pressure. Note that the *off-work days* indicator is included even though it is less relevant at the extremes of the age structure. The justification is that the measure also includes days absent from school (72% of under 18's in the sample were either studying or in the workforce); and that a significant number of over 64's (25%) were still in work.





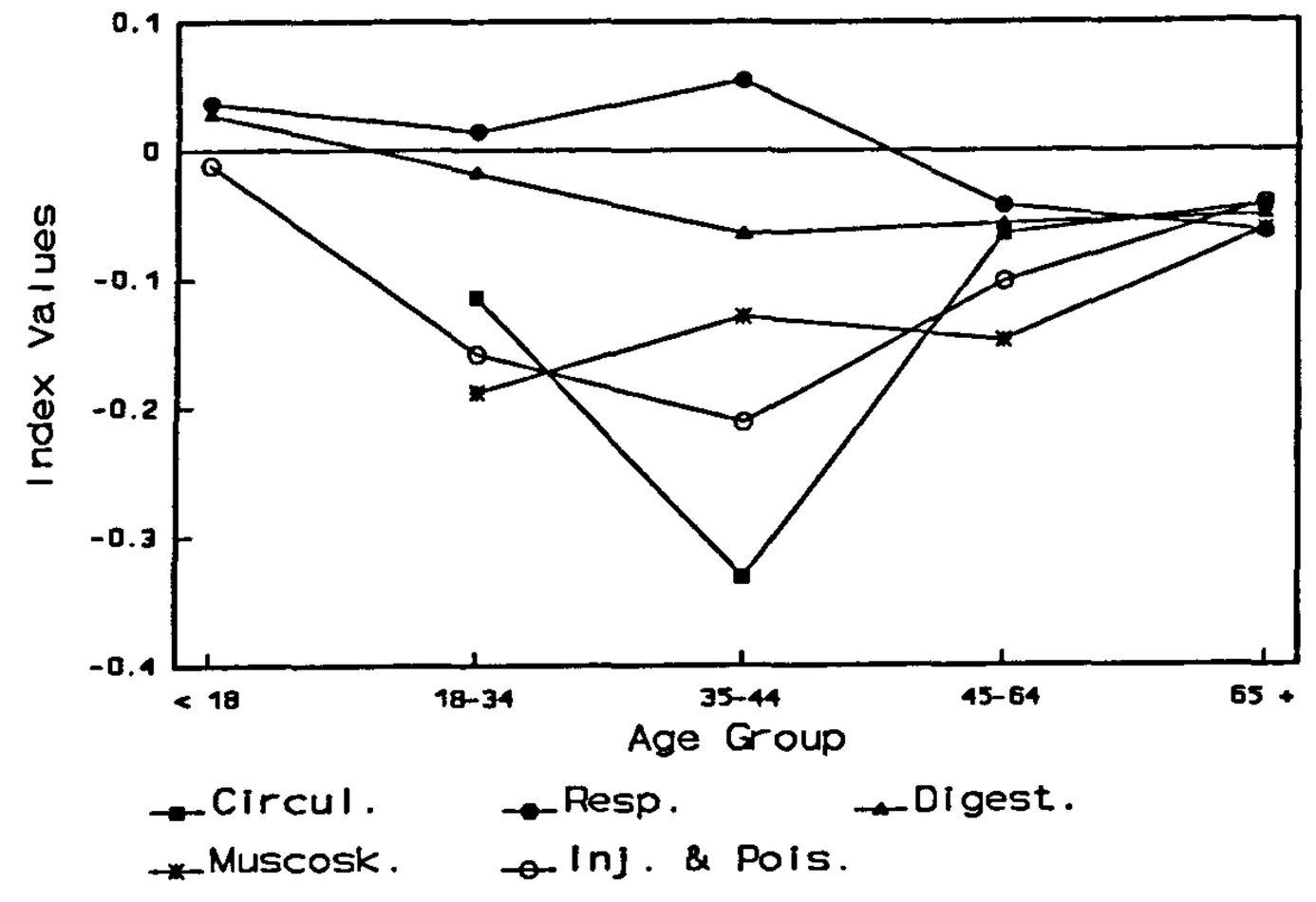


Figure 5.2: Morbidity concentration indices by age group

or poor are favoured.[19] In adulthood, however, poorer individuals report comparatively higher levels of the types of morbidity indicated. In middle age, the degree of inequity is large and highly significant. The high index value computed for the *circulatory* variable in the 35-44 age group may be a sign that poorer individuals are more prone to experience heart and cerebrovascular disease at an early age. The pro-rich bias declines in old age, but is still statistically significant for most morbidity categories.

The second age pattern is shown quite clearly by three indicators: non-incapacitating *illness* and the variables representing acute and digestive system conditions. In each case, negative values of the concentration index were computed for the adult population; and positive values for the under 18 age group. However, the degree of inequality is relatively low at all ages and the chi-squared test is seldom passed. Therefore, the measurements for these types of illness point to a generally weak association with income throughout the age cycle.

There are a number of points to note about these results. First, the two age configurations seem to be broadly associated with illness severity. Those types of morbidity that are on average less serious (*ie.* acute and non-incapacitating illness) belong to the second group; whereas, chronic and incapacitating illness belong to the first. It is noticeable too that the age-inequality U-shape is more pronounced for bed days than it is for days of restricted activity. Though both are a form of *disability*, the latter involves a less severe loss of social function. Therefore, pro-rich levels of inequity are greatest in the middle years the more serious the type of illness. If, as seems quite plausible, prevalence of chronic conditions and functional incapacity are better predictors of underlying health status, the results point to a more rapid deterioration of health in those who are economically disadvantaged. Future policy initiatives should perhaps pay particular attention to this situation, not least because of the resource consequences it holds (*eg.* lost production, increased use of public health and social services, etc.).

---

[19] The indices for circulatory and muscoskeletal disease at age < 18 are not plotted in Figure 5.2 since they are clearly affected by the small number of individuals reporting illness (7 and 34 respectively). For all other indicators, there are at least 80 persons reporting illness in each age group.

Another important finding is that there appears to be no inequity in the distribution of morbidity among children and adolescents. Similar evidence is available for the U.K. (*eg.* West, 1988; Macintyre and West, 1991), though none of these studies uses equivalent income as the ranking variable. An obvious implication is that the common practice of excluding children in overall assessments of health inequity in a particular country may lead to an overstatement of the pro-rich bias. Some caution is, however, necessary over these results. It may be that population-based survey measures of morbidity are not particularly appropriate for young, generally healthy individuals; in which case, future work should attempt to develop improved morbidity indicators that better capture the health experience of children and adolescents. Furthermore, richer parents may show a greater propensity to report their children as ill. It is noticeable that, even though the associations between income and morbidity are insignificant, there is a positive sign for the non-incapacitating *illness* variable, but negative signs for the more objective disability days measures.

Finally, there is one morbidity indicator which does not fit well into either of the two age patterns identified above. The respiratory disease measure is distinct in showing a *pro-poor* bias among adults aged under 45. It is possible that, were confidence limits to be placed on the index values, the age configuration would not be too different to that of the *illness*, *acute* and *digestive* variables. However, the statistically significant relationship found for the 35-44 age group (see Table A4.4) provides some evidence that the pro-poor bias is real. One possible explanation for this result is the differential in smoking habits between income groups. Unlike northern european countries, Portugal still has a higher prevalence of smoking among the better-off. The differences appear to be particularly high among 35-44 year olds.[20] Given that general levels of morbidity are not very high among this age group, habitual smoking may be the main factor behind the observed pro-poor distribution.

---

[20] The following smoking prevalence rates by equivalent income quintile (from poorest to richest) were computed from the INS data for the 35-44 age group: 17%; 21%; 19%; 23%; and 31%.

5.4 The effect of different equivalence scale relativities

The rest of the chapter is dedicated to measuring the effect of different equivalence scale relativities on computed index values. This has two purposes: (i) to check the robustness of the earlier results, and (ii) to inform future distributional comparisons of a cross-national or intertemporal nature. The analysis is relevant to all studies that seek to measure income related health and health care inequalities; and in particular, those that adopt the line of research initiated by Wagstaff *et al* (1989). It hardly needs to be stressed that equivalence scale adjustments are but one of the measurement issues for which sensitivity analysis is advisable in these type of studies.[21] Nevertheless, the fact that the procedure is common to all measurements of economic inequality in the domain of health, while at the same time no general agreement exists over the 'correct' equivalence scale to use, suggests that the space given over to the issue in this chapter is justified.

The impact of equivalence scales on measurements of health inequality is not immediately obvious. The effect depends crucially on the relationship between health, income and the family characteristics that go to make up the scale. Coulter, Cowell and Jenkins (1992b) have provided a rigorous analysis of the impact of different scales on measurements of income inequality and poverty. They note that the effect on the Gini coefficient is particularly difficult to ascertain. This is because concentration measures, unlike most other indices of inequality, aggregate incomes on the basis of rank ordering. In our case, the scope for establishing *a priori* results is even more limited given the inter-relationship with health. All that can be said is that changing the value of the equivalence deflator may lead to changes in the rank order of individuals; and that these will generally lead to changes in measured inequality.[22] The direction of the effect on index values is not clear. It, therefore, needs to be established empirically.

---

[21] Others include: the choice of inequality measures (examined in this study by means of the parametric indices); the method of aggregating health care consumption in studies of inequity in delivery; the choice of cut-off point in multiple category morbidity indicators [see Wagstaff and van Doorslaer (1994)]; incidence assumptions in health care financing analyses; and so on.

[22] The exception is where the configuration of health profiles is such that re-ranking effects are cancelled out.

For reasons of tractability and clearer recognition of scale effects, I assume that all equivalence scales can be characterized simply in terms of family size and a single key parameter. Buhmann *et al* (1988) have shown that several scales currently used in empirical work - including those that are based on other family characteristics in addition to size - can be conveniently summarized in this manner. Their scale is

$$M_j = S_j^e \tag{5.4}$$

with S_j representing the size of *j*th family and *e*, the elasticity of family 'need' with respect to size. Income values are equivalized by dividing observed family incomes by M_j. Larger values of *e* correspond to smaller economies of size. A value of $e = 0$ implies no adjustment for size, while $e = 1$ corresponds to taking *per capita* income. Scales based on subjective evaluation of what is needed "to get along" (*ie.* the Leyden School approach) tend to produce relatively low values of the elasticity; those based on econometric analysis of consumption patterns or on the relativities implicit in social welfare payments produce intermediate values; and normative scales (which Buhmann *et al* call 'expert statistical'), are represented by high size elasticities typically greater than 0.70.[23]

Figure 5.3 shows the empirical effect on income related health inequality of assuming equivalence elasticities in the range 0 to 1. Computations were carried out for four of the morbidity indicators considered earlier: *bed days, off work days* and *limiting chronic* and *acute* disease. The data points shown are concentration index values for the all-sample unstandardized distributions.[24] The results show clearly that different equivalence scale relativities have an appreciable effect on inequality estimates. Making

---

[23] Unique representations of the Buhmann *et al* scale (or slight variations thereof) have been used in a number of income (re-)distribution studies (*eg.* Rainwater, 1992; Aronson *et al*, 1994); and also in health economic research [*eg.* the Italian, Spanish and Portuguese country studies in the volume by Van Doorslaer *et al* (1993)]. Coulter *et al* (1992a, 1992b) use the formula to measure the impact of scale relativities on computations of income inequality and poverty, in very much the same way as it is used here to measure the impact on illness concentration indices.

[24] The analysis was also repeated on the age-sex standardized distributions of the *bed days* and *off-work days* variables. This was found to have no appreciable effect on overall conclusions. Since the objective of the analysis is to illustrate the effect of scale relativities, rather than obtain precise measurements of the extent of inequity, all measurements shown are for the unstandardized distributions. The advantage of this option is that it is less demanding in terms of computations.

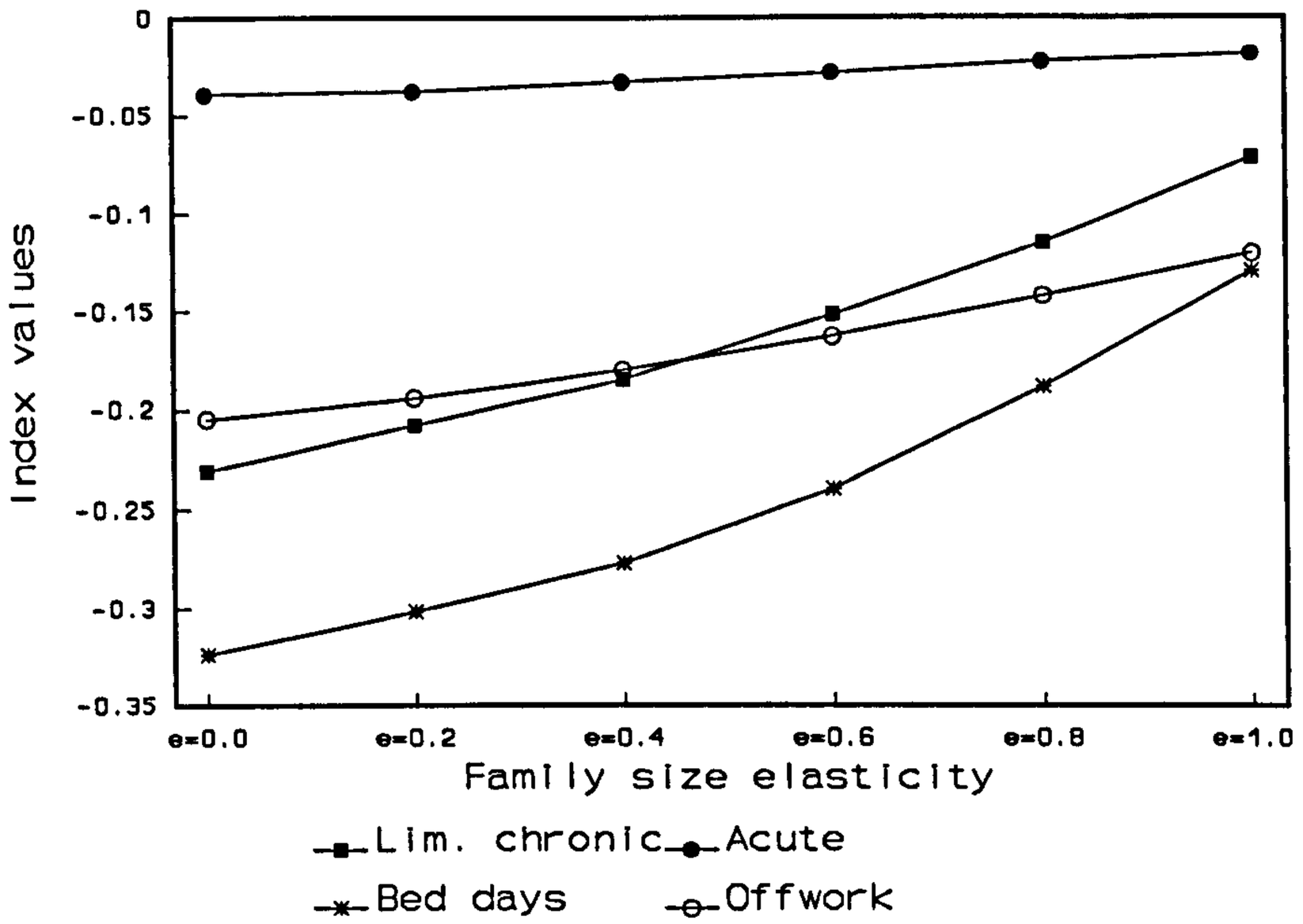


Figure 5.3: Impact of equivalence scale relativities on measured health inequity
Unstandardized concentration indices

no adjustment for family size leads to health inequality estimates which are between 70% and 220% greater than if *per capita* income was taken as the ranking variable. It is noticeable also that the inequality ranking of the *limiting chronic* and *off work days* variables changes as e is increased, suggesting that the earlier conclusions regarding inequality in different types of illness could change if a different equivalence scale had been adopted. Perhaps the most important feature of the results is that the extent of inequality declines quasi-linearly as the equivalence scales become more generous to larger families. This is in marked contrast to the effect on the Gini coefficient of equivalent income, the graph of which against e is distinctly U-shaped. The values for this measure in the range of elasticities shown in Figure 5.3 are: 0.358, 0.351, 0.345, 0.342, 0.347 and 0.357. A similar pattern is shown by Coulter *et al's* (1992a,b) calculations using the UK 1986 FES. Note also that the relative effects on computed index values appear much larger in the case of morbidity concentration coefficients. The difference between the maximum and minimum estimates for the Gini are only 5%.

These results may be challenged on the grounds that they are essentially driven by the categorical income variable available in the INS. Changing the value of the

equivalence elasticity changes the income ranking. If the number of income observations is small very few rank changes can have a pronounced effect on measured inequality. I cannot provide conclusive evidence on this issue, but there are a number of points which suggest that the findings are valid. First, the number of distinct income observations produced by the interpolation procedures is quite high. Though many people are left with equal income levels this is also true of real world situations. Second, if the induced reranking is thought to have a crucial influence on the disparity of health inequality estimates then it should affect the measure of income inequality in a similar manner. However, the 5% difference found for the Gini coefficient is less than the 8% reported by Coulter *et al* (1992a, p. 107) for continuous income data. Third, the results in this section are confirmed in Chapter 7, where a similar procedure is used to measure the impact of scale relativities on progressivity estimates. The analysis uses continuous income data from household budget surveys. The effect of increasing e on health care payments *concentration* indices is also to reduce the degree of measured inequality in a quasi-linear fashion. Generally, the impact of different scale relativities is quantitatively smaller than in the present case, but the most likely explanation is *not* the income measure as such, but the nature of the health or health care variable under analysis. The strong effect shown in Figure 5.3 may be due to the scale on which ill-health is measured. Rank changes in the income distribution will have pronounced effects on measured health inequality because the scale for morbidity may be a simple dichotomous category. Since most morbidity indicators are categorical rather than continuous, there is likely to be a significant impact on health inequality measures when equivalence scale relativities are changed. Finally, it is also the case that many health surveys are restricted to categorical income data. If researchers are to adequately measure income-indexed health inequality, then they have little choice but to adopt the interpolation procedures used in this chapter. Therefore, there is a case for accepting the findings of Figure 5.3 and indeed for extending the analysis to draw further implications for future research.

Although a systematic relationship between income related health inequality and equivalence scale relativities has been established, the analysis does not indicate conclusively if the earlier results are robust or not. The alternative scales that might have been chosen may well be characterized by a narrow range of e values. It is

unlikely, for example, that non-adjusted or *per capita* income would be used since both provide an unrealistic measurement of individuals' potential command over resources. On a more general level, it is also important to know if the effect on index values is likely to affect cross-country comparisons. It may be that observed differences in levels of health inequity between countries are so great that the choice of equivalence scale is irrelevant to final conclusions. One may also take the view that different countries (or the same country at different points in time) should indeed adopt different scales, reflecting variations in family 'needs'.

Definitive answers to these questions are beyond the scope of this study, but some light may be shed on them by further sensitivity analysis. Drawing on the work of Buhmann *et al* (1988), I calculated the implicit family size elasticities of the scales used in previous portuguese research and in the ECuity study as reported in van Doorslaer *et al* (1993). Values of the parameter *e* were estimated by ordinary least squares from the INS data file through the equation:

$$Log(E_j) = e\ Log(S_j) + \varepsilon\ . \qquad (5.5)$$

E_j represents the number of equivalent adults, S_j family size and *e*, the equivalence elasticity. The full results are shown in the Appendix as Table A4.5, together with a description of each equivalence scale. Judging by the values of the computed R-squares, the Buhmann *et al* formula does indeed provide a reasonable empirical approximation to all equivalence scale relativities. The explained variance for six of the scales is over 98% and only slightly less for the other two. Age gradation in the scales appears to make no difference to the goodness of fit of the regression line.

The estimated *e* values for the scales used previously in portuguese research are 0.77 for the OECD scale; 0.82 for the ILO scale; and 0.78 and 0.86 for the Santos (1984) econometric scales. Clearly, the family size elasticities of these scales are very similar, suggesting that measurements of income-indexed health inequity are unlikely to be greatly affected by choosing any of the four.[25]

---

[25] It is interesting to note that the estimate of *e* for the OECD scale is greater than that obtained by Buhmann *et al* who report a value of 0.73. This is due to different distributions of family size in the samples used. I replicated the analysis on both the 1980/81 and 1989/90 Portuguese *Family Budget Surveys* and the same value of $e = 0.77$ was revealed. The most likely explanation is that average

Table 5.4: Implicit (or explicit) family size equivalence elasticities adopted in the ECuity study

Country study	Scale used	Fam. size elasticity
Denmark	OECD	0.769
France	OECD	0.769
Ireland	Irish scale	0.677
Italy	None	0.000
Netherlands	Dutch scale	0.454
Portugal	Buhmann *et al* formula	0.600
Spain	Buhmann *et al* formula	0.400
Switzerland	Swiss scale	0.590
United Kingdom	CSO/DSS	0.690

The elasticities of the scales used in the ECuity study are summarized in Table 5.4. Two countries (Spain and Portugal) used explicit values of the Buhmann *et al* elasticity to adjust the distribution of incomes in their surveys. The Italian study used a value of $e = 0.40$ in the finance side analysis, but left family incomes unadjusted for the delivery side due to lack of information on family structures. France and Denmark used the OECD scale; and the remaining countries used national scales. There is a much larger variation in the e values of the ECuity scales than was found for the "Portuguese" ones. Note, however, that if adjustments based on arbitrary values of e are excluded, the range of elasticities is (0.454, 0.769).[26]

---

family size is higher in the Portuguese samples than in that used by Buhmann and others. Similar exercises on different data sets are also likely to reveal small variations in the estimates of e.

[26] A recent paper by the ECuity group (Wagstaff *et al*, 1994) has used a parametric scale suggested by Aronson *et al* (1994). This scale is defined as

$$M_j = (s_A + \phi s_C)^\theta , \qquad 0 \leq \phi \leq 1, \ 0 \leq \theta \leq 1$$

where s_A is the number of adults in the family, s_C the number of children, and ϕ and θ are parameters, the former determining the importance of children and the latter an equivalence elasticity with

As a final step in the analysis, illness concentration indices were computed by assuming each of the eleven different income adjustments used in portuguese research and in the ECuity study. The results are shown in Table 5.5. Consider first the implications of these measurements for the question of robustness of the earlier results. The obvious frame of reference is the results produced by the other "portuguese" scales. These show that the degree of pro-rich inequity is slightly less than that estimated under the OECD assumptions. However, the effects are not strong enough to warrant changing any of the conclusions reached by the earlier analysis. The measurements for the *bed days, off work days* and *limiting chronic* variables continue to show relatively high degrees of inequality, whereas the *acute* indicator still shows a very slight pro-rich bias. It may be argued, nevertheless, that none of the "portuguese" scales provide an accurate picture of current equity-relevant non-income differences between persons. This was the view taken by the portuguese report in the ECuity study, which used a value of $e = 0.60$ to adjust family incomes (Pereira and Pinto, 1993). It was conjectured that the Santos (1984) scales were unduly generous to larger families, because the econometric analysis was based solely on food and clothing expenditures. If other items of consumption had been included (*eg.* housing, energy, transportation, etc.) economies of scale would in all probability have been found to be higher. While the argument is relevant, the choice of $e = 0.60$ may seem too strong a departure from other scale relativities used in portuguese research. Such a judgement is, of course, only possible now that the implicit e values have been estimated. Nevertheless, it is reassuring to find that for the four morbidity variables considered, the earlier conclusions are still valid. At $e = 0.60$, the indices show values between 12 and 23 per cent greater than under the OECD assumptions, but the ranking of the indicators remains the same. Therefore, it is fair to say that the earlier results are robust in terms of other probable equivalence scale relativities.

With regard to the implications for cross-country comparisons, the first question that needs to be asked is whether the impact of changing e is quantitatively significant? An instructive comparison can be made with the variation in estimates of health

---

analogous properties to Buhmann *et al*'s e. Wagstaff *et al* (1994) set both parameters equal to 0.5. I calculated the implicit e-value of this scale using the INS data and found $e = 0.439$ ($R^2 = 0.986$).

Table 5.5: Impact of equivalence scale relativities on measured health inequity "Portuguese and ECuity study scales. Unstandardized concentration indices

Equivalence scale	Bed days	Offwork days	Limiting chronic	Acute
"PORTUGUESE" SCALES				
Santos I	-0.164	-0.131	-0.097	-0.022
Santos II	-0.199	-0.145	-0.121	-0.023
ILO	-0.178	-0.139	-0.108	-0.022
OECD	-0.203	-0.146	-0.124	-0.023
Percentual difference from OECD scale				
Santos I	80.8%	89.4%	77.9%	92.8%
Santos II	97.7%	99.1%	97.5%	99.7%
ILO	87.5%	95.2%	86.4%	95.4%
OECD	100.0%	100.0%	100.0%	100.0%
*EC*UITY STUDY				
Italy	-0.324	-0.205	-0.231	-0.040
Spain	-0.278	-0.181	-0.185	-0.033
Netherlands	-0.270	-0.178	-0.178	-0.032
Switzerland	-0.240	-0.164	-0.152	-0.029
Portugal	-0.240	-0.163	-0.152	-0.028
Ireland	-0.222	-0.154	-0.137	-0.025
United Kingdom	-0.218	-0.152	-0.134	-0.024
Denmark & France	-0.203	-0.146	-0.124	-0.023
Percentual difference from OECD scale				
Italy	159.3%	139.9%	185.3%	170.5%
Spain	136.7%	123.5%	148.8%	143.0%
Netherlands	133.0%	121.4%	142.9%	138.7%
Switzerland	118.3%	111.8%	122.4%	123.6%
Portugal	118.3%	111.7%	122.2%	122.8%
Ireland	109.1%	105.0%	110.1%	108.3%
United Kingdom	107.0%	103.6%	107.9%	104.5%
Denmark & France	100.0%	100.0%	100.0%	100.0%

inequality for nine european countries presented in van Doorslaer *et al* (1993).[27] The absolute differences between the lowest and highest C_m values reported in that study are 0.14 for chronic illness and 0.20 for self-assessed health and limiting chronic illness. The latter indicator is particularly relevant since it is common to both studies. Table 5.5 suggests that, for a given joint distribution of health and income, the range of scale relativities adopted by the nine countries can itself lead to an absolute variation of 0.11 points in inequality estimates. This is more than half of the observed inter-country variation in Van Doorslaer *et al* (1993). Furthermore, there is a strong possibility that rankings would change if particular countries adopted a scale other the one chosen. For example, Denmark which has a mid-ranking score on the chronic illness indicator might well emerge as the most inequitable country if its incomes were deflated by, say, the Dutch scale; Switzerland, which has less inequality than Italy for self-assessed health, would very likely reverse positions if it adopted the Italian scale. Therefore, the quantitative impact of scale relativities on cross-country comparisons is far from negligible.

A more important question, however, is whether these effects are conceptually significant? Should we view the scale-induced variation as a form of statistical bias or as the inevitable result of differing circumstances between countries? The answer depends crucially on whether the scales chosen to deflate each country's incomes are an accurate representation of the relationship between income, family composition and levels of economic well-being in different countries. If they are, then use of country-specific scales is legitimate. However, it is not too difficult to raise doubts about the choices actually made in van Doorslaer *et al* (1993). Six countries (DK, F, I, P, E, and CH) used scales which bear no *known* empirical relationship to the joint distribution of income and 'needs' among their citizens. Of the remainder, the Irish scale is based on the relativities implicit in that country's welfare payments. However, as Coulter *et al* (1992a, pp. 99-101) argue, such scales may also be inaccurate (*eg.* they vary substantially according to policy choices rather than actual 'needs'; are typically inconsistent with relativities implicit in other parts of the tax and transfer system; and seem more appropriate for distributional assessments concerning the bottom tail of the

---

[27] See Table A4.6.

income distribution rather than its full range). This leaves the UK and Dutch scales which were in fact derived empirically from budget data. As mentioned earlier, however, this apparent advantage may not mean a great deal: econometric equivalence scales also rest on potentially controversial normative judgements and their results clearly depend on the measurement approach and particular model chosen [Bradbury (1989); Coulter *et al* (1992a, 1992b); Nelson (1993)].

In short, no *single* scale used by the ECuity researchers can categorically be assumed to truly represent the underlying relationship between income and 'needs' in a particular country. This does not mean that the scales were wrongly chosen. Rather, it accepts that the state of the art is such that a 'correct' equivalence scale is well-nigh impossible to find. "A range of equivalence scales is ... not only inevitable but also legitimate" [Coulter *et al* (1992a, p. 79)]. In these circumstances, and given that different scale relativities quite evidently contribute to observed variations in health inequality estimates, it is advisable that future cross-national comparisons seek to assure that the results are not unduly affected by those differences.

What then should be done? Recent cross-national studies of income distribution have invariably opted to use a common scale [*eg.* O'Higgins *et al* (1990); Nolan and Callan (1992); Rainwater (1992)]. However, this approach seems too restrictive. The obvious recommendation is that a major role be assigned to the parametric scales examined in this section. Quite how that role is materialized is a matter for future comparative studies. Calculations may be carried out for a range of scale relativities *ab initio*, or the parametric scales may be used to check the robustness of results derived by assuming country-specific or common scales. The main drawback is the increased number of computations that are required. However, this seems a price worth paying in order to achieve more robust results. In fact, the extra computational effort may not be too great. If the evidence for a quasi-linear relationship between health inequality estimates and scale relativities is generalizable, then checks for robustness may be carried out by choosing two plausible extreme e values.[28] Of course, a great many problems will remain to be solved if well-informed cross-national

[28] On the basis of the estimates shown in Table 5.4, those values may be set at 0.45 and 0.77.

comparisons of health domain inequity are to be made. The area of health inequality is notoriously problematic owing to cultural influences on the decision to report illness. However, by tackling the equivalence scale issue, at least one major area of uncertainty will have been removed.

5.5 Conclusions

The conclusions to this chapter may be stated simply enough. First, strong evidence has been found showing that the distribution of illness among the Portuguese population is generally unfavourable to poorer income groups. Second, the overall picture masks some important counteracting trends. For example, less serious illnesses show virtually a uniform distribution across income groups whereas more serious conditions are disproportionately concentrated among the less well-off. Also, the age dispersion of health inequality shows that poor middle-aged individuals are particularly disadvantaged. In contrast, there appears to be no economic differentiation in the health experience of children and adolescents, although this evidence may in part be due to the inappropriateness of the available morbidity variables for that sub-group of the population. Third, the overall results are robust in terms of two crucial methodological choices: the method of weighting income groups when aggregating inequality and the equivalence adjustments made to the income variable. Finally, it has been shown that a degree of caution is necessary when interpreting cross-national results based on different equivalence scale relativities. For a given joint distribution of health and income, the extent of measured inequity varies considerably according to equivalence scale generosity. The implications of these findings are discussed in the concluding chapter.

Chapter 6

Measuring Achievement II: The Time-trend of Inequity in Infant Mortality

---

"By the year 2000, the actual differences in health status between countries and between groups within countries should be reduced by at least 25%, by improving the level of health of disadvantaged nations and groups."
Target 1 of the common health policy adopted by WHO European Region member countries in 1980.

6.1 Introduction

Measurement of inequity in health is at its most useful when the prevailing time-trend can be established. The single year estimates presented in the previous chapter can signal the need for policy intervention (by virtue of the degree of inequity being judged undesirable); but they provide no way of knowing whether past or present measures are in any way successful. In contrast, a time series of inequity measures can inform the policy-maker if the country is moving in the desired direction, or not. It may also suggest the type of health and social phenomena that underly *changes* in the distribution, thereby indicating possible corrective measures.

In Portugal, as in most other countries, the only means currently available for examining the time-trend of inequities in health is through data on mortality. The use of such data is not without its drawbacks. On the one hand, information on deaths does not indicate health among the living (in particular, for diseases where the case-fatality ratio is low, the mortality rate will be a gross underestimate of the incidence of the condition in the community); and on the other, the scope and accuracy of socio-economic information collected at the time of death are somewhat limited. With the increasing availability of large scale micro-data bases, which include information on

morbidity that may easily be crosstabulated with demographic and economic variables, there has naturally been a trend away from using mortality data. This is particularly true of work by economists.[1] However, information on deaths does have some advantages. First of all, it provides detail on a particular (and important) dimension of the social concern for health, that of the length of life. Mortality statistics are, therefore, complementary to the morbidity indicators used in Chapter 5, which may be interpreted as measuring the "healthfulness" of life (Jazairi, 1976). Secondly, unlike morbidity data they do not raise the problem of cultural variability in the rates reported by (or on behalf) of different socio-economic groups (d'Houtaud and Field, 1984). The actual occurrence is accurately and completely documented in developed countries, although some variability remains in the accuracy of cause-of-death and demographic background of the deceased. Finally, it is also the case that specific types of mortality are highly correlated with increased risks of ill-health in survivors, when both are disaggregated by individuals' socio-economic characteristics. According to Tooley (1966, p. 18) and Blaxter (1991, pp. 34-35), this is especially true of mortality among infants. Therefore, examining inequities in death may, in some circumstances, be an informative means of establishing inequities in health.

The present chapter concentrates precisely on the time-trend of inequity in infant deaths (*ie.* children under 1 year). This age group has long been recognized as highly vulnerable to the adverse living conditions that go hand in hand with poverty. Numerous epidemiological studies, both at the aggregate and micro levels, have added support to this view (*eg.* Blaxter, 1981; Paneth *et al*, 1982; Nersesian, 1988; Stockwell *et al*, 1988). In Portugal, the issue has attracted a good deal of attention from researchers and policy-makers. The rapid decline in infant mortality rates witnessed throughout the 1970's and 80's is often put forward as an example of improved levels of general well-being, not least by government sources. Doubts have been raised, however, as to whether the decline was accompanied by improvements in the socio-economic and spatial distribution of those rates. All research results to date point to the equity gains having been negligible (*eg.* Kannisto; 1986; Leitão, 1988). These studies,

---

[1] The work of Le Grand and associates cited earlier (Le Grand and Rabin, 1986; Illsley and Le Grand, 1987; Le Grand, 1987) makes extensive use of mortality data, but it does not address the issue of socio-economic differences. It simply measures inequality in age-at-death.

however, have some important drawbacks discussed in Chapter 2. They cover limited time-spans; do not distinguish between the different types of infant mortality; make limited use of statistical techniques available for summarizing distributions; or simply fail to address the issue of socio-economic differences.

The present analysis maintains the ethical base of Chapter 3, identifying inequity with income-indexed mortality differentials. By implication, the concentration curve methodology is also adopted. Routinely published data are used throughout and an empirical approach is presented that may be useful for countries with an insufficiently developed information base. I first set out the methodology (section 6.2) and then present the results (6.3). Section 6.4 discusses the main implications of the findings.

6.2 Methodology

The empirical approach is based on the use of geographical observations *ranked* by economic position. The infant population is first disaggregated by mother's place of residence, and then ranked by an index of average income levels in the spatial units. The cumulative percentiles of these observations together with the respective infant death percentiles serve as inputs to the construction of concentration curves. The ranking element assures that we are measuring socio-economic differences, albeit at a high level of aggregation, and not merely geographical inequalities.

The procedure is justified by the incompleteness and inaccuracy of more direct alternatives. For example, the obvious approach to the problem would be to rank the infant population by some proxy of families' economic circumstances, such as the father's profession or mother's level of educational achievement, both of which are recorded on Portuguese death certificates. This method is adopted by Wagstaff *et al* (1991a), who use data grouped by occupational class, to compute standard concentration coefficients for *adult* mortality in England and Wales, Finland and Sweden. However, in Portugal the quality of socio-economic information on death certificates is known to be particularly poor. The National Statistical Institute (INE) only began publishing crosstabulations of deaths by occupational class in 1987, and then only for adults of

working age. Furthermore, even if the necessary data were readily available, it would then have to be linked with data on births, decomposed by the same categories. These statistics are also generally not available. Even if they were, it is by no means certain that this would provide a satisfactory basis for examining inequity. The linking of occupational class data from two sources (recorded in quite different circumstances) is, after all, one of the main reasons commonly advanced for doubting the value of mortality-based inequity analyses (Power *et al*, 1991, p. 17).

The present approach has no such drawbacks. Information on mothers' place of residence is generally reliable and complete on both death and birth certificates. The economic ranking of geographical units is an area that is relatively well researched. It needs to be stressed, nevertheless, that the relationship between mortality and income levels is being measured at a high level of aggregation. Persons living within the geographical units will obviously have unequal command over resources, to which may be associated further inequalities in mortality. If this is true, then the computed indices are likely to be biased, showing *lesser* inequity than is in fact the case. Yet this is a problem that is shared by studies that measure socio-economic status directly. Virtually all work on mortality has used grouped data, usually by occupational class. Within those classes there will also be inequalities. Therefore, given the underdeveloped information base and the limitations of alternative methods, there seems to be good reason for pursuing the current approach.

The empirical analysis considers four conventionally used components of infant deaths:

(i) *Perinatal mortality*. Stillbirths occurring at 28 weeks or more of gestation and deaths of live born infants occurring in the first week of life.

(ii) *Neonatal mortality*. Deaths occurring in the first four weeks of life;

(iii) *Post-neonatal mortality*. Deaths occurring from the fifth week of life up to one year; and

(iv) *Infant mortality*. Deaths occurring in the first year of life (*ie*. the sum of neonatal and post-neonatal mortality).

It is common to express the perinatal mortality *rate* per thousand live births and stillbirths at 28 or more weeks of gestation, and the *rates* for the other three components per thousand live births. The main reason for considering the components separately is that they are differentially affected by specific causal factors (WHO, 1970; Jazairi, 1976; Blaxter, 1981; Edouard, 1985). In particular, neonatal and perinatal deaths are more sensitive to hereditary factors, the natural environment and the impact of ante-natal and obstretric and perinatal care. Post-neonatal deaths, on the other hand, mainly reflect poor conditions in dwellings, bad nutrition and lack of sanitation and hygiene - in short, the type of situations that are associated with poverty.

The unit of observation is the *district* (n = 18), which is the main geographical entity used for administrative and statistical purposes in Portugal. The period of analysis is comprised by the years 1971 to 1991 for the infant mortality indicator and 1976 to 1991 for the other three components.[2] As Table 6.1 shows, there are considerable differences in death rates between the districts, both at the beginning and end of the period. For example, in 1971 the infant mortality rate varied between 25.9 and 65.7 deaths per 1000 live births. By 1991, there had been a marked decline in rates throughout the country, but the highest recorded value (20.5) was still almost three times greater than the lowest (7.4). The ratios between extreme values are greater for the neonatal and postneonatal components and smaller for perinatal mortality.

The economic ranking of the observations was established via a 3-year moving average of an index of wages in the construction industry, derived from a continuous employment survey carried out by the INE.[3] This is the *only* indicator of living standards in the districts that is available for all years under consideration. Although the nature of the index might raise doubts as to its ability to proxy the underlying

---

[2] The different periods of analysis are due to inconsistencies in the published data. For some years in the period 1971-75, the published statistics for the neonatal, post-neonatal and perinatal components refer to the *de facto* district of death/birth. I use only data broken down by the mother's place of residence. The data sources are: INE, *Estatísticas Demográficas* and INE, *Estatísticas de Saúde*, for the period 1971-86; and Ministério da Saúde (DGCSP), *Natalidade, Mortalidade Infantil e Mortalidade Perinatal 1987/91*, for the remaining years. The district borders are shown in the map presented as Appendix 2.

[3] The source is the *Anuário Estatístico* (various years).

Table 6.1: Mortality rates at the endpoints of period of analysis

		Infant	Neonatal	Post-neonat	Perinatal
1971/76* -	National rate	48.3	19.8	13.1	29.1
	Maximum	65.7	30.4	30.5	39.3
	Minimum	25.9	8.0	8.7	15.7
1991 -	National rate	10.7	6.9	3.8	12.0
	Maximum	20.5	10.6	9.9	14.5
	Minimum	7.4	3.3	2.1	6.9

* The infant mortality rate refers to 1971 and the neonatal, post-neonatal and perinatal rates to 1976.

variable of interest, concordance tests with other possible indicators for the years when these are available revealed fairly strong positive associations. Values of the Spearman correlation coefficient, Kendall's Tau-b and Goodman and Kruskal's Gamma in the range {0.65, 0.84} were computed for associations between the rank of the chosen index and those of 3 standard of living indices (census- regional product- and market research-based) in selected years when these were available. This suggests that the analysis is robust in terms of choice of available income proxies.[4]

The concentration curves and indices were estimated by using data on the numerators (number of deaths) and denominators (live births and still-births) of the four mortality indicators. For example, the infant, neonatal and post-neonatal mortality curves were constructed by measuring, on the horizontal axis, the cumulative percentage of live births with the districts ranked from poorest to richest; and on the vertical axis, the cumulative percentage of deaths occuring to infants of mothers residing in the

[4] Nevertheless, actual inequality estimates based on a more general measure of living standards are provided in section 6.4. It needs to be stressed that the construction industry wage variable is used because there are no other indicators available for all years under analysis.

income-ranked districts. A visual representation is available in Figure 6.2, which presents the observed concentration curves in the years 1976 and 1991. For the case of perinatal mortality, one obviously has to add still births to live births and deaths in order to calculate the percentiles. The standard concentration indices (C_m) were computed using the linear approximation formula given in equation (4.19), with p_j representing the cumulative proportion of live births (plus still births in the case of perinatal mortality) and m_j representing the cumulative proportion of respective deaths (plus still births for perinatal mortality).

6.3 Results

Table 6.2 presents the series of standard concentration indices calculated for the four infant death components. All the indices have negative signs, suggesting that lower levels of average income are associated with higher mortality rates. The burden of infant deaths is, therefore, generally unfavourable to the poorer districts. There is also a clear pattern in the relative values of the four indicators. Apart from a short period in the late 1980's - which detailed inspection of the data suggests is not an artefact effected by the ranking variable - inequity is always greatest for post-neonatal mortality and least for the perinatal and neonatal components. Thus, the closer one gets to birth the smaller the degree of inequity. It is, however, worth noting that a larger number of deaths are associated with these smaller income differentials (see Table 6.1).

The most salient finding is, nevertheless, the significant decline in the degree of inequity over the two decades. For example, the concentration index for infant mortality, which stood at -0.117 in 1971, had by the end of the period risen to -0.030. Between 1976 and 1991, the absolute values of all component indices are at least halved, suggesting that the WHO precept cited earlier has already been surpassed. The average annual rates of change (reported in the lower part of Table 6.2), suggest that the greatest relative 'equity gains' were achieved for the neonatal and perinatal components. Indeed, by 1991 there is virtually no income related inequity in the distribution of these deaths. It is also clear that in the late 1970's, the overall decline in inequity levels is primarily attributable to improvements in the post-neonatal distribution, but that in the

Table 6.2: Standard concentration coefficients.
Infant mortality and components, 1971-1991

Year	Infant	Neonatal	Post-neonatal	Perinatal
1971	-0.117	na.	na.	na.
1972	-0.102	na.	na.	na.
1973	-0.119	na.	na.	na.
1974	-0.103	na.	na.	na.
1975	-0.109	na.	na.	na.
1976	-0.080	-0.028	-0.158	-0.046
1977	-0.098	-0.068	-0.148	-0.062
1978	-0.097	-0.063	-0.149	-0.066
1979	-0.080	-0.034	-0.153	-0.031
1980	-0.068	-0.054	-0.092	-0.073
1981	-0.091	-0.075	-0.124	-0.071
1982	-0.078	-0.056	-0.129	-0.052
1983	-0.065	-0.043	-0.109	-0.034
1984	-0.036	-0.030	-0.047	-0.038
1985	-0.034	-0.023	-0.058	-0.006
1986	-0.055	-0.061	-0.043	-0.041
1987	-0.058	-0.044	-0.087	-0.012
1988	-0.048	-0.050	-0.045	-0.055
1989	-0.063	-0.070	-0.048	-0.033
1990	-0.050	-0.032	-0.082	-0.008
1991	-0.030	-0.007	-0.072	-0.001
Average rate of annual change				
1976-91	3.9%	4.7%	3.4%	6.1%
1971-75	1.3%	na.	na.	na.
1976-80	3.0%	-18.4%	8.3%	-11.9%
1981-85	12.6%	14.0%	10.6%	18.3%
1986-91	7.6%	14.7%	-11.2%	16.2%

na. = not available

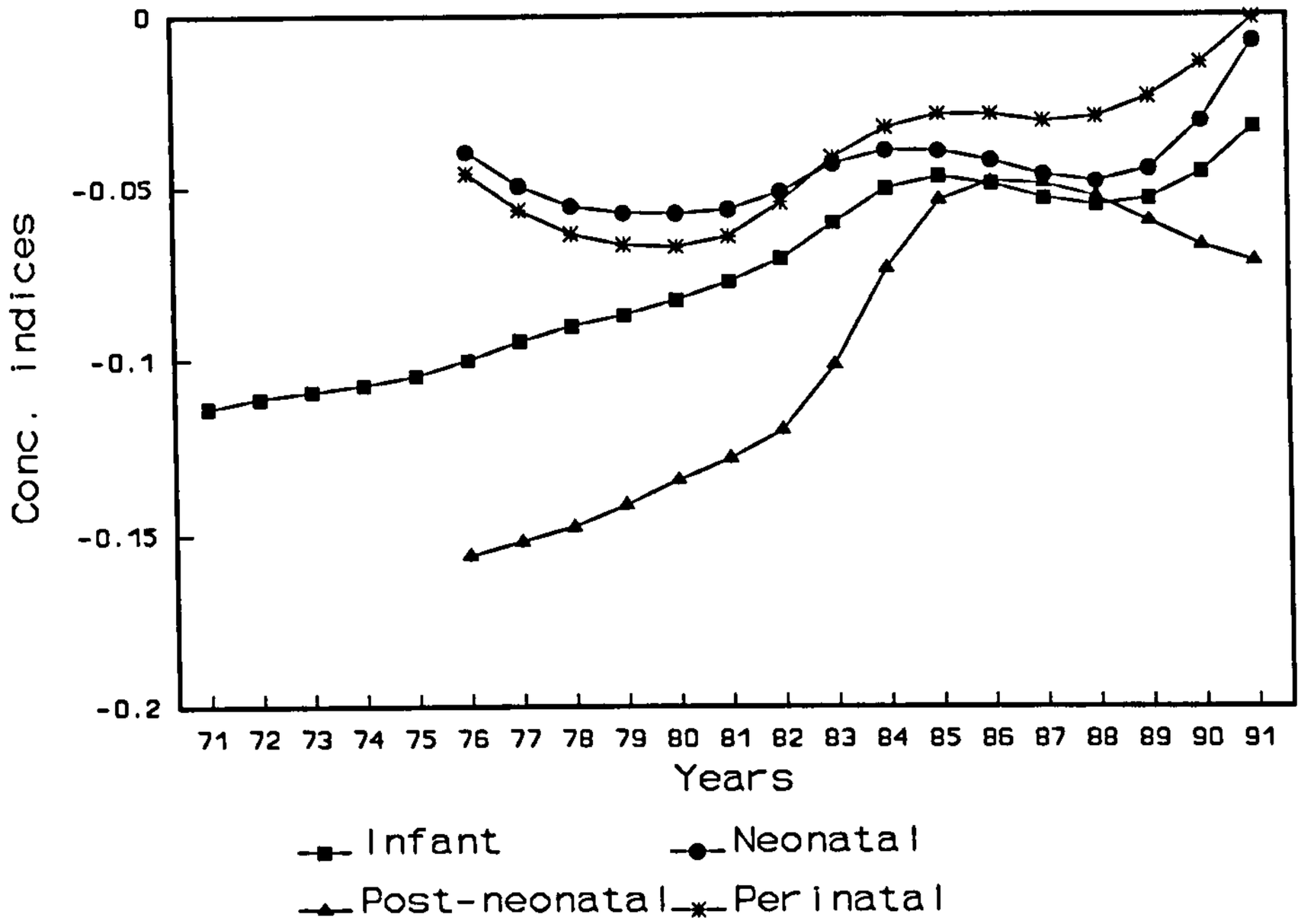


Figure 6.1: Smoothed time-trend of standard concentration indices. Infant mortality and components. 1971-1991.

1980's this role is taken over by the close-to-birth components.[5]

The time-trend is best examined through Figure 6.1, where smoothed plots of the index values are presented.[6] Pro-rich inequity in *infant* mortality shows a steady decline up to the mid-eighties; it then stagnates at $C_m \approx -0.06$, but there is a renewal of the declining trend after 1989. This general tendency is the result of quite distinct movements in the component indices. The distribution of post-neonatal mortality is initially very unfavourable to the poorer districts. However, the degree of inequity falls rapidly up to the mid-eighties, whereupon the tendency is inverted. By contrast, pro-rich inequity in the neonatal component, which starts off at a much lower level, hardly

---

[5] The relative weight of component mortality does not change greatly throughout the period. Post-neonatal deaths accounted for 40% of all infant deaths in 1976 and 36% in 1991.

[6] Smoothing of the time series was carried out via a technique sometimes referred to as *T4253H smoothing*. It involves repeated smoothing with running medians of size 4, 2, 5 and 3, followed by hanning (running moving averages). The procedure is available on some statistical packages (*eg.* SPSS for Windows).

changes up to the late 80's. It is, however, the moving force behind the renewed overall decline at the turn of the decade. Noticeably, this recent inflexion in the neonatal tendency is also apparent in the perinatal distribution, which is not surprising given the degree of overlap in the indicators.

An explanation for these trends clearly requires evidence based on multivariate analyses. However, bearing in mind the differential impact of causal factors referred to earlier, it is quite plausible that the early sharp fall in post-neonatal mortality inequity is the result of improved living conditions, particularly in the poorest regions. Prior to the mid-1970's, the remotest parts of the country lacked basic sanitation and income was very unevenly distributed. With the advent of democracy in 1974, major improvements were rapidly brought about in these areas, and these may have had a decisive impact on the distribution of post-neonatal deaths. By the same token, the recent inversion of the tendency may be a result of the well-documented increase in poverty in the mid-eighties (*eg.* Franco, 1990). There is also a plausible explanation for the recent decline in neonatal and perinatal inequity in terms of improved access to health care. Since the early 80's, a number of new hospitals have opened in previously underserved areas and the spatial distribution of ante-natal care facilities has also improved (Abel-Smith, 1992; Urbano *et al*, 1993). It may be that these factors provided the necessary stimulus to make the distribution of close-to-birth mortality insensitive to underlying income differences (as reflected by the income-ranked geographical distribution).

The concentration curves at the beginning and end of the period shed further light on the time-trend of the distributions (Figure 6.2). In 1976, the distribution of infant, neonatal and post-neonatal mortality unambiguously favours the better-off districts.[7] Furthermore, the post-neonatal distribution dominates the infant distribution which in turn dominates that of the neonatal period. In 1991, the dominance relationship is maintained, but there is a conspicuous decline in pro-rich inequity. Indeed, the curves for both infant and neonatal deaths now actually cross the 45° line. It is therefore inadvisable to conclude that, in 1991, there is still a pro-rich bias in these distributions.

---

[7] This is also true of perinatal mortality. Its concentration curve is not shown for presentational reasons, but it closely follows the configuration of the neonatal curve.

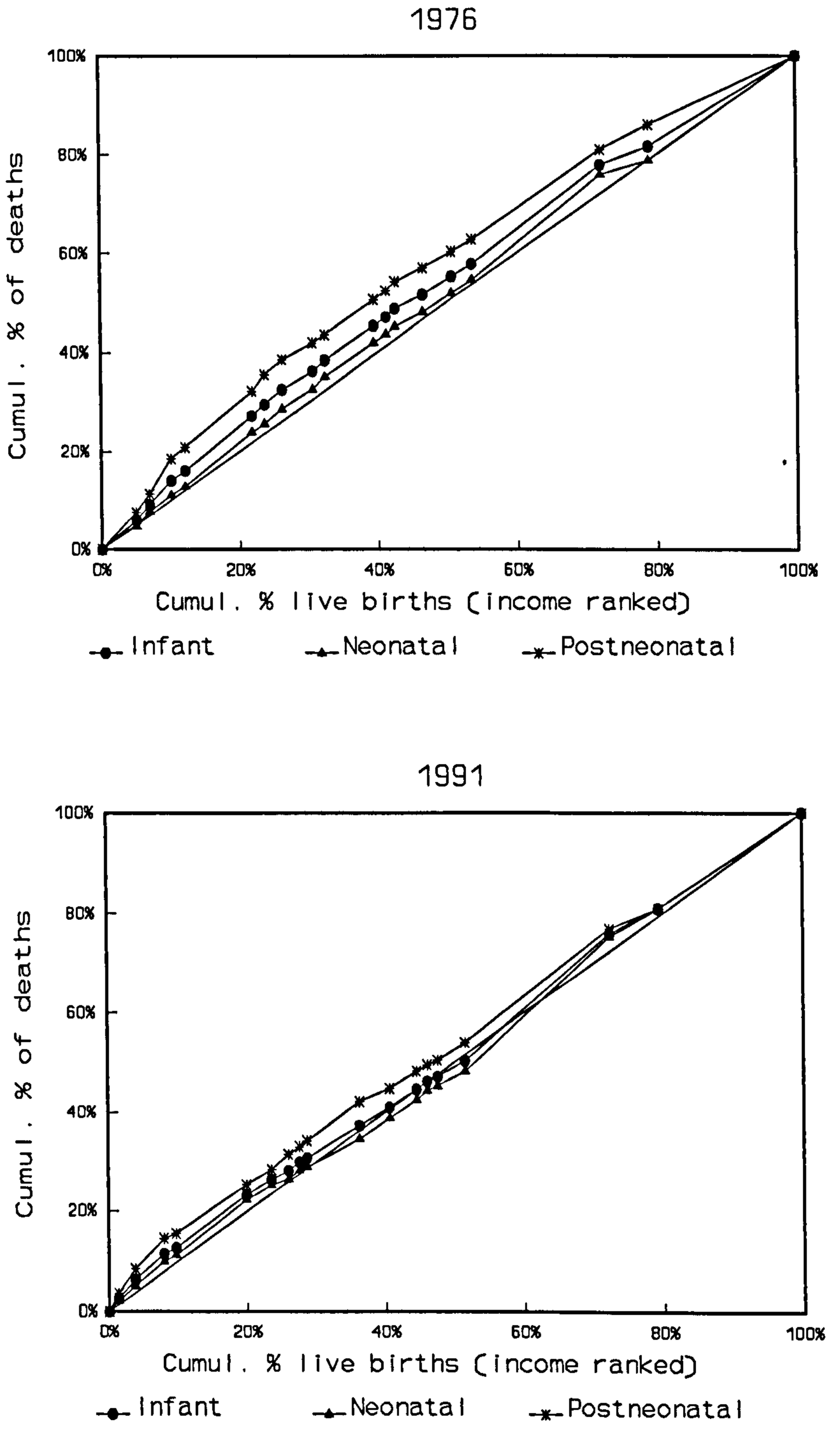


Figure 6.2: Concentration curves for infant, neonatal and post-neonatal mortality. 1976 and 1991

Even so, the fact that the curves initially cross the diagonal from *above*, means that the very poorest districts continue to experience a comparatively high burden of mortality. Therefore, it would seem that the equity gains observed throughout the period are primarily accounted for by districts where average levels of income are highest.[8]

In order to check the robustness of the results in terms of alternative distributional judgements, generalized concentration coefficients (with δ set at values between 1.0 and 5.0) were also computed. Given that, in this case, the computational requirements are considerably more demanding, the analysis is restricted to five year intervals. The results are shown in Table 6.3 and Figure 6.3, which plots the index values presented in the Table.[9] [10]

As expected, all indices tend to zero when the distributional judgement parameter is set at $\delta \rightarrow 1$ (computed for $\delta = 1.001$). The observer's indifference to economic related inequality leads him to infer that the distribution of mortality favours neither rich nor poor. At $\delta = 2$ the indices are equivalent to the standard coefficients reported in Table 6.2. As the parameter is increased beyond $\delta = 2$ (reflecting greater weight being given to the lower tail of the income distribution), successively greater degrees of pro-rich inequity are shown. For example, the sequence of increases in the 1976 perinatal distribution is: -0.046 at $\delta = 2$; -0.071 at $\delta = 3$; -0.088 at $\delta = 4$; and -0.102 at $\delta = 5$. However, it is noticeable that increased levels of inequality aversion have a differential impact, depending on the type of mortality under consideration.

The degree of measured inequity in the post-neonatal distribution is particularly

---

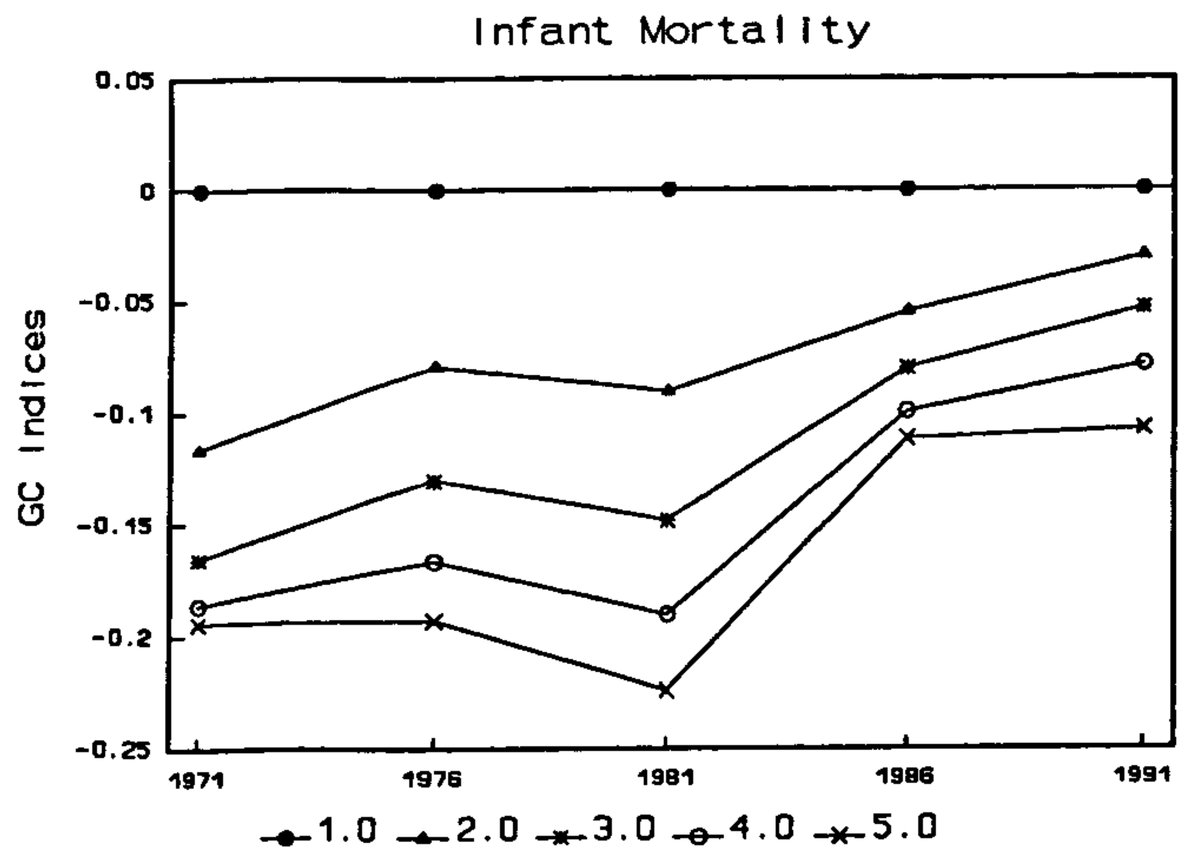
[8] The exception to this trend is the ante-penultimate observation (the Oporto district), which has both high living standards and high levels of mortality. Given that it also has the largest number of births, a further crossing of the diagonal is induced. This finding - allied to the fact that widespread social inequalities are known to exist *within* the Oporto district - suggests that the geographical classification may need to be further disaggregated in order to serve as a better proxy of income differences among the Portuguese population.

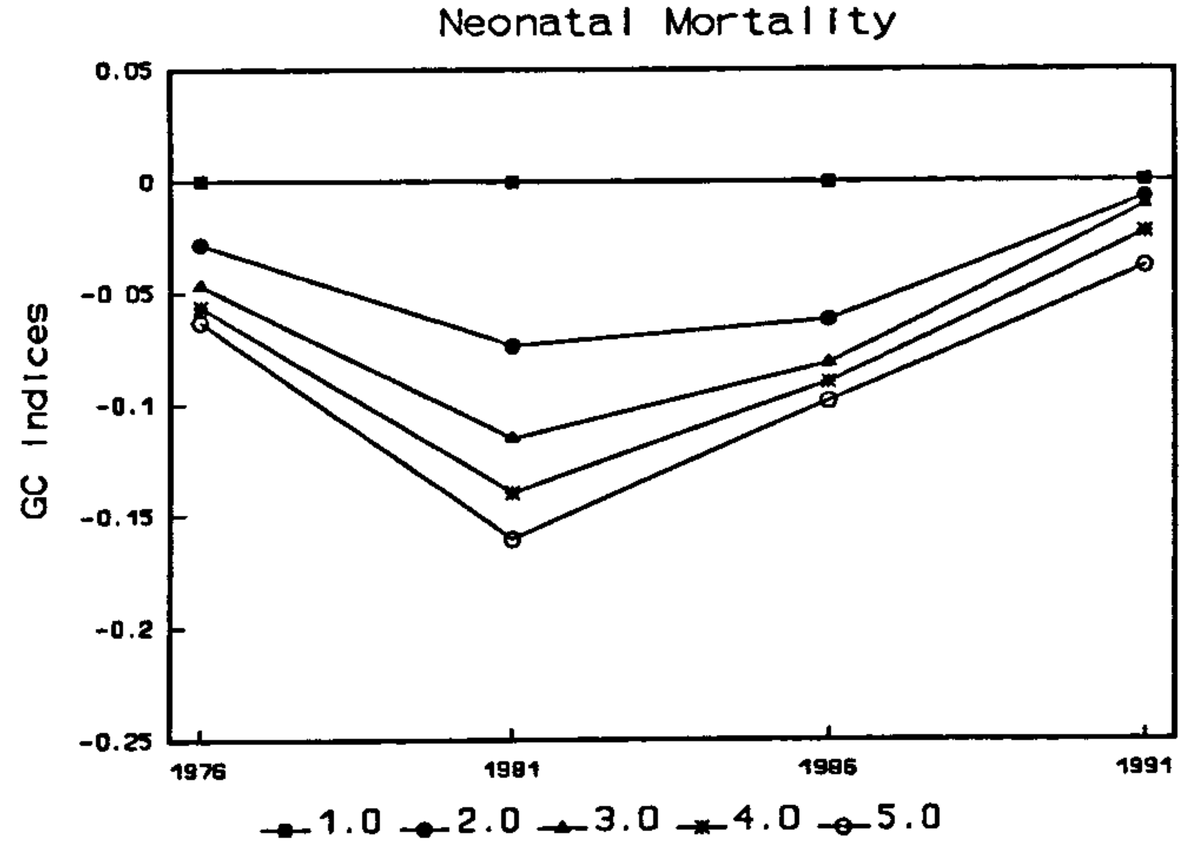
[9] Note that the post-neonatal diagram in Figure 6.3 has a different scale.

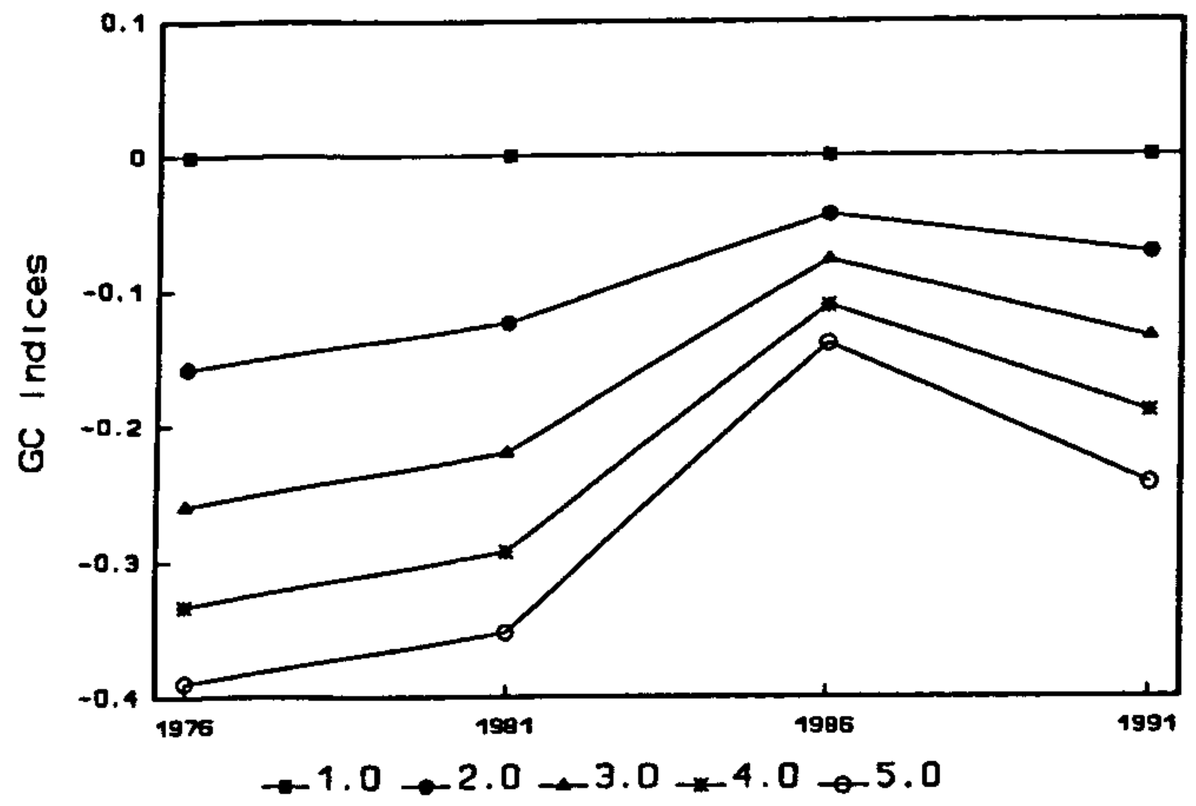
[10] The indices were computed according to equation (4.20). Note that μ_m and m_k represent average and district mortality *rates* rather than number of deaths. The w_k and E_k elements represent the same denominators used for calculating the standard concentration indices (*ie.* live births and still births).

Table 6.3: Generalized concentration indices
Infant mortality and components. 1971-1991

	δ	1971	1976	1981	1986	1991
Infant	1.0	-0.000	-0.000	-0.000	-0.000	-0.000
	2.0	-0.117	-0.080	-0.091	-0.055	-0.030
	3.0	-0.166	-0.130	-0.149	-0.080	-0.053
	4.0	-0.186	-0.167	-0.191	-0.010	-0.078
	5.0	-0.195	-0.193	-0.224	-0.112	-0.107
Neonatal	1.0	na.	0.000	-0.000	-0.000	-0.000
	2.0	na.	-0.028	-0.075	-0.061	-0.007
	3.0	na.	-0.047	-0.115	-0.082	-0.011
	4.0	na.	-0.057	-0.140	-0.090	-0.023
	5.0	na.	-0.063	-0.161	-0.099	-0.038
Post-neonatal	1.0	na.	-0.000	-0.000	0.000	-0.000
	2.0	na.	-0.158	-0.124	-0.043	-0.072
	3.0	na.	-0.259	-0.220	-0.077	-0.134
	4.0	na.	-0.334	-0.293	-0.111	-0.189
	5.0	na.	-0.390	-0.353	-0.139	-0.242
Perinatal	1.0	na.	-0.000	-0.000	-0.000	-0.000
	2.0	na.	-0.046	-0.071	-0.041	-0.001
	3.0	na.	-0.071	-0.108	-0.059	-0.004
	4.0	na.	-0.088	-0.133	-0.072	-0.010
	5.0	na.	-0.102	-0.150	-0.084	-0.017







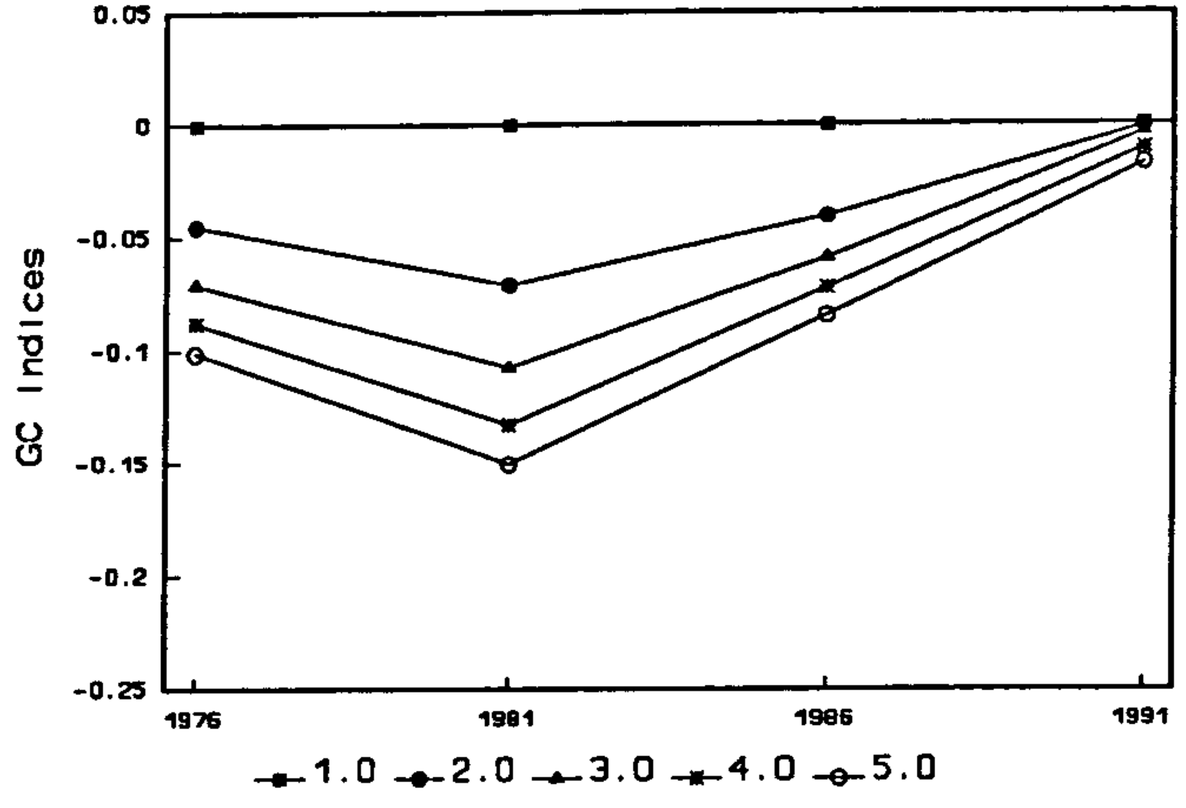


Figure 6.3: Generalized concentration indices. Infant mortality and components. 1971-1991. Values of δ = 1.0, 2.0, 3.0, 4.0 and 5.0

sensitive to changes in the value of δ. Compare, for example, the sequence of generalized indices for post-neonatal and perinatal mortality in 1986; or for that matter, the sequence of post-neonatal indices in 1991 with the neonatal values in 1981. In both comparisons, the standard concentration coefficients hardly differ, but the effect of increasing inequality aversion is considerably greater for the post-neonatal measurements. This suggests that the poorest districts bear a particularly high burden of post-neonatal deaths, and tends to confirm the view that chances of survival in the 1-12 month age group are highly susceptible to the degree of poverty.

The basic direction of the time-trend for the component indices does not change as greater weight is attached to the tail of the income distribution. Inequities in neonatal and perinatal mortality first increase and then decline. Similarly, post-neonatal inequity shows the reversed tendency identified earlier. Nevertheless, the degree of change in measured inequity is much less at higher levels of inequality aversion. For example, between 1976 and 1991, the rate of annual improvement in the $\delta = 5$ infant mortality indices is 2.8%, which compares unfavourably with the 3.9% change identified in Table 6.2. For the other indicators the differences between annual rates of change at $\delta = 2$ and $\delta = 5$ are: 4.7% - 2.4% (neonatal); 3.4% - 2.4% (post-neonatal), and 6.1% - 5.2% (perinatal). A greater concern for the tail of the income distribution, has a conspicuous impact on the course of inequity in total infant deaths. At $\delta = 5$, the 1981 distribution of infant mortality is actually more inequitable than that of ten years earlier. Furthermore, the post-1986 improvement, identified by the standard measurements, now appears insignificant. Finally, note that at $\delta \geq 4$, the inequity ranking of the components in 1986 reverts to that found for other years at all levels of inequality aversion (*ie.* post-neonatal > infant > neonatal >< perinatal). This suggests that the reversal identified by the standard measures for the late 80's is not robust to different distributional perspectives. All these results are indicative of the very poorest districts supporting a relatively high proportion of infant deaths. They also suggest that the generalized concentration coefficient is particularly sensitive to inequities that are the result of poor socio-economic conditions. If there is a policy concern for the very poorest in society, and not merely for inequality as such, then the generalized concentration indices are clearly a useful measurement tool.

6.4 Discussion

The problem of inequity in infant deaths has, for some time, captured the attention of researchers and policy-makers in Portugal. Recently, concern has focused on whether socio-economic and regional disparities, first identified in the 1970's, have shown signs of abating. Given that time series data are available, this question lends itself well to the concentration index approach. There is a problem, nevertheless, with regard to the choice of variable(s) used to partition the population. Our interest is in mortality inequalities that are related to economic status. However, information recorded on birth and death certificates that might be used to proxy this variable is generally of poor quality and, in any case, has only been published since 1987. Consequently, the present analysis has drawn on geographical observations of mortality and births ranked by an index of economic position.

Given that economic status is measured in an indirect manner, the validity and reliability of the inequity estimates may well be questioned. Such reservations cannot be fully resolved while a suitable alternative is unavailable. Nevertheless, there is no doubt that the data approach *does* address the question of interest. The economic ranking of the districts provides an ordering of the population in terms of average income levels. Admittedly, there exists a wide variance in commodity choice opportunities within the districts, but this is also true of occupational classifications (*eg.* 'skilled manual' may include long-term unemployed bricklayers and high earning electricians). The important point about the data approach is that a high level of aggregation is involved. It is possible that trends in the aggregate-level relationship between the socio-economic status and mortality of populations in the districts is quite different from the individual-level association between the same variables. However, we have no way of knowing since longitudinal individual-level data are not available. In these circumstances, the procedure of using geographical observations ranked by income levels is a valid method for measuring inequities in the health sphere. For countries with an insufficiently developed information base, it may be the only means for shedding light on the important question of time trends.

Doubts may also be raised about the legitimacy of using the construction industry

Table 6.4: The effect of using an alternative measure of living standards in the districts. Infant mortality

Year	Living standards ranking variable	
	Census-based	Wage index
1971	-0.089	-0.117
1981	-0.082	-0.091
1991	-0.021	-0.030

wage index as a measure of living standards in the districts. As I have stressed the choice is due to unavailability of more refined proxies in all years under consideration. Table 6.4 reports the results of using an alternative proxy. The new measure is based on population census information collected in 1971, 1981 and 1991. The ranking was established by a simple weighted average of the district orderings for three indicators: the percentage of population living in urban areas; percentage of dwellings with access to electricity, running water and refuse collection; and the adult literacy rate. Though the new measure produces lower inequality estimates, the general trends shown by the main results are clearly visible. Judging by the concordance tests reported earlier, other proxies would also show similar results.

The results obtained in section 6.3 point to three main conclusions. First, the degree of pro-rich inequity is currently much lower than it was in the 1970's. Irrespective of the indicator under consideration or the level of inequality aversion implied by the measures, the 25% reduction prescribed by the WHO has already been surpassed. Secondly, there are important differences in the trends of the four mortality components. Close-to-birth mortality inequities seem to have been largely erradicated. In contrast, inequity in the distribution of deaths to children aged 1-12 months is once again on the increase, after initially showing the greatest fall. The actual rates of these deaths are, however, much smaller than those in the earlier period of life. Finally, although the overall degree of inequity is much reduced, there are signs that the position

of the very poorest has not shown much improvement. This conclusion is suggested by both the configuration of the concentration curves and the trend of the inequality averse measures.

There is an apparent disagreement between these results and those of previous research. The studies by Kannisto (1986), Pereira *et al* (1987) and Leitão (1988) all suggest that the degree of inequity in infant deaths has remained more or less stagnant. The reason for this discrepancy is plain to see, but it needs spelling out not least for the important policy implications it holds. Previous studies have simply measured geographic mortality differentials and not whether these are associated with inequality in command over resources. They, therefore, address a quite different question to the one examined in this chapter. Nevertheless, geographic inequality results are often assumed to be a straightforward reflection of economic differences (*eg.* Santos and Hespanha, 1987), in much the same way as the 'north-south divide' in British mortality differentials is accepted as clear-cut evidence for inequities related to living standards (*eg.* Whitehead, 1987). The actual time-trend of the two types of inequity may, however, be quite different.

Figure 6.4 compares the evolution of geographic and income-related inequities in infant mortality between 1971 and 1991. The plots in the negative region reproduce the data presented in Table 6.2 and Figure 6.1 (*ie.* the observed and smoothed trends of the standard concentration indices). The plots in the positive area represent analogous trends for the Gini coefficient, which is used to summarize geographic inequalities.[11] Quite clearly, spatial inequalities have not diminished. Throughout the 1970's and 80's the trend of the Gini measures hardly diverges from a value of around 0.13. Furthermore, the Gini is insensitive to changes in income related inequality. In six of the years under analysis (75, 80, 81, 82, 87 and 88) the change in the coefficient (*vis-à-vis* the preceding year) is actually sign-antagonic to the change in the concentration index. The obvious implication of these findings is that geographical inequalities in infant deaths are still pervasive but that they no longer reflect, so decisively, known differences in standards of living. If there is a policy concern for area differences, then

---

[11] The Gini was calculated by ranking the districts in ascending order of *mortality rate*.

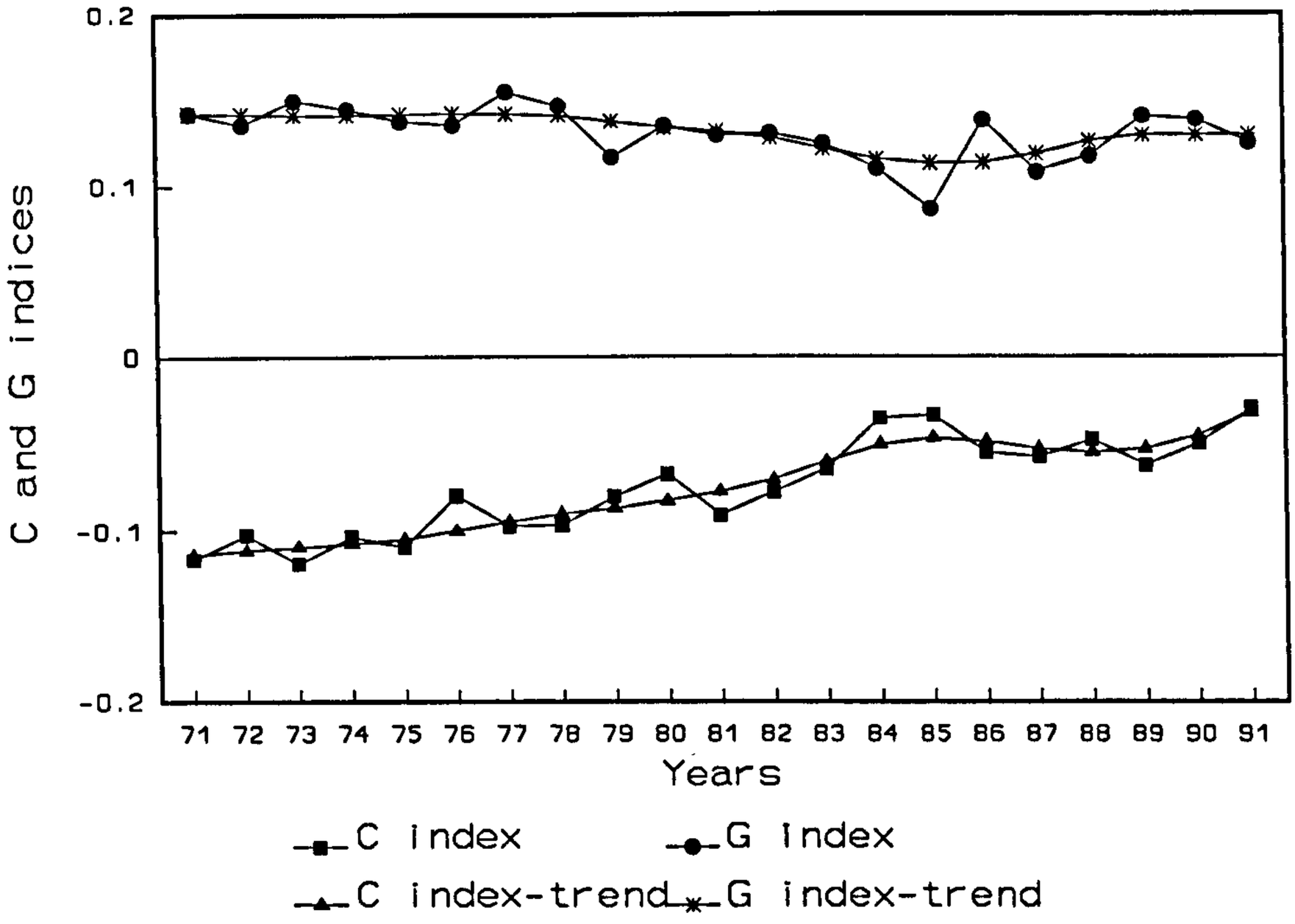


Figure 6.4: Comparison of concentration and Gini indices for infant mortality, 1971-91

it may be useful to target other factors which are known to affect the survival chances of infants (*eg.* very young mothers, mothers of high parity, the distribution of medical technology and so on).

It is important to stress, however, that the overall decline in income related inequity should not be interpreted as a sign that the issue no longer merits attention. The recent downturn in post-neonatal inequity and the position of the very poorest are two specific questions which the results suggest should be monitored carefully. Although many of the social and economic circumstances which lead to infant deaths may no longer apply in Portugal, some deaths could probably be avoided. This may best be achieved by targeting social support and health care resources at persons in poverty.

Finally, although the analysis has been useful in identifying past trends in infant health inequity, there is obviously a need for more refined measurement in the future.

Essentially, this means access to better data. For example, if small-area observations of both mortality and economic variables were available, then the reliability of inequity estimates would be improved. Such data have indeed become available in the late 1980's and could therefore be used in future analyses. It is also important that the results are compared with those obtained from individual-level data. This will require concerted efforts on the part of researchers and data collecting agencies to guarantee acceptable standards of data quality. If possible, the analysis of individual level data should be based on linked records, in order to avoid the problem of numerator/ denominator bias. Ultimately, however, the results of this chapter point to a new situation which requires alternative measures of infant health. The number of deaths ocurring each year are now quite low (at the levels of other EU countries) and the degree of inequity in their distribution is much reduced. If we still wish to monitor the relationship between economic disadvantage and infant health, then more sensitive but equally objective health indicators are required. These might include birth weight or other measures based on clinical examination, many of which could be easily produced as part of routine assessments.

Chapter 7

Measuring Advantage: Inequity in Health Care Payments, 1980-1990

"In general, the poorest members of the population ... have to be supported entirely by the State, and richer individuals - for reasons of social justice, equity and social solidarity - have to make a financial effort towards the maintenance of [health] services. ... That, for me, is the great principle of health care financing".
Paulo Mendo, Minister of Health (1993, pp. 10-11).

7.1 Introduction

This chapter provides estimates of the degree of vertical inequity in health care financing in Portugal, both at the beginning and end of the 1980's. On the basis of the earlier normative discussion, a vertically equitable distribution is interpreted as one where families' health care payments are positively related to ability to pay. The analysis - which basically conforms to the approach adopted in the ECuity study (van Doorslaer *et al*, 1993) - makes use of the global progressivity indices and associated graphical representations discussed in Chapter 4. Values of the Kakwani and Suits indices are computed for different sources of finance; namely, direct and indirect taxation, social insurance contributions, insurance premiums and out-of-pocket payments. The data are drawn from two household budget surveys carried out by the National Statistical Institute in 1980/81 and 1989/90. The main result is that, throughout the 1980's, the distribution of health care payments evolved from being overall progressive to overall regressive. This change is shown to be robust in terms of a wide range of methodological choices.

The main focus of the analysis is on providing relevant empirical information to the ongoing debate on health care finance in Portugal. When the NHS was created in 1979, one of the main arguments put forward in its favour was that tax-financing would

lead to greater equity in the burden of payments borne by different income groups (Assembleia da República, 1979). However, the following decade witnessed a marked escalation in the share of out-of-pocket financing, as the supply of NHS care proved insufficient to meet the growth in demand (see Chapter 2). There were changes too in the distribution of particular types of financing. The tax system, for example, underwent profound alterations with the introduction of VAT and the creation of a unified income tax. Patient co-payments within the NHS were also widened to various forms of care, after initially being circumscribed to pharmaceuticals. The redistributive impact of these changes has not been studied in any great detail. Only one empirical study has addressed the issue and that by means of simulation analysis of 1980/81 grouped data [Pereira and Pinto (1992, 1993)]. These data limitations mean that the evidence may easily be challenged. By using micro-data drawn from two periods, the present analysis is able to provide greater detail and accuracy in measurement, and ultimately, improved understanding of the effect of policy choices.

Recently, health care finance reform has been very much on the political agenda. The two main parties appear committed to enlarging the role of direct payments and private insurance [Mendo (1993); Campos (1990)]. However, this strategy is by no means consensual. Strong criticism has been raised in consumer, professional and academic circles, and legislation has continually had to be put back [Diário Notícias (1993); Ordem dos Médicos (1994)]. Amid the disagreement, however, there is some common ground. Virtually everyone agrees that health care financing should be guided by the ability to pay principle (see, for example, the quote at the head of the chapter); and all have limited knowledge of how the actual burden of payments has evolved over recent years! Therefore, the present chapter may help to clarify the current debate by providing empirical evidence on how, and in what degree, vertical equity in financing has been achieved.

There are two other aspects of the analysis that are worth underlining, each with implications beyond the Portuguese context. First, in common with previous chapters, there is a concern to measure the impact of what may be judged controversial methodological choices. I return, for example, to the effect of equivalence scale relativities and find a different impact to that on health concentrations indices. Other

issues that are discussed include the choice of the ability-to-pay proxy, standardization of family expenditures and aggregation of inequality in different parts of the distribution. I show that although there may be disagreement about particular methods, or even about attitudes to differential treatment of unequals, some relatively robust conclusions can be drawn. On the other hand, where alternative specifications are shown to have an impact on results, the analysis provides useful information to future empirical work.

Secondly, as far as I am aware, the present analysis is the first attempt to measure inequity in health care financing over time. This has more than mere curiosity value. The main purpose of distributional analysis is to monitor and inform policy. Previous studies of inequity in financing have either compared the progressivity of different sources at a single moment in time for a given country [*eg.* Long *et al* (1982)], or extended that analysis to take in various countries [*eg.* Gottschalk *et al* (1989); Van Doorslaer *et al* (1993)]. The latter are clearly more informative to the policy-maker since they allow the relationship between finance-mix and the degree of overall progressivity to be established. However, they also present serious problems of comparibility due to variation in the quality, scope and accuracy of the data in different countries (Schieber and Poullier, 1991; Maynard, 1993). In contrast, such data defficiencies as do exist in a particular country are unlikely to change their relative importance over a period of time such as the one considered here. Therefore, one-country time comparisons may provide a sounder basis for examining the effect of reforms on inequity in financing. Many of the changes which occurred in Portugal throughout the 1980's (*eg.* growth of patient co-payments, narrowing of income tax bands) are similar to those experimented in other countries. By measuring the degree of inequity at the end points of the period, the present analysis provides a valuable indication of the impact of these policy choices.

The chapter is organized as follows. Section 7.2 describes the data sources, the assumptions made in assigning the health finance burden and the definition of variables for empirical analysis. Section 7.3 presents and discusses the main empirical findings. Section 7.4 compares these results with those which might have been obtained under alternative assumptions. Finally, section 7.5 provides some concluding remarks drawing particular attention to policy implications of the findings.

7.2 Data and methods

7.2.1 The data sets

The analysis is based on microdata from the last two household budget surveys carried out by the National Statistical Institute: the 1980/81 Family Income and Expenditure Survey (FIES 80 for short) and the 1989/90 Family Budget Survey (FBS 90). The sample sizes are respectively 8039 and 9640 households, corresponding to 26753 and 29622 individuals.[1]

The surveys cover the non-institutionalized population of Portugal and yield representative estimates at the national and regional level. Their sampling and data collecting techniques are similar to those employed in budget surveys in other countries (*eg.* the UK FES). The basic sampling unit is the household, defined as a set of persons at the same address with common food and housing expenditures. Sampling is based on a multi-stage probability design, using information from the most recent decennial census. Interviewing takes place through a period of twelve months, with both surveys having run from March to February of the following year. The interviews were evenly distributed throughout the periods and the households were observed only once. A wide range of detailed information was collected, including income levels and sources, taxes and other contributions, public and private transfers, expenditures on goods and services and socio-economic characteristics of household members. Much of the expenditure data is derived from detailed diaries which respondents are required to keep during one week.

The FIES 80 and the FBS 90 are the only data sources available in Portugal that permit the overall health financing burden to be measured. Naturally, they are not free of drawbacks. It might be argued, for example, that the incidence of health care financing should be measured over a life-cycle rather than at a single moment in time. This is particularly relevant if, as in Portugal, a large portion of health service revenues

---

[1] In order to maintain consistency with the empirical analysis of previous chapters, observations for residents of the Madeira and Azores islands (roughly 5% of the country's population) were excluded from the analysis.

are raised through taxation. For instance, sales taxes may appear regressive at any point in time because the poor tend to have a higher average propensity to consume than the rich. However, over a lifetime average propensities to consume tend to unity, so the taxes would look close to proportional on a lifetime incidence basis (Davies *et al*, 1984; Poterba, 1989). Many such problems have been raised in the literature, but no convincing solutions have been proposed based on readily available data (Atkinson, 1990). Unless one has access to longitudinal data, analysis of the health financing burden must be restricted to providing "snapshots" of current distributions.

The other main drawback concerns data reliability. There are three potential sources of bias: recording errors, differential non-response and an atypical year of comparison. Only in the latter case does the available information permit any reasonable form of correction. The FBS 90 coincides with the introduction of a new income tax system which led to a temporary distortion of the underlying distribution of taxes. As a general rule, from 1989 onwards taxes on income were retained at source, whereas before an interval of one year elapsed between income accrual and payments. This meant that in 1989 many households paid income taxes under the 'old' and 'new' systems; the main exceptions being the large proportion of poorer households who make no such payments (*eg.* the unemployed, persons on state pensions and those whose incomes did not reach the payments threshold). Therefore, payments reported in the FBS 90 are likely to overstate the true degree of progressivity of the income tax system. In order to correct this deficiency I have excluded payments of taxes abolished under the new system. This option is not without its problems (see further on), but is likely to provide a more accurate picture of the underlying distribution than if all payments reported in the survey were included.

With respect to the other problems, there is insufficient documented evidence to allow accurate corrections to be made. Experience with similar surveys in other countries has shown that recording errors may include the reporting of expenditures incurred outside the reference period and the under-reporting of certain types of income [*cf.*, *eg.* Kemsley *et al* (1980), Atkinson and Micklewright (1983), Smeeding and Schmaus (1990)]. To a significant degree, the latter problem may arise because of differential non-response among sampled households. British evidence, for example, suggests that "older households, households where the head is self-employed, those

without children and higher income households, are less likely to cooperate than others" [Atkinson and Micklewright (1992, p. 62)]. Unfortunately, information on the magnitude of these biases in the portuguese surveys is rather limited. The only evidence comes from Pereirinha (1988) who compared the FIES 80 income data to National Accounts and found significant shortfalls with respect to self-employment and investment income. However, there has been no substantiation of data on the payments side and the level of differential non-response has not been documented. In these circumstances, any adjustments to the data would be arbitrary.

7.2.2 Assigning the financing burden

Estimates of progressivity are made for the four sources of finance identified in Chapter 2, as well as for the health care system overall. The sources are:

(i) *general tax revenues*, which are used to fund the NHS and to subsidize occupational insurance schemes operating in the public sector;

(ii) *social insurance contributions* to occupational schemes, of which by far the largest is the ADSE scheme for public servants and their families;

(iii) *private insurance premiums*; and

(iv) *direct expenditures*, including NHS co-payments and payments to the private sector.

The distribution of these payments according to 'ability to pay' is derived directly from the surveys. However, given that tax revenues are not earmarked for health care, a method of weighting the sources when calculating the overall burden of payments is required. There are basically two options. One, where non-earmarked payments are allocated *pro rata* to the shares of relevant revenues going to finance health care, and the resulting survey proportions are used as weights; and another, where all sources are weighted by shares derived from aggregate data. Both these procedures are somewhat arbitrary. However, the first has the added drawback of leading to biased estimates if overall levels of different types of contribution are recorded with different degrees of accuracy and completeness.[2] Because of this problem, previous studies (including the

---

[2] There is evidence that this may be so from a number of sources (*eg.* Kemsley *et al*, 1980; Borooah *et al*, 1991). The discrepancy arises, in part, because budget surveys are primarily designed to measure expenditures and hence the level of detail on this component is much greater than, say, for

ECuity research) have opted to aggregate the source distributions according to macro shares. The same procedure is followed here.[3]

Irrespective of whether a particular source is earmarked for health care, the question arises as to who bears the economic burden of the payment. The theoretically correct approach to this question is quite clear (Atkinson and Stiglitz, 1980; Aaron, 1992). One wants to compare the general equilibrium (GE) of the economy before the payment is made to the equilibrium which occurs afterwards. Ideally the GE model would be sufficiently disaggregated and dynamic, and would allow for market and other distortions. This would permit the true shifting of taxes to be identified, deadweight losses to be accounted for and consumer surplus associated with expenditures to be excluded from calculations. In practice, of course, this method is overly aspiring given data limitations and the relative underdevelopment of computable GE models. Consequently, one has to opt for the less ambitious approach of specifying certain plausible incidence assumptions within a partial equilibrium framework.

The approach used here and in the ECuity research is basically an application of the tax incidence evaluations carried out by Pechman, Musgrave and others [*eg.* Pechman and Okner (1974), Musgrave *et al* (1974), Reynolds and Smolensky (1977) and Pechman (1985)]. The advantages of this approach are its transparency, in the sense that assumptions are made explicit, and applicability, which means that detailed evidence on a matter of considerable interest to policy makers can actually be provided. In assigning the financing burden, I have adopted a standard set of incidence assumptions.[4] Income, property and capital taxes are assumed to be borne fully by tax-payers. This conjecture implies that factor supply is either fixed or fairly inelastic. The corporate income tax is assumed to be divided equally between capital income recipients and consumers. Given the controversy which surrounds the incidence of this tax and the fact

---

social insurance contributions.

[3] The macro weights are derived from Ministério das Finanças (1981, 1989), OECD (1993) and Pereira *et al* (1994). They are shown in Table 7.1 along with the main results.

[4] See Pechman (1985, Chapter 3) for a detailed explanation of the reasoning behind those related to taxation.

that household budget surveys may provide an incomplete picture of its distribution, two other alternatives are considered later in the chapter: (i) that the incidence falls completely on dividend recipients and (ii) that the burden is passed on to consumers in the form of higher prices. Indirect taxes on both final and intermediate goods and services are assumed to be fully shifted to consumers. The incidence is therefore allocated according to the consumption propensities of households. Payments to occupational insurance schemes are assumed to be borne by the employees' households. Note that employer contributions are largely irrelevant to the portuguese case given that general social insurance is not used to finance health care expenditures. The "social insurance" component measured in the analysis refers to mandatory contributions made overwhelmingly by public sector workers. It is assumed that any eventual deficits in the relevant schemes are borne by tax-payers. Finally, earmarked payments - private insurance premiums and net direct expenditures - are assumed to fall entirely on the households who make the payments.

7.2.3 Variable definitions

A household's *ability to pay* is measured by its gross income adjusted by the number of equivalent adults. This variable serves as the benchmark for assessing the progressivity of payments. The gross income definition is fairly comprehensive. It includes wage and self-employment income, cash property income, public and private cash transfers, the value of home-produced consumption, in-kind earnings and imputed rent on owner occupied housing. It does not include capital gains or employer contributions to social and private insurance schemes. Household incomes were converted to a per-equivalent-adult basis in order to provide a measure of the standard of living available to the household. The equivalence scale used for this purpose is that of the OECD, also used in Chapter 5.

Personal *direct taxes* (*eg.* income tax, property taxes, inheritance tax) have been allocated on the basis of actual payments reported by households. The surveys do not, however, inquire about corporate taxes, which make up around 30 per cent of direct tax revenues. These have been allocated half in proportion to capital income and half in proportion to household expenditure. The total direct tax variable is weighted in

accordance with the revenues raised from non-corporate and corporate taxation. For the reasons stated earlier, the 1989/90 analysis excludes taxes abolished under the new income tax system. This means that whereas in 1980/81 the full burden is imputed to households, in the later period around 20% of direct tax revenues are left unallocated.

The allocation of *indirect tax* financing is based on work carried out for the government commission which supervised the introduction of VAT in 1986 (Domingues *et al*, 1984).[5] This is the most recent study of indirect tax incidence in Portugal. The authors estimated the tax burden both before and after the introduction of VAT. The before-data, which I have used for the FIES 80 analysis, refer to estimates for 1979 based on the 1973/74 FIES. The after-data, used for the FBS 90 analysis, are a simulation of the same information admitting VAT rates very similar to those actually in place during 1989/90. Domingues *et al* calculated effective tax rates by eight total expenditure percentile groups for seven classes of goods and services. In order to generate the indirect tax distributions, I have applied these rates to expenditures reported in the two surveys by matching by the relevant categories.[6] The Domingues *et al* estimates are based on taxes which comprised roughly 60 per cent of all indirect tax revenues in 1980 and 1989. Stamp duty and excise taxes on tobacco, petrol and other goods were not considered. Given that in the present analysis the overall health care financing burden is weighted by the full share of indirect taxation, I make the implicit assumption that ommited taxes are distributed as those that are included.

The variable termed *social insurance* represents mandatory contributions to occupational schemes. Although the health care financing component is usually autonomous in these schemes, no such separation is available in the data sources. This is unimportant for the present analysis given that both health and non-health contributions are typically proportional to earnings with no ceiling being applied. I have assumed that the distribution of health related payments reflects that of social insurance contributions made by civil servants. This seems valid given that the civil servants'

---

[5] I was unable to calculate the incidence of indirect taxes directly from the surveys because the available data files provided an insufficient level of expenditure disaggregation.

[6] The information from Domingues *et al* (1984) is taken from Tables I-A1 and VI-A1. I have used Scenario E for the FBS 90 analysis.

health fund (ADSE) accounts for around 75% of all health related social insurance financing. Thus, the only major source of error in the calculations is if the distribution of earnings among other insured households is radically different to that of civil servants. This is unlikely to be the case.[7]

The distributions of *private insurance premiums* and *direct payments* are derived directly from the actual values reported by households. In the second case, the data provide a considerable amount of detail with regard to the type of care consumed (*eg.* pharmaceuticals, doctor visits), but not with respect to the sectoral mode of consumption (*eg.* NHS or private). In both years, direct payments have been computed net of reimbursements.[8] However, because of data limitations, no account is taken of tax rebates that households might receive in respect of their health care expenditures.

Two further definitional issues of some importance concern the weighting of units and allowance for differences in household structure in the payments variables. In both cases, I adopt the assumptions made in van Doorslaer *et al* (1993). That is, each household is given equal weight irrespective of the number of individual members; and health care payments are not adjusted for household size and composition. The rationale for these choices is that the rules governing health care payments typically relate to families or households rather than individuals, and that economies of scale are unlikely to apply in health care consumption (Wagstaff *et al*, 1992). There are, however, some grounds for disagreement on these issues, and hence they are among those for which I modify the assumptions in section 7.4.

---

[7] The other main insurance schemes are for bank and insurance personnel, the armed forces and public utility workers. In the FBS 90, the Gini coefficient for the comprehensive income measure used in the analysis is, 0.19 for households where the head works in one of these sectors, and 0.21 where the head is a civil servant. Note that the definition of civil service in Portugal is much broader than in the UK. For instance, all NHS, local government and public sector teaching personnel are classed as civil servants and therefore contribute to the ADSE fund.

[8] The FIES 80 provides specific information for this purpose. However, in the FBS 90 health sector reimbursements are grouped with other transfers. I have estimated the 1989/90 values by matching the data files according to gross household income vingtiles and assuming that equal proportions of expenditures were reimbursed for each matched income group in both periods. A restriction of non-negative expenditures was also imposed.

7.3 Empirical Findings

The main results are shown in Tables 7.1 and 7.2. Table 7.1 compares the proportion of source-disaggregated health care payments borne by each decile of equivalent income in the two periods under analysis.[9] Table 7.2 reports standard progressivity indices estimated for the source distributions and for the health care system as a whole. These, and all other index values shown subsequently, have been computed from microdata using the covariance methods discussed in Chapter 4.

7.3.1 Taxes

Direct taxes are globally progressive in both periods, but over time the extent of their progressivity declines substantially. The larger fall computed for the Suits index suggests that the change is mainly attributable to reduced levels of progression in the top end of the income distribution.[10] It is noticeable too that the 1980/81 and 1989/90 concentration curves do not cross, implying that there was an unambiguous shift in the direct tax burden favouring households that are better off. Although other factors may have contributed to this change, it is likely that the main explanation lies in the reform of the income tax system brought about in 1989. Among the alterations then enacted were a reduction in the number of personal income tax brackets from eleven to five, with the difference between the minimum and maximum rates becoming 24 percentage points, against 76 in the previous system. The reform also set a uniform rate for capital income and widened the tax base considerably.[11]

The distribution of indirect taxes also becomes more favourable to the rich, with

---

[9] A graphical illustration, by means of relative concentration curves, is also provided in Appendix 5 (Figure A5.1). These curves have been drawn by taking the 1989/90 pre-tax income shares as the reference distribution. As Table 7.1 shows there was virtually no change in the distribution of pre-tax income between 1980/81 and 1989/90. This means that the curve comparisons in Figure A5.1 provide a useful graphical representation of distributional changes in the payments distributions from one period to the next.

[10] This may easily be verified by computing measures of local progression from the data shown in Table 8.1 [see Lambert (1989, pp. 159-163)].

[11] See OECD (1991) for a detailed survey of Portuguese fiscal reform in the 1980's.

Table 7.1: Distribution of health care financing
1980/81 - 1989/90

Income deciles	Pre-tax income	Direct taxes	Indirect taxes	Total taxes	Social insur.	Private insur.	Direct paym.	Total payments
1980/81								
Poorest	2.8%	0.5%	2.9%	2.0%	0.2%	0.0%	7.2%	3.4%
2nd	4.3%	1.2%	4.0%	3.0%	0.9%	2.6%	7.4%	4.1%
3rd	5.5%	2.3%	4.7%	3.9%	1.9%	0.6%	7.0%	4.6%
4th	6.5%	3.6%	6.8%	5.7%	2.7%	3.1%	8.9%	6.4%
5th	7.6%	5.7%	7.6%	7.0%	3.9%	4.1%	9.4%	7.5%
6th	8.9%	6.5%	8.0%	7.5%	7.7%	4.9%	9.3%	8.0%
7th	10.3%	8.2%	9.9%	9.3%	8.8%	22.7%	10.3%	9.6%
8th	12.2%	11.3%	11.3%	11.3%	12.0%	10.5%	11.9%	11.5%
9th	15.5%	18.8%	16.6%	17.3%	20.7%	16.8%	12.4%	16.1%
Richest	26.4%	42.0%	28.2%	33.0%	41.2%	34.7%	16.1%	28.7%
Percent from each source		23.2%	42.8%	66.0%	5.2%	0.6%	28.2%	100.0%
1989/90								
Poorest	3.0%	0.7%	2.5%	1.9%	0.7%	0.4%	6.5%	3.5%
2nd	4.3%	2.0%	4.1%	3.4%	3.2%	0.5%	7.9%	5.0%
3rd	5.3%	2.9%	5.0%	4.2%	3.4%	5.2%	6.9%	5.2%
4th	6.3%	5.0%	6.7%	6.1%	4.5%	4.6%	8.2%	6.8%
5th	7.5%	6.2%	7.6%	7.1%	4.3%	9.2%	8.6%	7.5%
6th	8.7%	9.0%	9.1%	9.1%	4.0%	7.2%	10.5%	9.3%
7th	10.1%	10.6%	10.5%	10.6%	6.5%	10.6%	10.9%	10.4%
8th	12.1%	12.9%	12.5%	12.7%	7.9%	5.8%	9.0%	10.9%
9th	15.6%	17.7%	16.1%	16.7%	15.5%	16.7%	12.7%	15.1%
Richest	27.2%	33.0%	25.9%	28.5%	50.1%	39.8%	18.6%	26.3%
Percent from each source		20.7%	34.5%	55.2%	6.0%	1.4%	37.4%	100.0%

Table 7.2: Health care financing progressivity indices

1980/81 - 1989/90

	Concentration		Kakwani		Suits	
	1980	1990	1980	1990	1980	1990
Equivalent income	0.343	0.351				
Direct taxes	0.570	0.479	0.227	0.127	0.253	0.123
Indirect taxes	0.362	0.350	0.019	-0.002	0.021	-0.011
Total taxes	0.436	0.398	0.092	0.047	0.103	0.040
Social insurance	0.588	0.595	0.245	0.244	0.260	0.299
Private insurance	0.519	0.503	0.175	0.152	0.177	0.173
Direct payments	0.147	0.166	-0.196	-0.186	-0.201	-0.188
Total payments	0.363	0.323	0.019	-0.027	0.026	-0.028

the Kakwani and Suits measures suggesting a change from global progressivity to regressivity. However, the change is quantitatively small so that in both periods the distributions hardly diverge from the proportional benchmark. This suggests that differential sales tax rating - a feature of the portuguese tax system both before and after the introduction of VAT in 1986 - may be an effective instrument for achieving vertical equity.

The overall progressivity of the tax system is computed by weighting the direct and indirect tax distributions by the respective State revenue shares. Naturally, the results point to a reduction in progressivity over time. The Kakwani index declines to 51 per cent of its 1980/81 value and the Suits index to 39 per cent. This shift in the taxation burden appears to have penalized middle income groups the most. Table 7.1 indicates that local progression for the lowest decile actually increases, but that it declines considerably at the highest levels of income. Thus, whereas in 1980/81 only the richest quintile paid a higher share of taxes than their share in total income, in the

later period this condition is shown by the top half of the distribution. It is significant, however, that despite the large fall in progressivity, richer households continue to support a higher than proportional share of the tax burden.

It is worth considering the extent to which the data problems referred earlier may undermine the results. Because of the overlap between the FBS 90 and the income tax reform, it is possible that the fall in direct tax progressivity is slightly overstated. To see this, recall that pre-reform income taxes paid during 1989 were excluded from the analysis. There seems little doubt that this is justified, given that the 'double payment' transitional period led to a temporary distortion of the income tax distribution. However, reported payments of the 'new' taxes may also provide an imperfect basis for measuring the underlying progressivity of post-reform income taxation. For example, although under the new system income tax is generally retained at source, there are situations when this is not feasible (*eg.* self-employment income). In such cases, taxpayers are required to make pre-payments at the lowest tax rate based on self-assessed liabilities. Obviously, once the yearly assessments are made the effective tax rate for different units will vary considerably. It is unlikely that the FBS 90 captures this correction, given that the period of observation only lasted until March 1990. If effective tax rates for the self-employed were, on average, higher than the lowest marginal rate, then the 1989/90 estimates will understate the true degree of progressivity. Unfortunately, the data do not allow for such bias to be accurately measured, but a simulation admitting that income tax payments by the self-employed are double those actually reported led to the Kakwani index for direct taxes in 1989/90 rising to 0.137. Since this value is not much greater than the reported estimate and because other potential biases may in fact operate in the opposite direction (*eg.* tax rebates that are proportionally greater for higher income households), there seems little reason to believe that a large fall in progressivity did not take place.[12]

---

[12] Subsequent to writing this chapter two papers have been brought to my attention which may raise doubts about the estimates provided for direct tax progressivity. Rodrigues (1993) and Gouveia and Tavares (1995) both measured the distribution of disposable income in Portugal using the FIES 80 and FBS 90. They concluded that inequality had decreased in the period, which seems incompatible with the reduced levels of progressivity identified in this chapter. However, the reason for their results is precisely the failure to consider the biasing effect of the 1989 tax reform. The two papers subtracted taxes levied under the "old" and "new" systems when computing individual disposable income levels.

The indirect tax results can also be questioned on the grounds that they depend on the VAT Commission estimates. However, the error induced by the allocation procedure may not be so great. There are two potential biases: one, if expenditure patterns changed considerably since the mid-70's (*ie.* the period from which Domingues *et al* (1984) drew their data to calculate effective tax rates); and another, if the distribution of excluded taxes is radically different from those that are included. As to the first issue, the analysis provides some form of correction to the extent that the actual expenditure distributions observed in the FIES 80 and FBS 90 are used. With regard to excluded taxes (*eg.* excise taxes), there is no recent evidence as to their incidence. However, Tanzi and de Wulf's (1976) estimates for 1973 suggest that they were marginally progressive. Nothing guarantees that this pattern remained stable in later periods, but given that included taxes account for roughly 60 per cent of indirect tax revenues, some very large changes would have had to take place for the present results to provide a grossly inaccurate indication of the overall burden of indirect taxes in the 1980's. Therefore, although a degree of caution is advisable in relation to the tax estimates, it would be surprising if the general trends suggested by the results were contradicted by better data.

7.3.2 Social and private insurance

Social and private insurance contributions are globally progressive in both periods, with their distributions being similar to those found for direct taxes. However, the explanation is somewhat different. The progressivity of direct taxation is largely attributable to increasing marginal rates of tax. In the case of social insurance, the progressive structure stems from the fact that the funds are not universal. Those employees that are covered make contributions in proportion to earnings, but because they tend to be in higher income groups (see Chapter 2), the source emerges as highly progressive. Similarly, the progressivity of private insurance is due to selective

---

They also subtracted contributions which are not included in the present analysis (*eg.* employee social insurance contributions). It is these factors which apparently explain the discrepancy in the results. For the reasons given above, I believe that the approach followed in this chapter is the more useful if one wishes to estimate the underlying distribution of direct taxes (or income net of those taxes) at the beginning of the 1990's.

coverage; in this case, because insurance is mainly purchased by richer households, as a supplement to NHS cover.

Levels of progressivity for insurance payments remained relatively stable throughout the 1980's. This is unsurprising given that there were no significant changes in the rules governing social insurance contributions, and that private insurance continued to be a relatively unimportant form of finance. The results by income decile do suggest some changes at different points of the income distribution but it is doubtful that they are of any great significance. For instance, the social insurance distributions show a fall in local progression at low to middle ranking incomes and a rise for the two richest deciles. In the absence of changes in the rules governing payments, these results are likely to reflect the evolving structure and relative value of public sector earnings.[13] There are very similar changes in the local progression of private insurance payments. However, it is not advisable to read too much into these results since the number of households in the decile cells is rather small (overall, the number of households reporting private insurance expenditures is 69 in the FIES 80 and 305 in the FBS 90).

It is important to note that the insurance results do *not* imply that greater reliance on these sources would induce greater progressivity in health care finance. A significant expansion of either source would likely involve increased coverage among lower and middle income groups, which would tend to reduce progressivity levels. The cross-national results presented in van Doorslaer *et al* (1993, pp. 40-42) suggest that in countries where social insurance payments are not restricted to specific professions, the source tends to be, at most, marginally progressive. Wherever social insurance is the main source of revenue to the health service, it emerges as regressive. The same study also indicates that it is only in countries (such as Portugal) where private insurance is taken out as a supplement to comprehensive cover provided by the State, that this form of payment is progressive. If private insurance is (or is nearly) the sole source of cover for a part of the population, then it tends to be regressive.

---

[13] Civil service pay differentials widened throughout the 1980's and average levels declined in relation to the private sector (Ministério das Finanças, 1988). These trends will have altered the relative ranking of public sector workers in the income distribution, and hence the progression results.

7.3.3 Direct payments

In contrast to other forms of finance, direct payments are shown to be highly regressive. In 1980/81, the poorest 30% of the population held 13% of income but paid 22% of all direct expenditures. The richest 30%, by comparison, held 54% of income and paid 40% of expenditures. The 1989/90 results present virtually the same distributional pattern (*cf.* the concentration curves in Figure A5.1), so that global regressivity levels remain stable at around -0.19. This suggests that the growth of private financing in the 1980's was secured by all income groups increasing their expenditures in roughly equal proportion.[14]

The interpretation of these results, and their implications for equitable financing of health services, are far less clear than for other sources. Since direct payments reflect consumption choices, the only way that the regressive structure may properly be understood is by focusing on the factors that affect demand for care by means of a behavioural model. Such an undertaking is beyond the scope of this thesis. Nevertheless, the present framework can yield a number of useful insights which may be followed up in later work.

The first point to note is that the results take no account of the fact that health care expenditures are tax-deductable. Pinto and Santos (1993) examined tax returns in 1989 and found that fiscal savings associated with health care expenditures were highly progressive, which in turn suggests that the distribution of direct payments in Portugal is even more regressive than the present results indicate. Because of data limitations, I have been unable to measure the precise extent of this effect. However, it may well be of some importance, particularly in the later period. In 1980/81 there were limits on the amounts which could be deducted (50% of expenditures at most) and certain expenditures were ineligible (*eg.* pharmaceutical expenditures). Following the 1989 income tax reform no such limits prevailed. In principle, therefore, the richest households might be able to recoup as much as 40 per cent (the highest marginal tax

---

[14] Recall from Chapter 2 that the increased weight of out-of-pocket financing shown in Table 7.1 corresponds to an increase in GDP share while the share of public financing remained stagnant.

rate) of their health care expenditures, whilst the poorest (if their taxable income does not reach the payments threshold) recoup nothing at all. Future studies should look closely at this issue.

A further element of interest is that by international standards the level of regressivity is rather high. For example, it is higher than in all countries in the ECuity study that rely on tax-financing as the main source of revenue (*eg.* Denmark, UK, Ireland).[15] Given that in Portugal the share of out-of-pocket payments is higher than in these countries the regressive structure neutralizes the equitable impact of revenues raised through the tax system. It is noticeable too that countries such as Spain or Italy - which have high shares of direct expenditures and health systems that operate in a manner similar to the portuguese - reveal expenditure distributions that are close to being proportional.

These observations raise the question of whether public policy may be contributing to high levels of expenditure regressivity. For example, could limited NHS supply in some medical specialities mean that less well-off patients have to resort to the private sector? To what extent does the practice of charging uniform rates of co-payment irrespective of the patient's economic condition lead to regressivity? The level of disaggregation available in the data allows only limited light to be cast on these issues. Table 7.3 presents a disaggregation of the Kakwani indices by type of care consumed. As for overall expenditures, these results show minimal variations from one period to the next, suggesting that policy and behavioural changes had a negligible impact on actual distributions. However, there are rather large differences in levels of progressivity/regressivity across types of care consumed. Direct expenditures on hospital care are highly progressive, suggesting that these outlays are largely borne by the rich. This is unsurprising given comprehensive cover by the NHS and absence of co-financing. Expenditures on diagnostic procedures are marginally progressive. In this case, NHS co-payments (introduced in 1987) are waived for certain patient groups such as the poor, unemployed, pregnant women and children. This may be a crucial factor

---

[15] See Figure A5.2 in the Appendix which compares the Kakwani index values computed for each financing source with the measurements made in the country reports in van Doorslaer *et al* (1993).

Table 7.3: Progressivity of direct payments by type of care consumed
1980/81 - 1989/90

	Percentage Share		Kakwani index	
	1980	1990	1980	1990
Pharmaceuticals	61.0%	54.0%	-0.309	-0.323
Therapeutic appliances	5.0%	12.0%	-0.030	-0.026
Medical services and diagnostic procedures	27.0%	29.0%	-0.049	-0.051
- Doctor consultations	na.	18.0%	na.	-0.077
- Dental care	na.	4.0%	na.	-0.042
- Nursing/paramedical	na.	1.0%	na.	-0.201
- Diagnostic procedures	na.	6.0%	na.	0.045
Hospitalization	6.0%	4.0%	0.138	0.154
Other	1.0%	1.0%	-0.101	-0.113
Total direct payments	100.0%	100.0%	-0.196	-0.186

- na. = not available
- Reimbursements are assumed to be divided between each type of care in proportion to their share in total direct payments.

determining the progressive structure in 1989/90. All other forms of care show regressive distributions. Of these, the most noteworthy results are for pharmaceutical expenditures, where values of $\pi_K < -0.30$ were computed in both periods. It would seem that these outlays are the main contributory factor to high levels of regressivity in the total expenditure distribution. Despite various changes in NHS co-financing arrangements, no exemptions on demographic, social or economic grounds have ever operated (see Chapter 2). Naturally, the results do not merely reflect uniform rates of co-payment within the NHS. Some expenditures are on OTC medicines and others are made by privately insured individuals. However, the market share of these categories is relatively small, fluctuating each year around 27 per cent (Infarmed, 1993). Data from the INS 87 also indicate that pharmaceutical outlays account for over three quarters of direct expenditures made by NHS patients (Pereira *et al*, 1994). Thus, it is likely that the high levels of drug expenditure regressivity shown by the two budget surveys reflect,

to an important extent, policy choices and that these have inequitable consequences. Generally, however, the results of this section call for further research to be undertaken.

7.3.4 Overall health care financing

The aggregate results show that, over the 1980's, health care financing in Portugal became unequivocally more favourable to the rich. In 1980/81, the financing system was marginally progressive, with both the Kakwani and Suits indices displaying values in the region of 0.02. By 1989/90, the estimates show a decline of roughly five points, suggesting that health care finance had become slightly regressive. An indication of the gainers and losers can be gathered from Table 7.1, by considering which decilescontributed to the health system in greater proportion than their share in total income. In 1980/81 this condition is met by the 1st, 9th and 10th deciles; whereas in the later period, all bar the 3rd, 8th, 9th and 10th are in such a situation. Therefore, the burden of health care finance would appear to have shifted to middle income groups, with the principal beneficiaries being households who are situated in the richest quintile.

Under the concentration index approach, changes in the progressivity characteristics of a health care financing system depend on: (i) variations in the degree of progressivity of each finance source (the *progressivity effect*) and, (ii) variations in the proportion of total revenues raised from each source (the *revenue effect*). The first of these can be captured by allowing that a particular source maintains its earlier index value whereas all others change to the later measurements. Such a simulation is shown in Table 7.4.[16] It is evident that changes in the tax distributions were the major contributory factor in terms of progressivity variations. The revenue effect is shown in the table as the difference between the sum of progressivity effects and the overall change in index values. The Kakwani measurements suggest that progressivity and revenue changes contributed in roughly equal proportion to the decline in health care financing progressivity. The picture is slightly altered for the Suits index on account of the greater weight it attaches to departures from proportionality at the top end of the

---

[16] I make use of the result that the overall Kakwani and Suits indices are weighted averages of the individual component indices (see Chapter 4).

Table 7.4: Impact of changes in source progressivity and revenue proportions on the overall burden of health care finance

	Kakwani index		Suits index	
	Simulated value[1]	Difference from observed[2]	Simulated value	Difference from observed
Source progressivity effects				
Direct taxes	-0.006	-0.021	-0.001	-0.027
Indirect taxes	-0.020	-0.007	-0.017	-0.011
Total taxes	0.001	-0.028	0.010	-0.038
Social insurance	-0.027	0.000	-0.030	0.002
Private insurance	-0.027	0.000	-0.028	0.000
Direct payments	-0.031	0.004	-0.033	0.005
Total progressivity effect[3]		-0.024		-0.031
Total revenue effect[4]		-0.022		-0.023
Overall change[5]		-0.046		-0.054

1. Admitting that in 1989/90 the respective source maintains its former index value but all others change.
2. Numerical difference of the simulated index value in relation to that actually observed in 1989/90.
3. Sum of source progressivity effects.
4. Difference between absolute change in index values and total progressivity effect.
5. Absolute change in index values between 1980/81 and 1989/90.

distribution (note in particular the change in the tax burden). From Table 7.1 it is clear that the principal change in the revenue structure is a fall in the proportion of tax-financing, matched by an almost equal rise in the share of out-of-pocket payments. Therefore, in the 1980's, the overall burden of health care financing became more favourable to the rich mainly as a result of reduced progression in the tax system and an increase in the revenues raised directly from consumers.

The change in the overall finance burden may be put into perspective by comparing the results with those of other countries. The obvious point of reference is the ECuity study (Van Doorslaer *et al*, 1993; Wagstaff *et al*, 1992), on whose methodology the present research is based. The ECuity results were arrived at using the

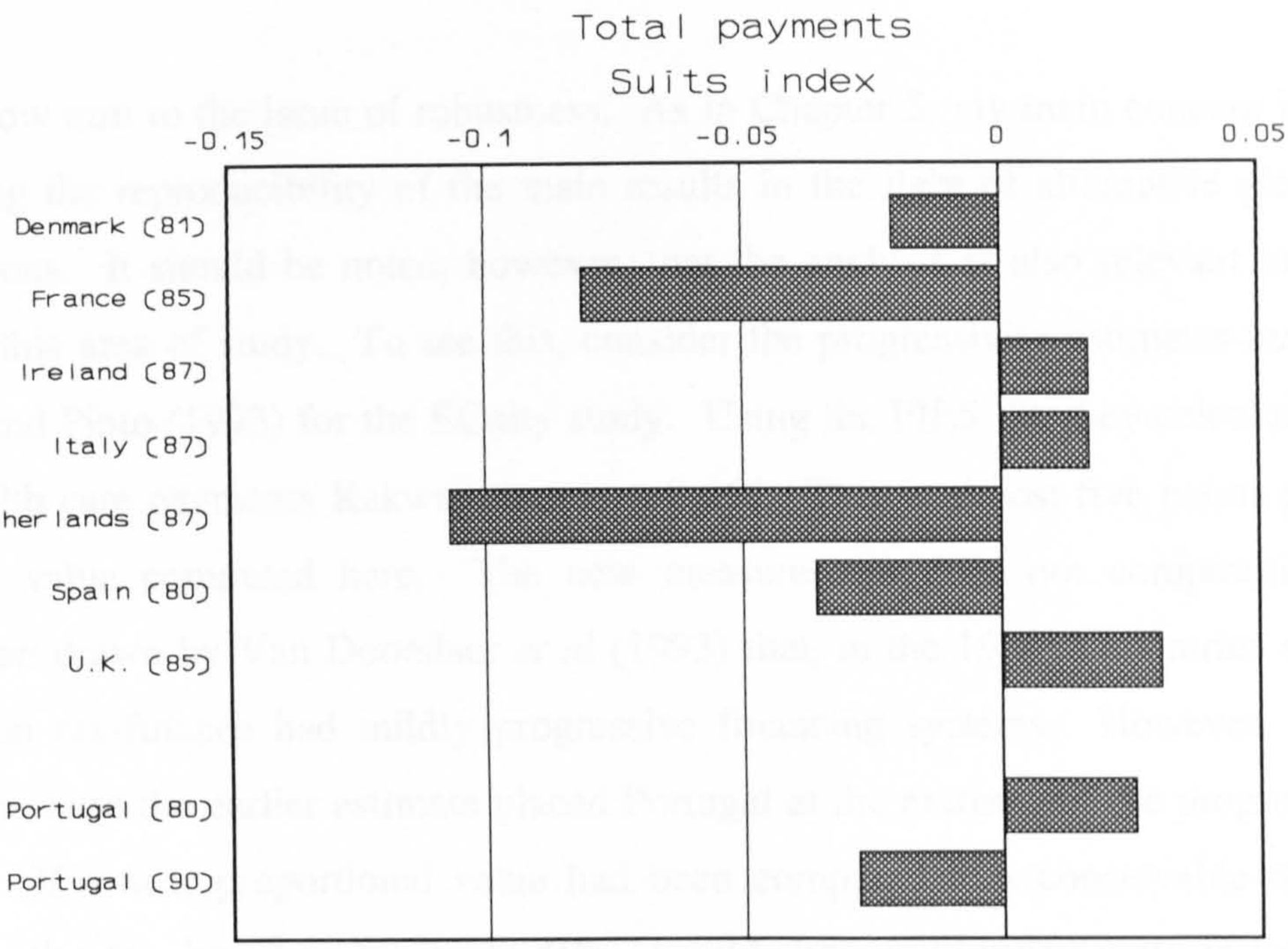


Figure 7.1: International Comparison of Health Care Finance Progressivity[17]

most recently available household budget survey in each country (in the 1980's). There are as yet no time comparisons, but the results provide an instructive means of situating the portuguese results. Figure 7.1 presents values of the Suits index for the health care financing systems of seven EU countries with the comparable estimates made in this chapter for Portugal in 1980/81 and 1989/90.[18] In the earlier period, Portugal had the second-most progressive health care system. By the end of the 1980's, however, it had slipped to fifth place in the ranking, with only the predominantly, social insurance based systems of France, Spain and the Netherlands showing higher regressivity values. Naturally, other countries may also have changed their relative position. However, the comparison does suggest that, throughout the 1980's, a fundamental change took place in the level of health care financing progressivity in Portugal.

---

[17] The results for countries other than Portugal are taken from Wagstaff *et al* (1992, p. 380).

[18] In the Appendix (Figure A5.2) similar comparisons are shown for the various sources of health care finance.

7.4 Sensitivity Analysis

I now turn to the issue of robustness. As in Chapter 5, my main concern is with examining the reproducibility of the main results in the light of alternative plausible assumptions. It should be noted, however, that the analysis is also relevant to other work in this area of study. To see this, consider the progressivity estimates made by Pereira and Pinto (1993) for the ECuity study. Using the FIES 80, they calculated the total health care payments Kakwani index as 0.065. This is almost five points greater than the value computed here. The new measurement does not compromise the conclusion drawn by Van Doorslaer *et al* (1993) that, in the 1980's, countries relying mainly on tax-finance had mildly progressive financing systems. However, this is mainly because the earlier estimate placed Portugal at the extreme of the progressivity ranking. If a near proportional value had been computed it is conceivable that the present estimate - based on the same data - would show mild regressivity in the early 1980's. The discrepancy in estimates between the two studies is largely due to the use of grouped and ungrouped data.[19] As shown in Chapter 4, estimation of concentration indices from grouped data leads to inaccurate results. However, the curve approximation procedure is not the only source of bias. In Pereira and Pinto (1993) equivalizing the income variable through application of the Buhmann *et al* (1988) formula to the mid-points of gross-income categories led to downward bias on the Gini and consequently to higher progressivity estimates. The results are also distinct because of alternative conjectures (*ie.* a different equivalence scale was applied and the incidence of corporate taxation was assumed identical to other direct taxes included in the survey).

What this example shows is not that the earlier results are invalid, but that they may be sensitive to methodological choices. Therefore I consider in this section the impact of alternative assumptions on the degree of measured inequity. So as not to overburden the analysis, measurements are made simply with regard to the Kakwani index. The investigation considers six issues over which there is uncertainty in the

---

[19] In the ECuity study, only the Portuguese and UK health care financing estimates were based on previously tabulated data.

literature: aggregation of health care payments inequality at different points of the income distribution; assumptions regarding the incidence of corporate taxation; equivalent scale adjustments to the income variable; the method of weighting observations; choice of ability to pay proxy; and equivalization of the payments distributions. The empirical strategy involves *ceteris paribus* simulations: for each issue only the variable under discussion is allowed to change; all other methodological assumptions made in section 7.3 are left unaltered. I do not address issues related to data quality, important as they are. Such questions are best handled by contrasting the results with those obtained from improved data sources, when these become available.

7.4.1 Different distributional perspectives

The first check for robustness involves the use of the parametrically weighted Kakwani indices, $\pi_K(\delta)$, defined in equation (4.14). These measures incorporate explicit assumptions about the weight attached to different points of the income distribution; and may therefore be seen as reflecting alternative judgements about the degree of progressivity preference. This is an important issue given that the normative basis of health care finance vertical inequity measurement is somewhat debatable (see, *eg.*, Aaron, 1992). There are no clear guidelines from policy-makers as to the desired degree of progressivity, merely a stipulation that payments be related to ability to pay. Economic studies (including the present) have chosen to examine vertical inequity by means of standard progressivity indices derived from the concentration approach. As suggested in Chapter 4, these measures have desirable properties, but they are simply statistical devices that measure deviations from proportionality. In using them, the researcher implicitly accepts that progressivity should be portrayed in terms of relative payments distances and relative income distances (between equidistant incomes) and that proportionality of payments is the neutral reference system. The measures also have specific weighting schemes for aggregating disproportionality which may, or may not, accord with the values of the policy maker.

Computation of generalized progressivity indices provides a partial, but important, response to these problems. By varying a single parameter, the robustness of progressivity estimates to different distributional judgements (including, perhaps, those

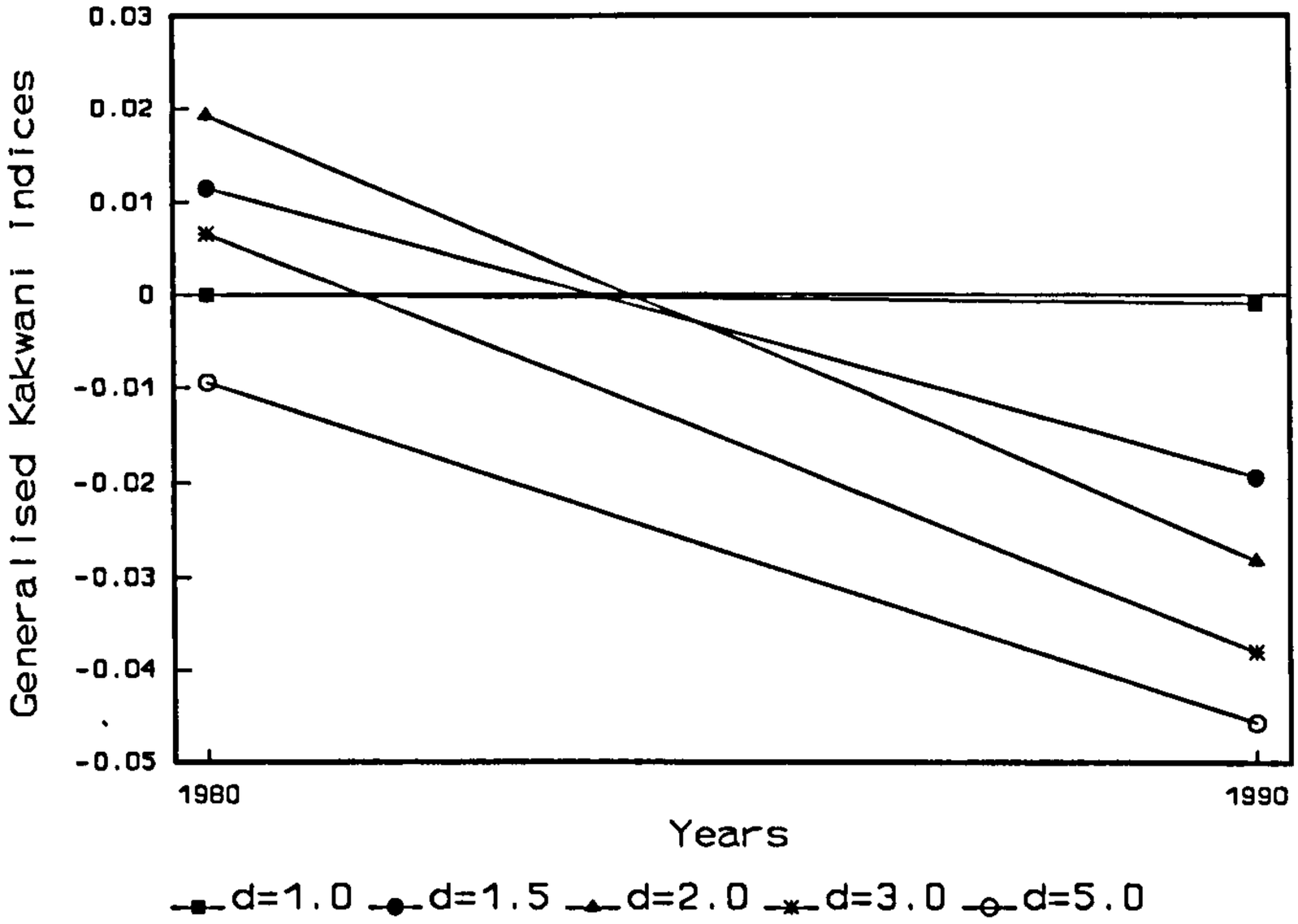


Figure 7.2: Generalized Kakwani indices for overall health care financing.
Values of δ = 1.0, 1.5, 2.0, 3.0 and 5.0

of the policy maker) can be explored. Furthermore, the alternative weighting schemes of the parametric measures may be seen as emulating the range of measures suggested by progressivity indices derived from an explicit normative approach [*eg.* those of Blackorby and Donaldson (1984) and Kiefer (1985)]. If the results obtained from varying the distributional judgement parameter point in the same direction, then the robustness of conclusions is improved.

Values of $\pi_K(\delta)$ in the range δ=1 to δ=5 were computed for each the financing sources as well as for the total payments distribution. The results are shown in Figure 7.2 and Table A5.1. For aggregate health care payments (Figure 7.2), the measurements generally suggest a shift towards regressivity between 1980 and 1990. The only exception is the inequality indifferent index, δ=1.01. Therefore, unless one is unconcerned about disproportionality of the payments distributions, it is hard to escape the conclusion that the distribution of health care payments became less favourable to the poor throughout the 1980's.

The estimates for the source distributions also tend to confirm the earlier results

(Table A5.1). Where there are exceptions, the new measurements may be seen to provide further informational content. For instance, the δ=5 index suggests that indirect taxes became more progressive in the later period, whereas at lower values of δ an opposite movement is suggested. This is because the δ=5 measure is highly sensitive to disproportionality in the lower end of the income distribution, leading it to detect an improvement in the position of the very poorest.[20]

7.4.2 Changing tax-incidence assumptions

Another area where the results may be challenged is with regard to incidence assumptions. Research in the taxation literature has shown that judicious choice of shifting assumptions can make a tax system appear either steeply progressive or sharply regressive (Whalley, 1984). It is also the case, however, that apart from three types of contribution - corporation taxes, property taxes and employer social insurance contributions - a set of 'standard' incidence assumptions appears to have gathered wide agreement in the partial equilibrium literature (Atkinson and Stiglitz, 1980). Of the exceptions, only the first is quantitatively important in the financing of portuguese health care. Therefore, two further assumptions are considered in this section with a view to evaluating their impact on the earlier conclusions: (i) assuming that corporate tax incidence falls exclusively on dividend recipients, and (ii) that the burden is fully shifted to consumers in the form of higher prices.[21]

The results (shown in Table 7.5) confirm that changing incidence assumptions can have an appreciable effect on the progressivity of the tax system. For instance, estimates of the direct tax Kakwani index for 1980 vary between 0.151 and 0.301; for the tax system as whole, between 0.066 and 0.118. Unsurprisingly, the full shifting assumption produces the least progressive results. However, the important result from

---

[20] This improvement may also be seen by careful reading of Table 7.1.

[21] The first of these would be likely to hold in the case of a long-run competitive equilibrium situation with intersectorally mobile capital. The second, if markets are oligopolistic and firms have the power to set their prices to cover what they regard as costs plus a margin for profits. Other justifications are also possible [see, *eg.*, Pechman (1985)]. The initial assumption of a 50:50 split between dividend recipients and consumers may be seen as an intermediate compromise between these extreme hypotheses.

Table 7.5: The impact of alternative assumptions concerning corporation taxes.
Kakwani indices

	Variant 1		Variant 2		Variant 3	
	1980	1990	1980	1990	1980	1990
Direct taxes	0.227	0.127	0.301	0.188	0.151	0.071
Total taxes	0.092	0.047	0.118	0.069	0.066	0.025
Total health care payments	0.019	-0.027	0.036	-0.016	0.002	-0.040

Notes: Variant 1 - Incidence falls 50% on capital income receivers and 50% on consumers.
Variant 2 - Incidence falls on dividend recipients.
Variant 3 - Incidence falls on consumers.

the point of view of this chapter is that the conclusions regarding the evolution of health care financing progressivity throughout the 1980's are not altered. Whichever of the three variants is chosen there is a shift of some 4 or 5 points in the Kakwani index towards regressivity. Thus, once again the earlier results are shown to be robust.

7.4.3 Sensitivity to equivalence scales

While shifting assumptions are perhaps the most crucial part of any incidence calculation, the 'ability-to-pay' measure is also very important (Whalley, 1984). The following three sensitivity analyses consider the impact of alternative specifications of this variable, beginning with the effect of equivalence scale relativities. The baseline estimates were arrived at by deflating gross household incomes by the OECD equivalence scale. For reasons set out in Chapter 5, the choice of this scale was largely determined by pragmatic considerations (*eg.* its frequent use in current portuguese research). This detail, together with the uncertainty surrounding the equivalence scale issue (Coulter *et al*, 1992), suggest a need to verify the robustness of the earlier measurements to different scale relativities. The approach adopted for this purpose is analogous to that followed in Chapter 5. Values of the Kakwani index are computed

assuming various equivalent income distributions; namely, through deflation of income values by the Buhmann *et al* (1988) parametric scale with family size equivalence elasticities in the range e=0 to e=1.[22]

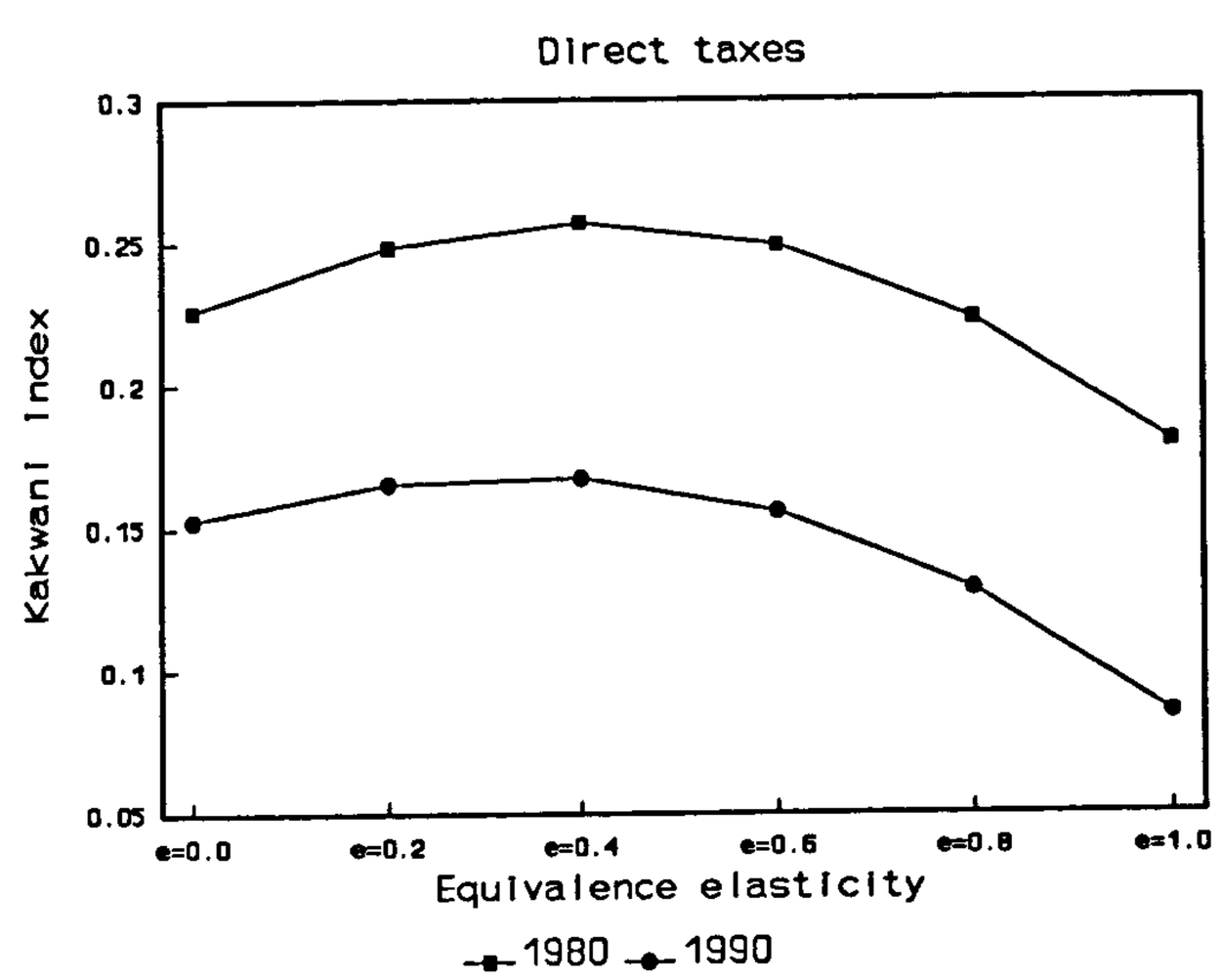
The results are presented in Figure 7.3. The most salient point is that the 1980-1990 change in the overall payments distribution is robust in terms of all scale relativities. Therefore, whichever the relationship between economic well-being, household incomes and 'needs' that is assumed, one is driven to the conclusion that health care finance became less favourable to the poor over the 1980's. The earlier conclusions regarding the time-trend of taxation and direct payments progressivity also appear to be upheld. Conversely, the direction of changes in the social and private insurance distributions seems to be affected by scale relativities. For example, at low e-values the social insurance estimates show a marked decline in progressivity but the effect is reversed with scales that are 'generous' to large families. The reasons for this are not immediately clear, but it may be noted that in both surveys insurance contributions are largely made by smaller families. A possible implication is that progressivity estimates for non-universal sources of finance may be sensitive to equivalence scale adjustments; particularly if non-income characteristics determining participation are correlated with the factors used to construct the scale.

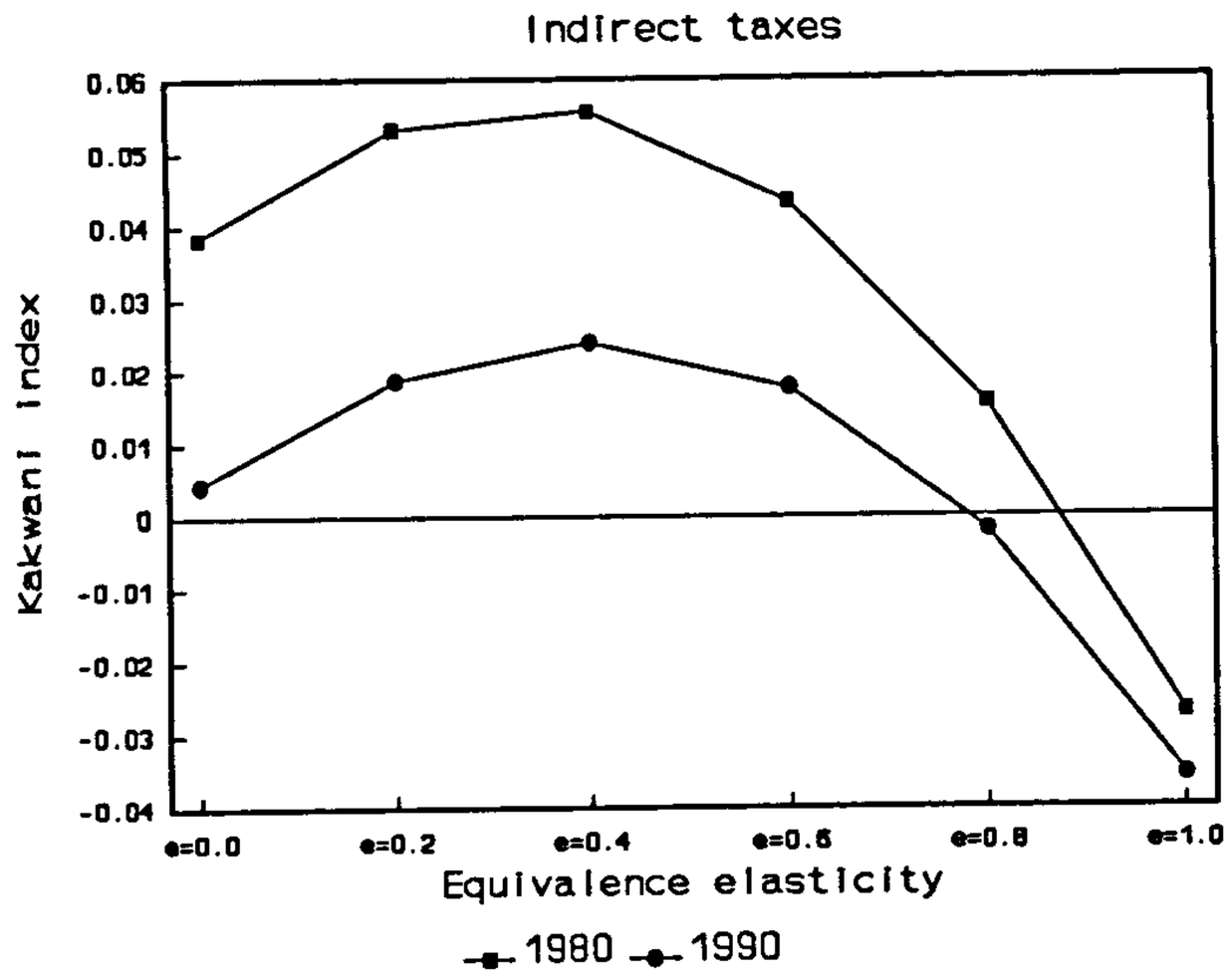
A further issue of interest is the shape of the relationship between levels of progressivity and choice of equivalence scale. For all sources there appears to be an inverted U-shaped relation, with progressivity first increasing (regressivity decreasing) and then decreasing (increasing) as the value of the equivalence elasticity is increased.[23] This finding is relevant to future research and, indirectly, to policy analysis. It means that one cannot simply compute progressivity measures for two extreme equivalence scales and assume that intermediate scales will lead to intermediate

---

[22] See equation (5.4) in Chapter 5.

[23] Generally, this result seems to be driven by the interaction of a U-shaped effect on the Gini coefficient (first noticed by Coulter *et al*, 1992a) and an inequality reducing quasi-linear impact on the concentration indices (see Figure A5.3 in the Appendix). Recall that for the illness concentration indices, examined in Chapter 5, the effect of increasing e was also to reduce the level of measured inequality.





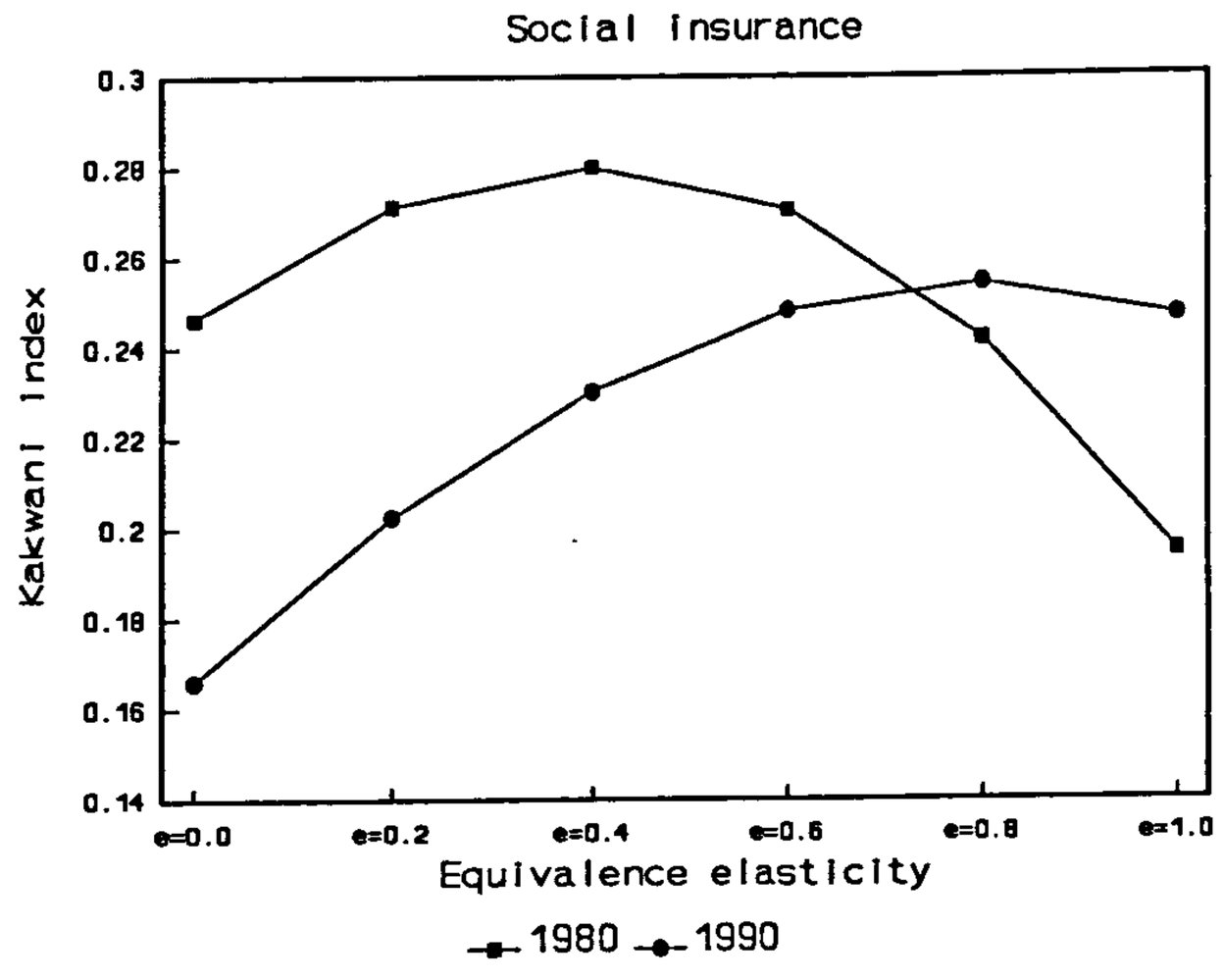
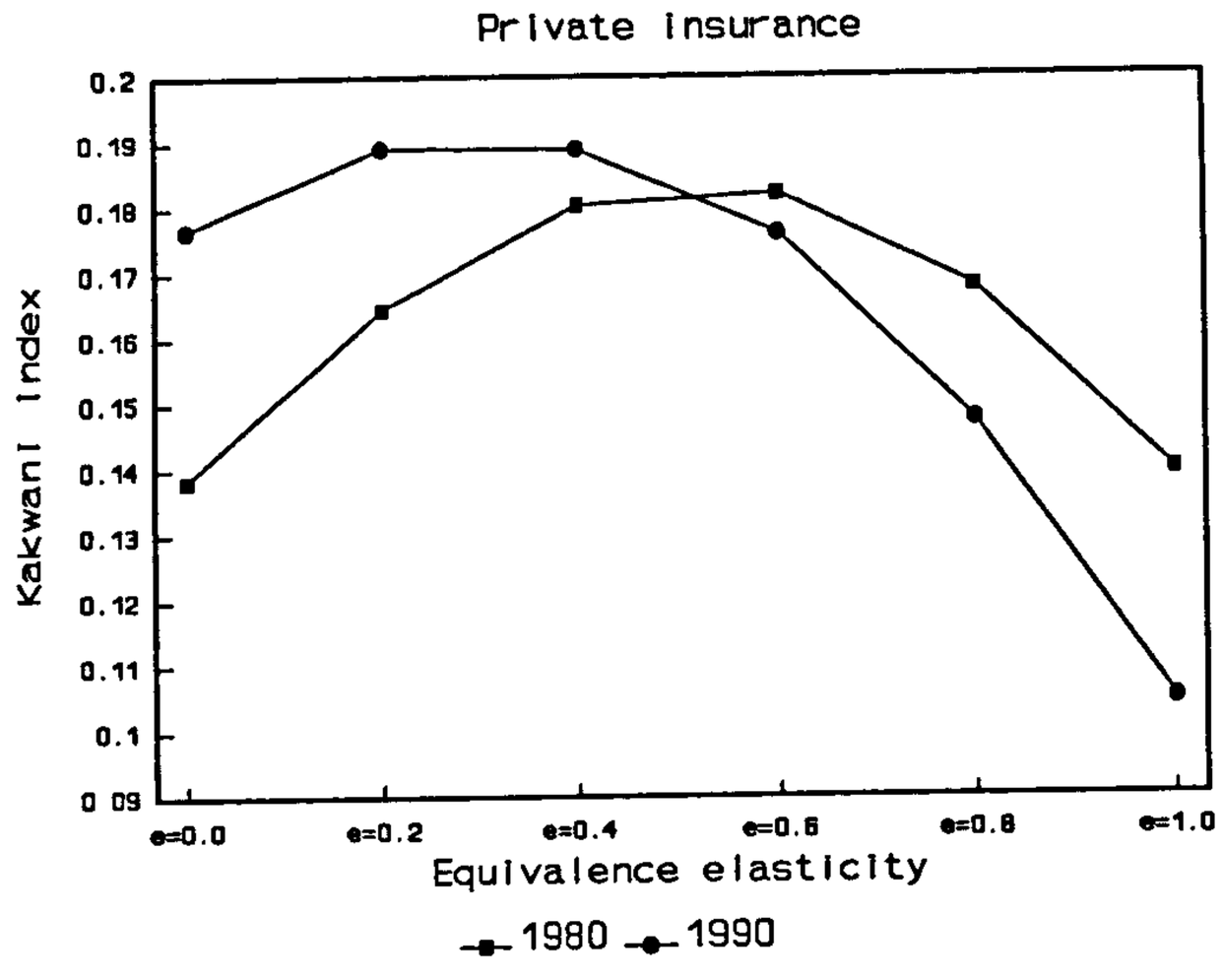
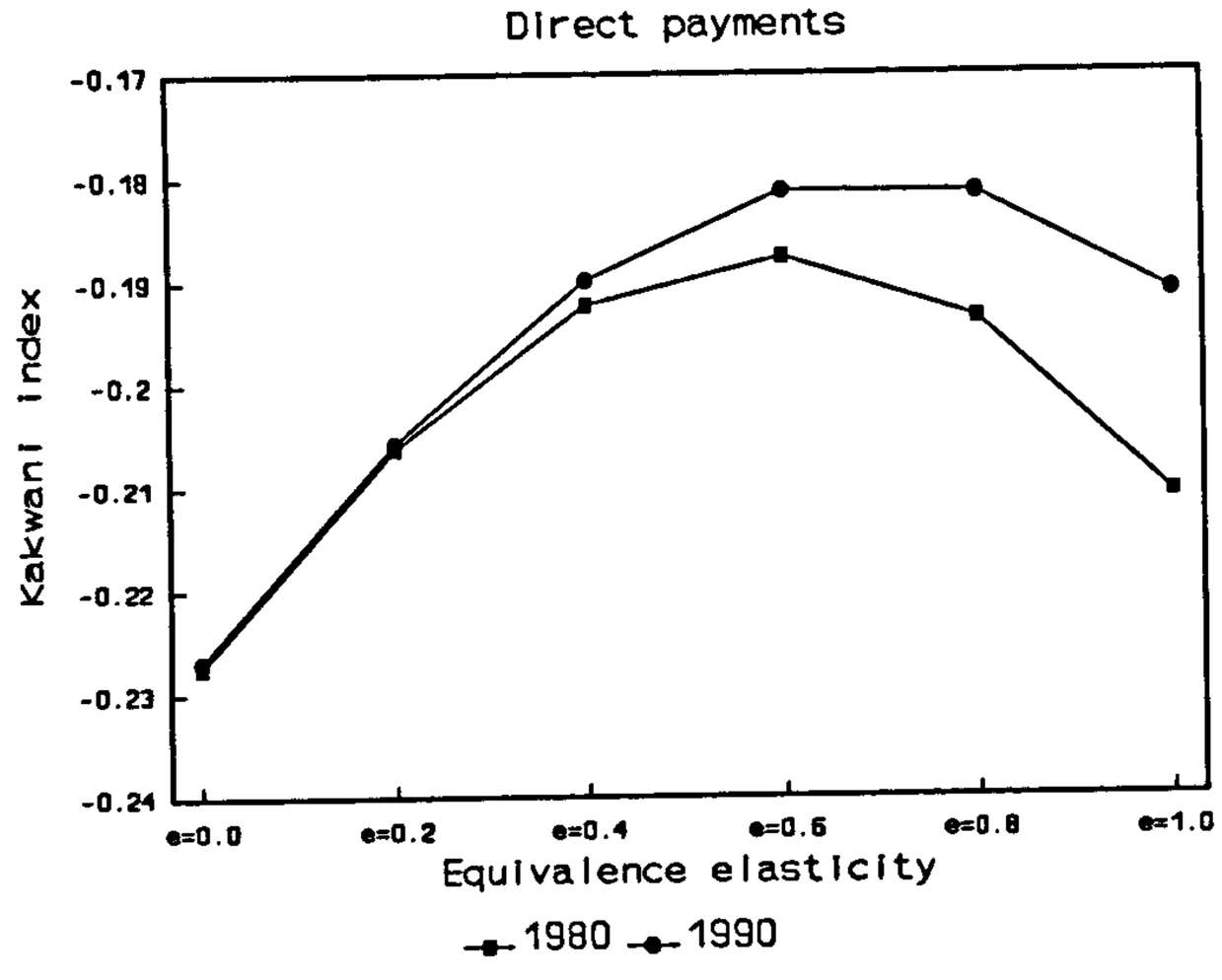


Figure 7.3: Impact of equivalence scale relativities on health care finance progressivity (cont.)





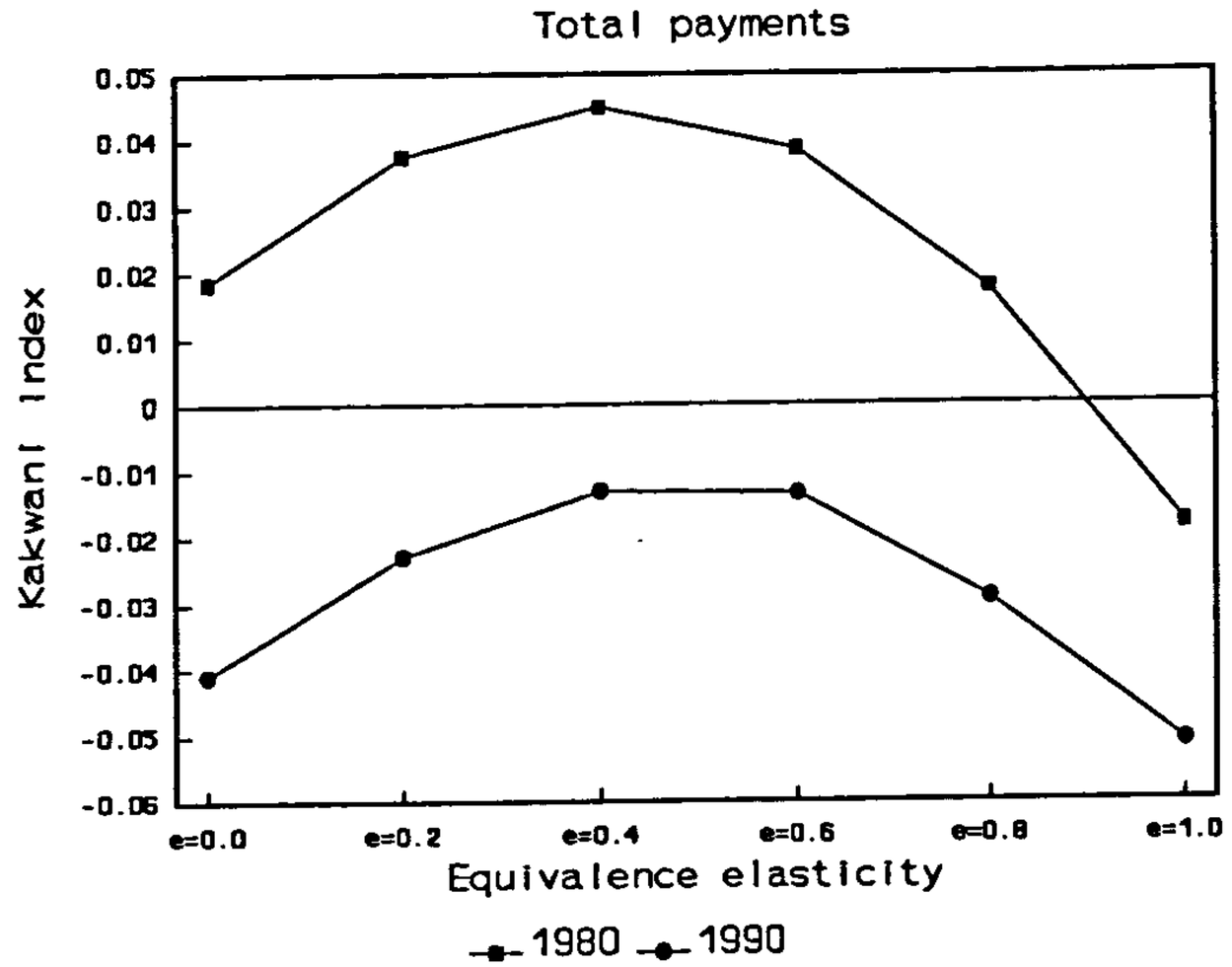


Figure 7.3: Impact of equivalence scale relativities on health care finance progressivity

progressivity estimates. The results also help to put into perspective other research findings. The most recent work by the ECuity group (Wagstaff *et al*, 1994) proposes to compare health care finance progressivity across countries by drawing on the equivalence scale used by Aronson *et al* (1994). As shown in Chapter 5, this scale has an implicit $e \approx 0.44$. The present calculations suggest that the scale provides a *higher* estimate of the extent of health care finance progressivity than do other scales currently used by economists.

7.4.4 Weighting by individuals

If income is used as the 'ability-to-pay' proxy, it raises the question of how the income receiving units are to be weighted. Compared to equivalence scale adjustments, this issue has received far less attention in the literature. However, it can have a considerable impact on the measure of income inequality (Danziger and Taussig, 1979), and consequently on progressivity estimates. The earlier results were weighted by household (*ie.* the equivalent household income is counted only once for each household irrespective of the number of individual members), a procedure which has been common practice in health care finance progressivity analyses. However, recent income (re-)distribution studies have generally opted to weight by the number of individuals, given that the former approach begs the question of the number of individuals affected by economic differences. Like all other questions considered in this part of the chapter, the correct approach is debatable. Hence, the value of modifying the assumptions to see if they do in fact affect the conclusions that are drawn.

Individual weighting requires that each individual in the household be attributed the household equivalent income (*ie.* the procedure adopted in Chapter 5). The results of admitting this conjecture are reported in Tables 7.6, A5.3 and A5.4 (along with those for remaining sensitivity analyses). Individual weighting has the effect of reducing the inequality estimates. For virtually all sources of finance, the reduction in the concentration index is greater than for the Gini coefficient. Consequently, levels of progressivity are reduced in relation to the baseline distribution. The more important point, however, is that the magnitudes of changes between 1980 and 1990 are basically

Table 7.6: Impact of different assumptions regarding weighting of income units, ability to pay proxy and equivalization of payments. Kakwani indices

	Baseline		Individual weighting		Net income		Equivalent payments		Combined assumptions	
	1980	1990	1980	1990	1980	1990	1980	1990	1980	1990
Equiv. income (Gini)	0.343	0.351	0.324	0.327	0.332	0.326	0.343	0.351	0.315	0.305
Direct taxes	0.227	0.127	0.217	0.087	0.162	0.110	0.257	0.166	0.188	0.096
Indirect taxes	0.019	-0.002	-0.001	-0.023	0.016	0.017	0.045	0.011	0.045	0.007
Total taxes	0.092	0.047	0.076	0.019	0.067	0.052	0.119	0.069	0.096	0.042
Social insurance	0.245	0.244	0.218	0.241	0.250	0.279	0.290	0.244	0.275	0.270
Private insurance	0.175	0.152	0.240	0.121	0.177	0.161	0.128	0.183	0.223	0.148
Direct payments	-0.196	-0.186	-0.181	-0.159	-0.191	-0.160	-0.209	-0.218	-0.154	-0.160
Total payments	0.019	-0.027	0.012	-0.034	0.004	-0.014	0.036	-0.028	0.036	-0.020

Note: Combined assumptions = individual weighting, net income as reference distribution, equivalization of payments (OECD equivalence scale).

the same as for the earlier results. Thus, individual weighting has no appreciable effect on overall conclusions.

7.4.5 Net income as a measure of ability to pay

A further issue on which there are grounds for disagreement is the use of gross income as the reference distribution. Once again this is common practice in the health care finance literature. However, it seems illogical to measure the distribution of direct payments, private insurance premiums and indirect taxes in relation to this distribution since what effectively constrains households are their disposable incomes (*ie.* after direct taxes and other contributions).[24] This suggests that distinct reference income distributions should be used for different types of payment. Future analyses may wish to consider this issue in greater detail. Meanwhile, it is useful to measure the impact of using disposable income as a measure of ability to pay to calculate the progressivity indices reported earlier. The precise income definition that is used is household income net of state and local direct taxes and social insurance contributions, equivalized by the OECD scale.

The results indicate quite naturally that the value of the Gini coefficient is reduced *vis-à-vis* the baseline distribution. Inequality in the disaggregated payments distributions also declines, leading to varied measurement and intertemporal effects (see Tables A5.3 and A5.4). The impact on the overall payments distribution is to make the 1980's change towards regressivity much smaller than under the baseline assumptions (two points of the π_K index as opposed to 5). This implies that in the event of 'source-appropriate' reference distributions being chosen, the fall in progressivity would not be as great as shown in section 7.3. However, the results also indicate that this is simply a question of the size of the effect; the general conclusions reached earlier would still be valid.

---

[24] Studies of the progressivity of indirect taxes do in fact adopt net or disposable income as the reference distribution [see, *eg.* Kakwani (1986, Chapter 10)].

7.4.6 The effect of equivalizing payments

Besides incidence assumptions, there are other plausible conjectures that might have been adopted for the payments variables. Chief among these is the equivalization of household payments to take account of differences in demographic structure. As in the majority of previous studies, the approach was not followed in section 7.3 given that it is generally accepted that economies of scale do not operate in health care consumption. However, this conjecture only seems valid for direct expenditures and even then it is possible to think of exceptions to the rule. With regard to other forms of payment, the no-equivalization hypothesis seems harder to defend (*eg.* the case of two families of size 1 and 5 each with a single wage earner making social insurance contributions that provide benefits to all family members). Moreover, recent health care financing progressivity estimates have opted to equivalize payments (*eg.* Wagstaff *et al*, 1994).[25] Consequently, it seems appropriate to measure the impact of adjustments for household structure in the payments distributions.

Household payments towards the health service were equivalized using the OECD scale. The progressivity estimates produced by adopting this assumption tend to have the opposite effect to the two previous sensitivity analyses. That is, the progressivity (regressivity) of progressive (regressive) sources is increased. The conclusions regarding the intertemporal change in health care finance progressivity are not greatly affected, though the magnitude of the changes is more pronounced than with other assumptions. There is almost a 7 point difference between the 1980 and 1990 overall payments Kakwani indices. Thus, equivalization of payments also leads to the later distribution appearing more inequitable than that of the earlier period.

7.4.7 Combining alternative assumptions

As a final step in the analysis indices were computed by combining the final three assumptions (see Table 7.6). Therefore, besides other conjectures previously laid out

---

[25] This option appears to have emerged simply because of the motivation to measure redistributive effect, in which case both income and payments distributions have to be equivalized in order to achieve comparability.

in section 7.3, the methodology now admits individual weighting of income units, equivalized net income as the counterfactual distribution and health care payments equivalized by the OECD scale. The purpose of this step is to verify if the results are robust to a combination of alternative plausible assumptions, rather than admitting a single variation at a time. The results confirm the general intertemporal trends identified earlier, in particular the shift from mild progressivity to mild regressivity of the overall health care payments distribution. Noticeably, the absolute change in the Kakwani index is greater than under the baseline assumptions.

Therefore, the sensitivity analyses generally show that the earlier estimates are robust in terms of alternative methodological assumptions. Although there may be disagreement among economists as to particular procedures, some relatively strong conclusions can be drawn about the structure and recent course of health care finance progressivity in Portugal.

7.5 Conclusions

In Portugal, as in many other countries, there appears to be widespread support for the objective of health care payments being related to ability to pay rather than to use of medical facilities. The goal was a major reason behind the creation of a National Health Service in 1979 and, somewhat surprisingly, has also been championed by those who now call for an expansion of alternative forms of finance. What is more striking, however, is that the debate has unfolded in a vacuum of evidence regarding the economic burden of payments. The present chapter has sought to inform the debate by examining how the progressivity characteristics of the main sources of finance have evolved in recent years and, more generally, how the system has performed with regard to the ability-to-pay goal.

It was shown that, throughout the 1980's, a fundamental change took place in the distribution of health care financing. Total payments to the health system, which in 1980 revealed a mildly progressive structure, had evolved by the end of the decade towards a moderately regressive disposition. The burden of health care finance appears

to have shifted to middle income groups, with the main beneficiaries being households situated in the richest quintile of the income distribution. This change is the result of two major trends: on the one hand, reduced progression in the tax system; and on the other, an increase in the share of revenues raised directly from consumers.

It was also shown that alternative forms of finance have quite distinct progressivity characteristics. Direct taxes, despite a noticeable change throughout the 1980's, are highly progressive. Indirect taxes tend to be close to proportional. The tax system as a whole clearly favours the least well-off. Social and private insurance contributions were found to be globally progressive in both periods, but this is largely explained by the phenomenom of selective coverage of households that are better-off. Out-of-pocket payments were shown to be highly regressive and there are signs that a significant part of the regressive structure is the consequence of policy choices. On the whole, these results tend to confirm the findings of previous studies in other countries.

It is possible to disagree with the results on two main fronts: the methods of analysis and the data used. As to the first issue, in contrast to previous research, considerable effort has been given over to measuring the impact of methodological choices on the results. Under all the alternative scenarios considered the conclusions tend to be very much the same, varying only with respect to the *degree* of progressivity/regressivity. Therefore, the main results would appear to be robust in terms of a number of important alternative assumptions. Clearly there are further methodological issues that need to be investigated. Future studies might, for example, consider the question of measuring progressivity of payments to the NHS, rather than for the system as a whole.

The second issue is far more complex. Doubts may be raised with regard to the suitability of budget survey data for examining the broad questions raised by vertical inequity analysis of health care financing systems. For instance, the estimated distribution of the corporate tax burden can only be seen as a very rough approximation to the actual burden. More specifically, trends in the progressivity of the tax system may be questioned because of the overlap of the 1989 reform with the survey observation period. Given that these trends were identified as a major driving force

behind the overall change in the health care financing distribution, there is clearly a need for further research using new data sources. However, given present data limitations, the estimates in this chapter are as reasonable an approximation to the evolving progressivity characteristics of the health care financing system as can be made. Future investigation should seek to verify the conclusions by replicating the methodology on improved or more disaggregated data. Alternatively, methods that are better able to *explain* the results (*eg.* econometric analysis based on behavioural models) should be employed. This is particularly apt with regard to the distribution of out-of-pocket payments.

Part 5

FURTHER METHODOLOGY

Chapter 8

Measurement of Social Welfare in the Domain of Health

"[N]o-one, as far as I know, proposes that the elimination of inequality ... is the sole objective of policy. For evaluating policy, the appropriate criterion has to be a measure of welfare which combines tendencies of the mean and of the distribution around that mean."
Pyatt (1985, p. 89)

8.1 Introduction

Parts 3 and 4 were concerned with formulating indices of relative inequality and applying them in the measurement of inequity in health and health care. Although analysis of this type is likely to form the backbone of an economic approach to studying the attainment of health-equity objectives in a particular country, it has to be said that, from the policy-makers' perspective, it provides a somewhat limited basis for developing and evaluating competing strategies. It cannot, for example, address key questions of policy such as the trade-off between equity and efficiency or the marginal impact of different inputs on the distribution of health. There is, however, a ready answer to this dilemma under the normative framework adopted for the thesis: measurement of the socio-economic distributions of health and health care constitutes only a *part* of the implied empirical analysis. It provides, in our terminology, incomplete assessments of both *achievement* and *advantage*. Once these attributes are measured with greater adequacy, the type of policy questions referred to can indeed be addressed.

The objective of the present chapter is to investigate how a general, coherent evaluation of society's *achievement* might be undertaken. In particular, the chapter seeks to develop an ethically-founded summary measure, that draws on the information content of the inequality indices discussed previously. In so doing it provides a further extension to concentration curve analysis in the domain of health.

For an individual, *achievement* is simply the actual level of the functioning 'good health' that he enjoys (*ie.* his health status). Clearly, if we are to measure how a given society fares in relation to this attribute, the first task must be to identify the values it holds *vis-à-vis* the health of constituent members. This amounts to specifying a Social Welfare Function (SWF) in the particular domain of health. The SWF is a concept much used in modern welfare economics but also sometimes misunderstood. A widely used textbook describes it as "*a relation between a distribution of utility levels among society's members and a judgement about the overall satisfaction ... achieved by that distribution*" (Friedman, 1985, p. 43). Under this definition, the evaluation of the social state is based exclusively on the individual utilities that are generated, an approach often referred to as 'welfarism' and which is dominant in modern economics. It is, however, unsuited to our purpose, given the limited and sometimes inaccurate nature of utility information as a basis for social evaluation in the field of health (see Chapter 3). Fortunately, the SWF need not be restricted to the 'welfarist' straitjacket. In Bergson's (1938) seminal article the concept is simply a representation of the 'goodness' of the social state. No specific restrictions are placed on its arguments or functional form and the appropriate determination of what is 'good' about a given situation is left completely open. In particular, despite the 'welfarist' sound of the term *social welfare function*, use of that concept does not presuppose a utilitarian or welfarist formulation of social objectives. Drawing on this eclecticism, it seems altogether more fitting to postulate a specific health domain SWF, defined directly over a vector of the population's health [*ie.* $W \equiv W(\mathbf{h})$, with $\mathbf{h} = (h_1, h_2, \ldots, h_n)$]. Naturally, there is no question of this function being taken to represent the overall well-being of society, but simply that part of it which is associated with the population's health. In other words, W represents *health social welfare*, which is a concept in every sense identical to *society's achievement.*

The following section considers how the representation given above might be further characterized with a view to developing an empirical summary measure of health social welfare. The discussion inevitably concentrates on the objectives of equity and efficiency, widely conjectured to be the major goals of social intervention in the sphere of health. A rare attempt by a health economist at integrating equity and efficiency measurement is discussed. Although that approach - proposed by Wagstaff (1991) -

shares some of the concerns of the present chapter it is rejected in favour of a more direct and ethically appealing method with some traditions in the income inequality literature. Two general principles for evaluating the social well-being associated with health are defined, and from these follows the specification of an *abbreviated social welfare function* of the form: $W(\mathbf{h}) = V(\mu_h, I_h)$, where μ_h is the average level of health in the community and I_h is a health inequality index. A question arises as to which of the many inequality indices available should be incorporated in this function. Section 8.3 presents a formal rationale, based on the concept of *income related health deprivation*, for using the health concentration index C_h (*ie.* an index identical to C_m in Chapter 4 but defined over health status). This analysis is important in two respects. On the one hand, given that our objective is to construct an empirically applicable measure of health social welfare, it is better to know the exact basis on which one distribution is to be judged superior to another. Equally important, the analysis establishes a logical role for concentration indices, which are widely believed, even by their proponents [*eg.* Wagstaff *et al* (1991b, pp. 200-201)], to have unappealing welfare implications. The principal result of the chapter is a rationalization of a family of welfare indices of the type, $W_h = \mu_h(1 - \kappa C_h)$, where κ is a distributional judgement parameter, that allows for a qualified trade-off between equity and efficiency objectives. Section 8.4 provides an empirical illustration - drawing on the INS data used in Chapter 5 - which compares levels of health social welfare for Portugal's five regions. Section 8.5 extends the analysis by disaggregating the family of welfare indices according to income class, a point that is of major interest given the overall objectives of the thesis. A further empirical illustration using the INS 87 is provided in 8.6, this time comparing levels of health social welfare by income groups in Portugal as a whole. The final section presents some concluding remarks.

8.2 Social welfare functions and the equity-efficiency trade-off

The Social Welfare Function is a construct that is rarely used to guide questions of measurement or policy evaluation in the health economics literature. Nevertheless, the implicit form that underlies the greater part of research in the field is relatively clear. In particular, the traditional 'welfarist' approach to evaluating society's well-being seems

to have been largely eschewed. There are some important reasons for this (Culyer, 1989). One is that a particular class of 'extra-welfare information' - the actual health of the community, or their *achievement* in our terminology - is of intrinsic importance when evaluating social states. Another reason, inextricably related to the nature of the health attribute, is that there is often a case for overruling individual judgements of value in favour of those adopted by a perceived or real 'ethical decision maker'. As a consequence, the bulk of health economics research that deals with questions of national policy would appear to be guided by an implicit SWF that is most naturally and directly defined in terms of the objectives of the 'decision maker' *vis-à-vis* the health distribution. There is also widespread consensus that two particular objectives - *equity* and *efficiency* - stand above all others (*cf.*, *eg.* Fuchs, 1987; McLachlan and Maynard, 1982), the implication being that variations in these attributes provide a suitable basis for monitoring improvements in society's health welfare.[1]

In this context, logic would suggest that macro-level evaluations of health well-being should be based on the *combined* measurement of equity and efficiency, thus enabling the possibility of trade-offs between the objectives to be addressed. Invariably, however, empirical studies treat evaluation of the two as separate issues. This is particularly true of research that draws on the techniques available for inequality measurement, where studies often include a brief discussion on the importance of efficiency criteria for policy evaluation, but then go on to measure inequities independently of such considerations. In much the same way, research that employs the tools of efficiency analysis tends to treat equity as "a competing dimension upon which decisions are made in addition to that of the efficient deployment of resources" [Drummond *et al* (1987, p. 33)]. Consequently, very little is done to incorporate suitable distributional criteria into actual measurement. As an example, recall the discussion of Chapter 3 on the QALY literature, where it was shown that acceptance of health maximization as a basis for allocating resources has been unduly hindered by the insensitivity of the QALY metric to widely supported distributional value judgements.

Of course, to an important extent, the dichotomy between equity and efficiency

---

[1] Another obvious implication is that policy evaluation and health social welfare evaluation essentially require the same information basis.

measurement in the empirical literature is intentional. There are, after all, a number of valid reasons for focusing exclusively on a particular objective (*eg.* current public concern, lack of previous research in that area, etc.). One suspects, however, that most health economists would readily opt for an integrated approach to policy evaluation, *if* the techniques for joint measurement of equity and efficiency were sufficiently developed. Yet it is only very recently that such a task has come to be adequately addressed (*eg.* Culyer, 1989; Wagstaff, 1991; Culyer and Wagstaff, 1991 1992). Quite naturally, this research has concentrated on probing conceptual issues. Wagstaff's (1991) article does, however, contain a practical suggestion, based on explicit construction of an SWF, as to how integrated measurement of the two fundamental objectives might proceed.

Wagstaff suggests that social welfare in the domain of health might be represented by the well-known *isoelastic* SWF, a general form of which is

$$W = \frac{\sum_{i=1}^{n} a_i (h_i)^{1-\tau}}{1 - \tau} \qquad if \quad \tau \geq 0, \ \tau \neq 1 \tag{8.1}$$

$$= \prod_{i=1}^{n} (h_i)^{a_i} \qquad if \quad \tau = 1$$

Individual health status, h_i, is assumed to be a continuous variable that can be suitably measured on an interval scale, an assumption that will be maintained, for ease of exposition, throughout the chapter. The coefficient a_i indicates the weight that society attaches to the *i*th person's health status. Recent contributions to the outcome measurement literature have suggested that weights such as these can provide all the information that is needed to capture specific equity concerns. It is argued, as a consequence, that society's health well-being is greatest when a suitably weighted sum of health states is maximized (*cf.* Williams, 1988a; Culyer, 1990).[2] Wagstaff disagrees that weights are an appropriate form of representing societal concern regarding equity, ostensibly because they cannot reveal aversion to inequality in health outcomes. This view does not seem entirely correct: if the weights are a decreasing function of

[2] This suggestion was discussed in Chapter 3. Note that if one adopted the view of the early QALY literature, that a unit of health is of equal value to whomsoever it accrues, the a_i's would be set at the same value for all individuals in the community.

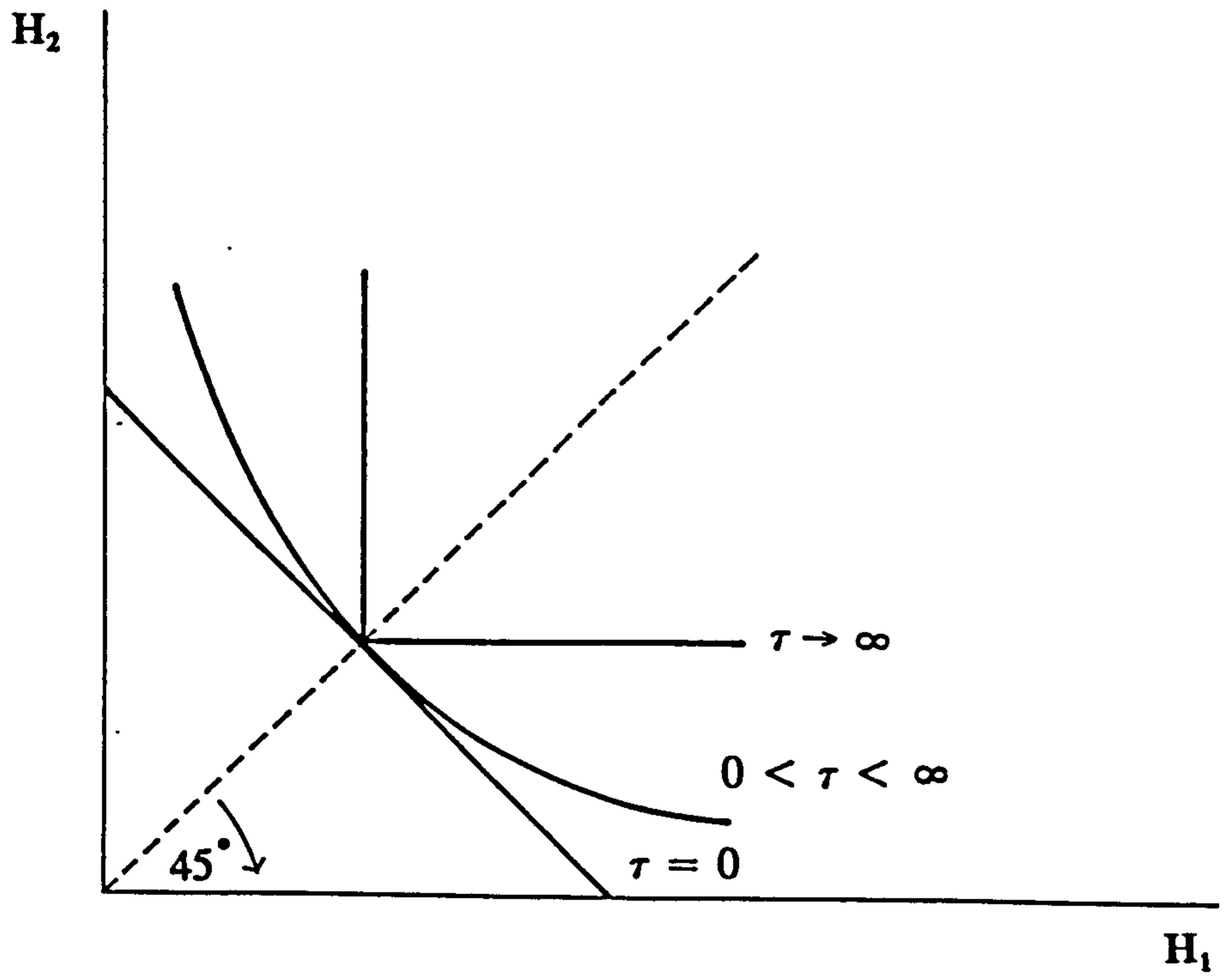


Figure 8.1: The isoelastic health social welfare function

the rank of the health distribution, then a particular form of aversion to inequality is implied. Nevertheless, Wagstaff appears to circumscribe the role of weights to indicating societal preferences regarding the process of health production (*eg.* the merit of some persons relative to others) or type of health improvements (*eg.* those that benefit a particular age group).

Specific health output equity concerns are incorporated into the welfare analysis through the mechanism of constant elasticity of substitution of the welfare contours ($1/\tau$), which allows the parameter τ to be interpreted as the degree of aversion to inequality in health outcomes. This suggestion is exemplified in Figure 8.1 which shows, for the two person society, an intermediate and two limiting cases of the isolelastic SWF. The axes indicate the health states of individuals 1 and 2, and it is assumed (for ease of presentation) that society attaches equal weight to the health of these two persons (*ie.* $a_1=a_2=1$).[3] When $\tau = 0$, social welfare is represented by a straight line with slope -1, which implies indifference to inequality in health outcomes.

---

[3] The effect of allowing $a_1 \neq a_2$ is to change the slopes of the welfare contours.

However, if $\tau > 0$, the contours of the welfare function become concave and some aversion to inequality is indicated. The limiting case $\tau \to \infty$ represents a Rawlsian-type criterion, whereby the only concern is with the health of the least healthy person in society. Thus, in an important sense, concavity of the SWF provides a means of capturing both efficiency and equity considerations - the former defined in terms of greater health output and the latter as less inequality in outcomes.

A potentially attractive feature of Wagstaff's approach is that it can be used to rationalize an empirical measure of health welfare.[4] Note that (8.1) is similar to the SWF underlying Atkinson's (1970) index of income inequality, discussed in Chapter 4. Building on Atkinson's method, suppose that the *mean equivalent level of health*, h_e, is defined as the level of health status, which if equally distributed, would produce the same amount of social welfare as that which actually exists. Thus h_e is defined implicitly by

$$W(h_e, h_e, \ldots, h_e) = W(h_1, h_2, \ldots, h_n) \quad . \tag{8.2}$$

By concavity of the SWF it follows that $h_e \leq \mu_h$, and hence the fraction of per capita health status which could be sacrificed with no loss of social welfare, if the rest were to be distributed equally, is given by

$$A(h) = 1 - \frac{h_e}{\mu_h} \quad . \tag{8.3}$$

A(h) is a measure of Atkinson-type inequality in health outcomes, bounded by zero (when everyone enjoys the same health status) and one.[5] Combining (8.2) and (8.3) it is clear that

$$W(h_1, h_2, \ldots, h_n) = h_e = \mu_h[1 - A(h)] \quad . \tag{8.4}$$

Thus, the equally distributed equivalent level of health provides a direct measure of

---

[4] It should be pointed out that Wagstaff does not address this issue directly. The argument presented here does, however, seem to be implicit in his approach.

[5] From (8.1), maintaining the assumption that $a_i = 1$ and substituting y_i, μ_y and ε respectively with h_i, μ_h and τ, it is possible to compute this measure through the form given in equation (4.2). Note that for a given health distribution, the computed value of A(h) is greater the higher the value of τ that is assumed.

health social welfare on the health scale, showing the extent to which health inequality reduces the level of welfare below that of per capita health status. Indeed, equation (8.4) can be viewed as specifying static trade-offs between efficiency (a higher average level of health) and equity (reduced inequality in outcomes).

Despite the apparent appeal of this argument, the Wagstaff approach has some obvious limitations which render the index implied by (8.4) largely unattractive as a measure of health social welfare. First of all, the concern for inequality derives essentially from the process of aggregating health levels: it is not *intrinsic* to the evaluative structure. Most modern societies, however, reveal an explicit and *autonomous* concern for inequity, which seems to suggest that relative differences between health levels should be given a meaning independently of how much they count in aggregation across persons. A second related drawback is that $A(h)$ does not actually measure inequality in the accepted sense: it is really an index of the welfare-loss or inefficiency that arises from a less than optimal distribution of total health.[6] Thirdly, Wagstaff's SWF assumes that the welfare of individuals is unaffected by feelings of sympathy, envy or deprivation in relation to others: it depends exclusively on own health levels. This seems unreasonable when the interdependence of individual welfare has long been recognized as an important issue in determining socially desirable distributions of health and health care (*cf.* Culyer, 1980). Presumably, such phenomena could be incorporated as weights a_i, but this is unsatisfactory given that no explanation is provided of how these weights come about. An alternative which allows individual levels of welfare to depend on the interpersonal distribution of health *or that of its determining variables*, seems altogether more adequate. Note that this does not necessarily imply sovereignty of individual preferences, but merely suggests that the 'ethical decision maker' takes into account commonly observed individual phenomena in defining society's health welfare.

A final decisive weakness of the Wagstaff approach is the particular type of health-equity concern which it embodies. It is assumed that society preoccupies itself with differences in health status *per se*, whereas what generally motivates concern for

---

[6] Recall the arguments of Chapter 4 rejecting the Atkinson method as a means for measuring inequality in health attributes.

inequalities is their systematic relation to economic status.[7] The problem which this issue raises is seen very clearly in the context of Figure 8.1. Consider the two situations, where only the following information is available : (i) individuals 1 and 2 are respectively 20 and 80 years old, and (ii) they are both 20 but one is poor and the other rich. Although concave welfare contours, indicating aversion to inequality in health outcomes, cannot be ruled out in the first case, they are obviously a far less accurate representation of societal concerns than for the second case. Yet the Wagstaff approach does not distinguish clearly between situations of this type: there is a *de facto* anonymity assumption, *unless* it is overruled via weights a_i. The obvious conclusion is that an isoelastic SWF can only handle income related health inequality adequately by attaching greater weight to the health of poor individuals. But once again we are faced with the problem of exogeneity of the weighting scheme. If socio-economic inequality in health is a fundamental concern - as indeed it appears to be in all modern health systems - then this concern should be built into the evaluative structure. As Culyer (1989) has argued, in the general context of normative analysis of health and health care:

> "it seems altogether more preferable to adjust the scheme of things so as to incorporate such considerations fully rather than leaving them as a kind of *ad hoc* 'escape clause' ... in order to prevent theoretical emasculation (... inability to discuss in a consistent normative framework some matters that are of evident normative importance)." (p. 36).

Hence, the conclusion to be drawn is that in spite of the importance of Wagstaff's proposal, indeed its innovation, the particular form which his argument takes does not faithfully represent widely held views on the social welfare associated with health. Consequently, an alternative must be sought, and this must begin by identifying the fundamental arguments of the health SWF.

In every evaluative structure, some types of factual matters are taken to be of *intrinsic* importance. They reflect the basic ends of the system and therefore constitute

---

[7] This oversight is surprising given that Wagstaff has been at the forefront of opposition to the use of health domain inequality measures without reference to the distribution of economic position. In Wagstaff *et al* (1991a), for example, he criticizes the work of Le Grand (1987), Leclerc *et al* (1990) and others, arguing that it "fails to capture the socioeconomic dimension to inequality in health" (p. 546).

the informational basis of evaluative judgements. For instance, in the traditional 'welfarist' approach the informational foundation includes only unit comparable individual utilities. In the realm of health welfare analysis it seems appropriate to view society as having two distinct fundamental aims with regard to the health distribution. The first and most basic is that more health is preferred to less. It is denominated here as the *aversion to sickness principle* (ASP). At the community-wide level, it implies that the aim will be to maximize the total health or, equivalently, to achieve the highest mean level of health status that is feasible. A second widely held value is that population differences in health status should not be systematically related to socio-economic position. It is referred to, in this chapter, as the *aversion to inequity principle* (AIP). Such an objective underlies much of the concern over health inequities expressed in policy documents (see in particular, DHSS, 1980; WHO, 1987). It implies that a distribution of health which reveals a lesser degree of socio-economic inequality will always be preferred to one where there is more.[8]

These two principles express what are widely acknowledged to be the major objectives of health policy: efficiency and equity. Quite clearly, AIP captures the essence of equity concerns regarding the health distribution. ASP, on the other hand, represents a particular notion of efficiency, that of health maximization. This is the definition implicit in Wagstaff's (1991) approach and also that of many other authors who have considered the equity-efficiency trade-off in the domain of health (*eg.* Culyer, 1988). It is worth reminding that the view is not unproblematic. As it stands, it reveals no more than a societal preference for health output growth and hence Le Grand (1991a) has a point when he asserts that "it would seem preferable ... to discuss the issue of any trade-offs between growth and equity explicitly, rather than obscuring the issue by reference to efficiency" (p. 32). However, if the costs of health producing commodities are assumed constant across two societies, then their respective levels of health output can logically be seen as measures of relative productivity or efficiency in transforming commodity characteristics into health. This is the view adopted here.

---

[8] At the limit, the implication is that social welfare is maximized at the point where all income groups share the same level of health status. Whilst this might be viewed as a somewhat extreme equity objective, there is no doubt that modern societies reveal widespread concern for the *degree* of income related health inequality. Therefore, the AIP value would seem to be applicable.

If it is assumed that the health variable can be suitably standardized for relevant non-income heterogeneity among the population (*eg.* age and sex), then one reasonable specification of social welfare incorporating the ASP and AIP principles is the SWF

$$W(h) = V(\mu_h, I_h) \tag{8.5}$$

where μ_h represents the average health of the population (*ie.* $\mu_h = \Sigma h_i / n$) and I_h is an index of income related health inequity, bounded by zero (when all income groups share the same level of standardized health) and one (when the very richest enjoy all the standardized health in the community). V is increasing in μ_h and decreasing in I_h, thus representing a societal preference for more health and less income related inequality in its distribution. Of course, it could be argued that the informational basis of (8.5) is somewhat limited, since it is blind to concerns about the relative merit or age of the individuals whose health is being assessed. For instance, the view could be taken that personal health losses that are a direct result of unhealthy activities over which the individual exercises full control, would not lead to a fall in social welfare (*cf.* the Le Grand approach in Chapter 3). Likewise, health gains experienced by the old might be valued less than those of the young (or *vice-versa*), which would once again have implications for overall well-being. Unlike the ASP and AIP principles, however, there is not a great deal of explicit, documented public support for these values. Even if such opinions are held quite strongly by some individuals (*cf.* Williams, 1988a), the truth is that they seldom figure in policy statements. It seems best, therefore, to exclude them from a general evaluation of health well-being.[9]

In the income inequality literature, specifications of the form given in equation (8.5), but defined over incomes, have come to be described as *abbreviated social welfare functions*.[10] Particular attention has been given to including the Gini index of

---

[9] This is not to say that other types of analyses should overlook these or other criteria. Indeed, there is one other category of information that might be deemed of particular interest: that which is *instrumentally* influential, through its role in determining the intrinsically valued variables implied by ASP and AIP. Taking account of these indirect influences on health welfare calls for causal analysis of the instruments and their consequences.

[10] The term 'abbreviated social welfare function' seems to have been coined by Kondor (1975). The volume by Lambert (1989) devotes the whole of Chapter 5 to these type of specifications in the context of income distribution.

inequality as an argument in the SWF, perhaps as a result of Atkinson's (1970) observation that the judgements implied by this measure do not generally accord with widely held social values. The new work has shown that the Gini coefficient is an admissable measure of inequality in the context of social welfare if one assumes that individuals suffer deprivation by virtue of others having income resources higher than themselves.[11] The impetus for this line of enquiry was provided by Sen's (1974) axiomatic approach to measuring well-being. He assumed a society of n individuals, arranged in ascending order of their incomes: $y_1 \leq y_2 \leq \ldots \leq y_n$, denoted by the vector $\mathbf{y} = (y_1, y_2, \ldots, y_n)$. He then considered the following social welfare function:

$$W = \sum_{i=1}^{n} y_i \, \upsilon_i(y) \tag{8.6}$$

where $\upsilon_i(\mathbf{y})$ is the weight given to person i with income y_i. It should be emphasised that $\upsilon_i(\mathbf{y})$ is a function of the whole income distribution and not of y_i alone. It is this feature which permits the notion of relative deprivation to be incorporated.

Sen's axioms for the measurement of welfare were as follows. First of all, *relative equity*, implying that if person i is considered worse off than person j, the income of the ith person should have higher weight than that of j. Secondly, *monotonic welfare*, whereby a person with lower income is always considered worse off than one with higher income. Thirdly and crucially, *rank order*, implying that weights $\upsilon_i(\mathbf{y})$ should be a decreasing function of the rank of the income distribution. Finally, *normalization*, which states that if income is equally distributed then $W = \mu_Y$, the mean level of income. The only SWF satisfying these four axioms is given by:

$$W(S) = \mu_y(1 - Z_y) \tag{8.7}$$

where Z_y is the Gini coefficient. This provides an interesting computable index of social welfare, which can be interpreted as the equally distributed equivalent income measure for the Gini coefficient (*cf.* Bishop *et al*, 1990). Note that it is similar to the measure of health social welfare derived in (8.4) for the Atkinson-type index of health

---

[11] An inverse rationale can also be given in terms of altruistic concern for those that are less well-off (*eg.* Layard, 1980; Pyatt, 1980).

inequality, apparently simply substituting one inequality index for another. There is an important difference, however, in that the Sen rationale is based on explicit consideration of a *separate* equity concern: the sense of relative deprivation. This is captured by taking into account the *number* of persons who are richer than the representative individual.

In recent years, various rationalizations of (8.7) have been put forward in the income inequality literature. Yitzhaky (1979) and Hey and Lambert (1980) placed the ethical justification of the Gini coefficient squarely within the theory of relative deprivation. This theory offers an explanation of attitudes to social inequality, based on an individual's position relative to some 'reference group' in society (*cf.* Runciman, 1966). Lambert (1985) and Kakwani (1986) provided an important extension to the analysis, by allowing the weight given to inequality, as measured by the Gini, to vary. Useful empirical applications of the indices were also presented in Kakwani (1986) and in Bishop *et al* (1990, 1991). This work is of great interest to the purpose of measuring society's health welfare. Recall that in Chapter 4 the Gini coefficient was shown to be a special case of the more general concentration index. This suggests that C_h, the health concentration coefficient, might be used as a measure of income related inequity in the SWF depicted in (8.5). If we can further incorporate the reasonable notion that different persons may weight the relative importance of ASP and AIP differently, then the possibilities for devising a suitable measure of the general welfare associated with health are considerably enhanced. The following section develops this suggestion, showing that concentration measures have an ethically-sustainable role to play in comparisons of well-being associated with the functioning *health*.

8.3 A new approach to measuring health social welfare

There is an interesting parallel between the notion of relative deprivation underlying the rationalizations of the Gini index cited above and the way people view inequities in the distribution of *health functioning*. Consider Runciman's (1966) classic definition of the relative deprivation condition:

"A person is relatively deprived of X when (i) he does not have X, (ii) he sees

some other person or persons, which may include himself at some previous or expected time, as having X (whether or not this is or will be in fact the case), (iii) he wants X, and (iv) he sees it as feasible that he should have X" (Runciman, 1966, p. 10).

Is such an explanation valid in the domain of health? In fact, at least two arguments can be made in its favour, the most obvious being where X is the health of any one individual in society. That person would be relatively deprived of the desired functioning 'good health' if she were sick, others of the same age were healthy and it was physiologically feasible that her health could be brought up to the level of the reference group. From the social point of view, however, it seems more appropriate to drop the anonymity assumption and allow that X is *a named-functioning* (*ie.* the same functioning going to two different persons).[12] Suppose that the population is homogeneous apart from differences in health states and income levels and that these differences are described by the categories sick-healthy and poor-rich, respectively. Runciman's statement above would then run as follows: an individual suffers *income related health deprivation* when (i) he is sick and poor, (ii) there are other persons who are healthy and generally rich, (iii) being in good health is desirable, and (iv) it is physiologically feasible that he could be a healthy person, but given his income position he has less opportunity to achieve that desired state.

This description is intuitively similar to much of the concern over inequalities in health, where observers have placed particular emphasis on the health gap that divides rich and poor, and on the view that such differences would not arise were it not for reasons that are the direct or indirect result of persons' income position. There is a concern that the poor-sick person is somehow deprived in relation to those that are rich and healthy. Such *income related health deprivation* can be defined as the well-being foregone because of not possessing the level of health enjoyed by others, when that difference is considered by society to be avoidable were it not for differences in socio-economic status.

---

[12] Hahn (1971) first used this terminology when he defined the concept of 'named-goods' to describe the same commodity going to two different individuals.

There are, admittedly, some problems with the argument. Relative deprivation is usually conceived as relating to the means for achieving a desired state. Sen (1983), in particular, suggests that functioning should be viewed in an *absolute* sense whereas a *relative* approach is called for in the space of commodities and resources. Culyer (1990) goes further by arguing that relativity in health-related commodities hardly matters at all and that for health functioning any deprivation that may exist is clearly absolute. Nevertheless, health is itself instrumental to other desired functionings (*eg.* earning a living, enjoying leisure time with one's family) and hence there is a case for considering socio-economic related health deprivation as worthy of analysis. Indeed, both Sen (1983, p. 168) and Culyer (1990, p. 21) unequivocally state that their arguments do not imply that *inequality* in functioning space is an irrelevant issue for public policy. Furthermore, the fact that it is the *relative* differences between rich and poor which have inspired preoccupation over inequalities in health cannot be avoided. Without a reference group it is doubtful that similar attention would have been drawn to the absolute level of health of the poor. It seems useful, therefore, to further explore the notion of *income related health deprivation* in the context of the social welfare associated with health. In what follows, a rationale based on this idea is presented for an abbreviated SWF of the form $W = V(\mu_h, C_h)$ and consequently, for an index of well-being in the domain of health. The deduction draws on the rationalizations of the Gini coefficient cited earlier and particularly on work by Kakwani (1986, pp. 197-202).

Suppose that income Y is a continuous random variable with probability density function $f(Y)$ and distribution function $F(Y)$; the proportion of individuals having an income less than or equal to y_i being $F(y_i)$. Assume further that health H is also a random variable, related to Y by means of a continuous function $h_i = g(y_i)$, with $g'(y_i) > 0$ for all $y_i \geq 0$.

Individuals are represented by the well-being functional $U[h_i, F(Y)]$, imposed by an ethical decision maker. This states that each person's health welfare depends on own health and also, crucially, on the income distribution. Given that health is itself a function of income, individuals will necessarily compare health states when they compare their own income with that of others. In this context, income related health deprivation (HD) felt by a person with income y_i relative to another with income y_j

may be defined as follows:

$$HD\,(y_i, y_j) = \begin{cases} 0 & \text{if} \quad y_i \geq y_j \\ \kappa(h_j - h_i) & \text{if} \quad y_i < y_j \end{cases}$$

where κ is a constant measuring the degree to which the deprivation is felt. If κ is 0 health deprivation is judged as completely unimportant; if $\kappa = 1$, the level of income related health deprivation suffered by an individual is exactly equal to the difference between his health and that of the richer individual; and if $\kappa > 1$, individual deprivation is greater than the health gap. The last of these cases might be justified on the grounds that health is a basic commodity required for living a fulfilling life, and hence, a greater weight than the actual health difference should be attributed to the deprivation. For any given value of κ, whenever compared income levels are equal or lower, no deprivation (or loss of welfare) is assumed, so that the individual's well-being associated with health is given by own health, *ie.* $h_i = g(y_i)$. Howèver, if compared incomes are higher, implying a theoretical observation of a higher health state $h_j = g(y_j)$, then individual *i* suffers a loss of welfare, which is assumed to be proportional to the differences in the two persons' health states. In this case, health welfare is equal to $[h_i - \kappa\,(h_j - h_i)]$.

Now assume that individual health welfare is determined by a procedure where pairwise comparisons are made with other persons. Given that the probability of selecting an individual with income y_j from the population is $f\,(y_j)\,dy_j$, it follows that in all pairwise comparisons the expected health welfare (HW) enjoyed by a person with income y_i will be:

$$E[HW \mid h_i] = \int_0^{y_i} h_i\, f(y_j)dy_j + \int_{y_i}^{\infty} [h_i - \kappa(h_j - h_i)]\, f(y_j)dy_j \,. \qquad (8.8)$$

As a final assumption allow that the level of welfare enjoyed by society is measured by the average of individuals' well-being.

A number of points are worth noting about the assumptions laid out. First of all, in contrast to Wagstaff (1991), social well-being associated with health depends directly

on the income distribution. This is an important feature given that what generally motivates concern for inequalities in health are not the differences *per se* but their systematic relation to socio-economic status. Secondly, in the present approach, there is no requirement that the deprivation effect have empirical validity at the individual level. $U[h_i, F(Y)]$ is best seen as a well-being functional imposed by an equity conscious ethical decision-maker. In this case, the value judgement implied by the framework is that society has as its goal that people are not put in a position whereby they feel income related health deprivation. A third point is that the parameter κ can be seen as measuring the degree of society's aversion to income related health inequality, or equivalently, the extent to which it values AIP. As κ rises greater weight is attached to this principle and less to ASP.

Finally, it is acknowledged that the present analysis does involve an apparently restrictive assumption about possible combinations of health and income chosen for comparison by the reference individual. By defining health to be an induced random variable which is positively related to income, we are in effect ruling out two empirically observable situations, namely $\{h_i \geq h_j \text{ and } y_i < y_j\}$ and $\{h_i < h_j \text{ and } y_i \geq y_j\}$. Clearly, it could be argued that there are some grounds for relaxing the assumptions $h_i = g(y_i)$ with $g'(y_i) > 0$. From the point of view of society, however, it is less important that exceptions can be found. What is at stake is whether this assumption *generally* reflects underlying views and factual evidence. In fact, the hypothesis that health is a positive function of income clearly underlies the inequality in health debate and it is one for which much empirical evidence can be advanced. For instance, in the volume by Van Doorslaer *et al* (1993) all the country reports show a positive (negative) relation between health (morbidity) and income no matter how the health variable is defined. The empirical analysis of chapters 5 and 6 also points in the same direction.

The assumption could be relaxed by defining $h_i = g(y_i) + \varepsilon_i$, where ε_i is a stochastic term with zero mean, constant variance and non autocorrelation. In this case HD might be interpreted as an expected measure, *ex-ante* to the realization of the health state. It is then possible to accomodate relevant *ex-post* cases such as $\{h_i \geq h_j$ and $y_i < y_j\}$. Since individual health welfare is determined by pairwise comparisons, this stochastic definition leads to the same result reached below (by virtue of the expected

value of ε_i being zero). Therefore, I use the deterministic specification for ease of exposition and because it reflects underlying views and general factual evidence.

Recall from Chapter 4 that the health concentration curve expresses the relationship between $F_1[g(y_i)]$ and $F(y_i)$. If individuals are arranged in ascending order of their incomes, $F_1[g(y_i)] = F_1(h_i)$, the proportion of health enjoyed by persons with incomes equal to or below y_i, is given by:

$$F_1(h_i) = \frac{1}{E[h_i]} \int_0^{y_i} H\, f(Y)\, dY \qquad (8.9)$$

where $E[h_i] = \mu_h$ = average level of health status. Substituting (8.9) into (8.8) gives the following health welfare curve (detailed mathematical steps are shown in Appendix 6):

$$W(h_i) = h_i - \kappa\mu_h[1 - F_1(h_i)] + \kappa h_i[1 - F(y_i)] \qquad (8.10)$$

where W (h_i) is the expected welfare enjoyed by an individual with income y_i and $F(y_i) = \int_0^{yi} f(y_j)dy_j$ is the probability distribution function of income.

Suppose now that income y_i corresponds to the 100pth percentile, then the health concentration curve is represented by a function $G_h(p) = F_1(h_i)$, where $p = F(y_i)$ and $0 \leq p \leq 1$. $G_h(p)$, which is obtained by eliminating y_i and h_i from the distribution function of income and the first moment distribution function, $F_1(h_i)$, is interpreted as the fraction of total health received by the lowest pth fraction of individuals. It satisfies the following conditions (Kakwani, 1980: pp. 157-8):

$$(a) \quad \text{if } p = 0, \qquad G_h(p) = 0$$

$$(b) \quad \text{if } p = 1, \qquad G_h(p) = 1$$

$$(c) \quad G_h'(p) = \frac{dF_1(h_i)}{dF(y_i)} = \frac{h_i}{\mu_h} \geq 0 \qquad (8.11)$$

$$G_h''(p) = \frac{d^2F_1[h_i]}{dF^2(y_i)} = \frac{h_i'}{\mu_h} \cdot \frac{1}{f(y_i)} > 0.$$

Given that $h_i' = g'(y_i)$ is positive, it follows that $F_1(h_i) < F(y_i)$ and that the concentration curve will fall below the egalitarian line. Using these relationships in (8.10) yields the

average health well-being enjoyed by an individual at the 100pth percentile:

$$W(h_i) = \phi(p) = \mu_h[\, G_h'(p)\,(1+\kappa-\kappa p) - \kappa + \kappa G_h(p)\,] \,. \qquad (8.12)$$

Given that social welfare has been defined as the average of individual welfare, the area under (8.12) can be interpreted as the average health well-being enjoyed by society:

$$\int_0^1 \phi(p)\,dp = \mu_h \int_0^1 [\, G_h'(p)\,(1+\kappa-\kappa p) - \kappa + \kappa G_h(p)\,]\,dp \qquad (8.13)$$

which on evaluating the integrals leads to:

$$W(h_i) = \mu_h [\, 1 - \kappa (1 - 2\int_0^1 G_h(p)\,dp)\,] \,. \qquad (8.14)$$

Since the health concentration index, C_h, is defined as one minus twice the area under the health concentration curve, $\int_0^1 G_h(p)dp$, this equation rationalizes a family of social welfare functions of the type:

$$W_\kappa = V(\mu_h, C_h) = \mu_h(1 - \kappa C_h) \qquad \kappa \geq 0 \,. \qquad (8.15)$$

The indices W_κ, implied by (8.15), provide the basis for empirical measurement of the social well-being associated with health. It should be noted that W_κ allows the weight to be given to the health concentration index to vary, and hence the terms of the trade-off between income related health inequality (equity) and health output (efficiency). Whenever $\kappa = 0$, implying that society does not value AIP at all (*ie.* does not recognize deprivation effects), W_κ is given simply by μ_h, a measure of well-being completely insensitive to changes in the social distribution of health. However, if AIP is valued (*ie.* $\kappa > 0$), the inequity-efficiency trade-off is given by:[13]

$$\frac{C_h}{\mu_h}\left.\frac{d\mu_h}{dC_h}\right|_V = \frac{\kappa C_h}{1-\kappa C_h} \,. \qquad (8.16)$$

This implies that a reduction of $\{\kappa C_h / 1\text{-}\kappa C_h\}$ percentage points in the average level of

---

[13] Equation (8.16) expresses the elasticity of substitution between the arguments of the SWF. It is obtained by totally differentiating equation (8.15) [See Lambert (1989), Chapter 5].

health status will be accepted by society, in order to achieve a 1% reduction in inequity. Clearly, the more unequal the existing socio-economic distribution of health, and the stronger the deprivation effect that individuals are assumed to feel, the greater will be the accepted reduction in mean health status.

8.4 An empirical application

The family of welfare indices developed above has a number of potential applications. The most obvious are in comparing the performance of health systems and in examining the intertemporal health achievements of a particular country. The data requirements for such analyses are relatively modest. One essentially needs access to: (i) a suitable *per capita* health indicator for each country or time period; (ii) an estimate of the respective health concentration coefficients, obtained by crossing individual or grouped information on health, income and demographic characteristics; and (iii) an assessment of society's relative valuation of the ASP and AIP principles. Items (i) and (ii) are readily available in a number of countries. For instance, the volume by Van Doorslaer *et al* (1993) presents data on the prevalence and income distribution of chronic illness in eight european countries that might easily be used for a cross-national comparison of the social welfare associated with chronic conditions. Similarly, Propper and Upward (1992) present average levels of illness and income-indexed concentration indices for a number of morbidity indicators available in various samples of the UK General Household Survey in the period 1974-87. Such data could be used to assess the time-trend of health social welfare in the UK.[14] Item (iii) is more problematical. Ideally, experimental methods should be used to try to elicit values of the inequality aversion parameter κ. In the short term, in the absence of firm guidance from policy makers or the public at large, one might opt for a strategy of calculating health welfare

[14] It might be argued that because of cultural and temporal variability in survey reporting, cross-country or longitudinal comparisons of morbidity levels will bias the social welfare estimates. However, such differences are not necessarily greater than those between persons from different income groups. The referenced studies implicitly assume that the decision to report illness is unaffected by cultural norms related to income position. As a number of sociological studies have shown, this is not necessarily true (d'Houtaud and Field, 1984; Blaxter, 1989). The obvious point to remember is that social welfare appraisals would benefit from more refined health status measures; but so too would estimates of the level of health inequality.

levels for each of several different values of κ.

The W_κ indices may also be used to compare levels of health social welfare across geographical areas of the same country. Such an application is illustrated in this section, with numerical estimates being provided for Portugal's five continental regions: the North, Centre, Lisbon and Tagus Valley, Alentejo and the Algarve. The data are drawn from the 1987 National Health Survey (INS) described in Chapter 5.[15] Table A7.1 (Appendix 7) shows various background statistics computed from the survey. The main points of interest for the analysis that follows are: (i) distinct demographic compositions of the regional populations, with average age increasing and average family size decreasing as one moves from northern to southern areas; (ii) higher than average health care consumption in the Lisbon and Tagus Valley region; and (iii) low income levels coupled with high levels of inequality to the north in contrast to the richer but less unequal regions of the south. These patterns are in agreement with macro-level observations (see, *eg.*, Pereira *et al*, 1987).

Given that the INS only collects morbidity data, the analysis actually measures the level of social 'illfare' (*ie.* the converse of social welfare). Three morbidity indicators were used:

(i) *limiting chronic illness*, a dummy variable representing persons who declared 'disability' (bed days, off work days or restricted activity) in a two week reference period, with the incapacity resulting from a medical condition present for at least three months;

(ii) *disability and illness*, a further dummy variable representing persons who declared 'disability' and 'illness' (no disability but felt unwell) in a two week r.p.; and

(iii) *number of sickdays*, computed by aggregating bed days, off work days and restricted activity days in a two week r.p., with bed days being subtracted for any individual that simultaneously reports off work days.

Since there is some overlap in the informational content of the measures, a degree of caution is advisable when considering congruence and dissimilarity in the results

---

[15] Recall that the sampling design yields representative estimates at the regional level. Geographical delimitations may be found in the map shown in Appendix 2.

provided by each one. Nevertheless, the measures provide some indication of particular aspects of an individual's health condition: respectively, the susceptibility to functional limitation due to longstanding conditions, general lack of health whether of a limiting or subjective nature and the degree of illness severity as measured by the time duration of incapacity.

If valid comparisons of health social welfare are to be made, it is necessary to remove collinearity with demographic factors. This was accomplished by means of direct standardization (see equation 5.3). Average levels of illness in the regions were age-sex standardized by the national population. The concentration indices were calculated (through the linear approximation method) using income decile groups of the population in each region and standardizing by the age-sex distribution of the national population. Thus, the regional estimates are made on the assumption that each region has the demographic composition of the country as a whole.

The income definition is the same used in Chapter 5: namely, individual equivalent disposable income using the OECD scale. Estimates of health social welfare were computed according to four different values of the inequality aversion parameter κ, in the range κ=0 to κ=3.

The results are presented in Figure 8.2, which shows regional variation in health social illfare relative to the national average (= 100).[16] The three morbidity variables show roughly similar patterns. The North reveals stable levels of social illfare, close to the national average, irrespective of the weight attached to the ASP and AIP principles. The Centre has high levels of social illfare which tend to be judged slightly less disadvantageous the greater the concern for income related health deprivation. The Lisbon and Alentejo regions perform relatively well at all values of κ, with social illfare generally below the national standard. For the general disability and illness variable the relative advantage of the Lisbon region declines as κ is increased, a fact which seems largelly attributable to high levels of inequality in subjective illness. The most striking results are, however, for the Algarve region which generally presents the worse health

[16] The actual computed values, measured on the morbidity scale, are shown in Table A7.2.

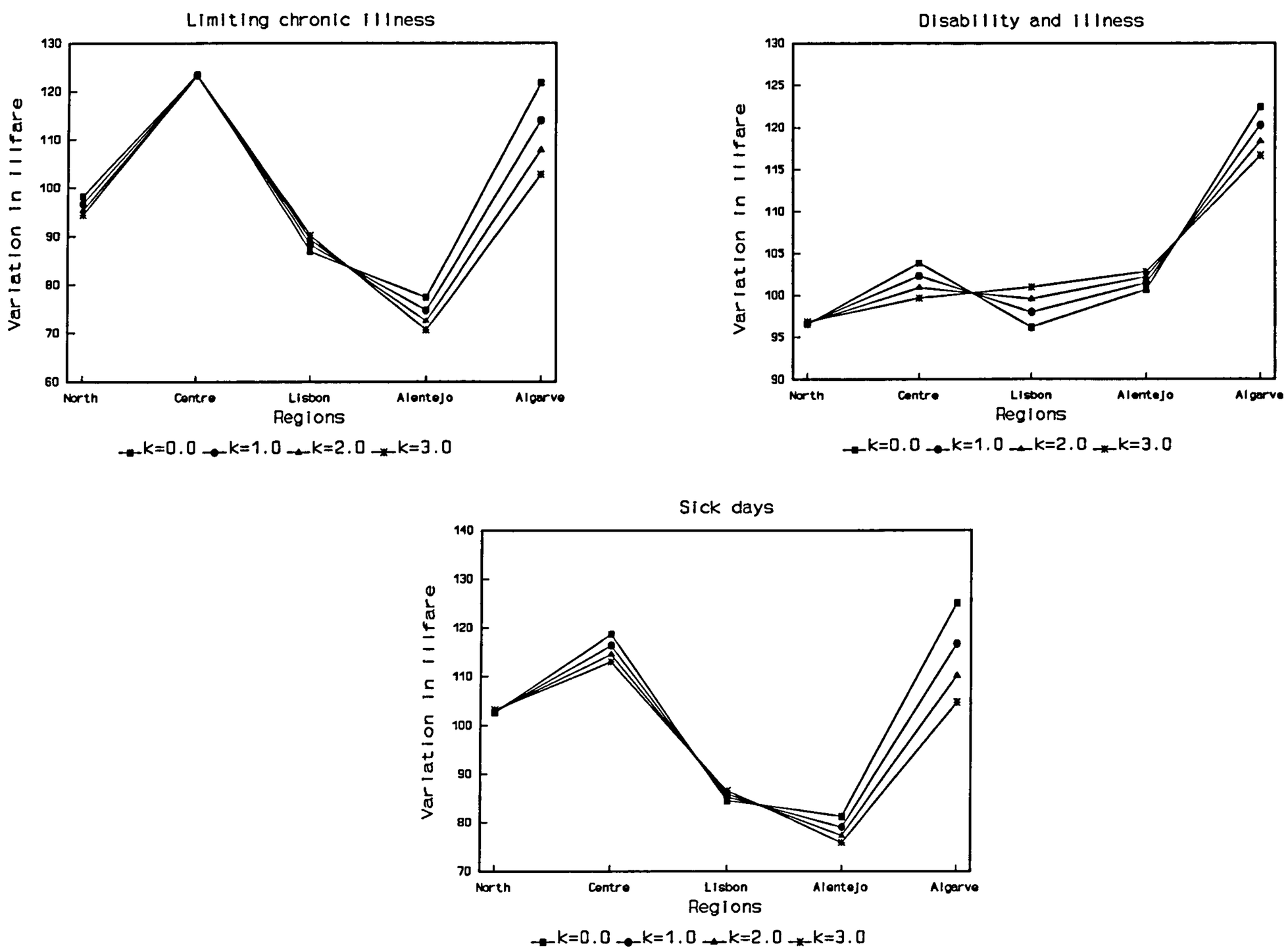


Figure 8.2: Variation in the social illfare associated with health based on three morbidity indicators. Age and sex standardized

illfare levels according to the deprivation indifferent index $\kappa = 0$. Once the value of κ is increased, the relative position of the Algarve population improves significantly. Thus, for the sick days variable at $\kappa = 3$, health social illfare is only 5 per cent higher than the national average against 25 percent for the deprivation indifferent index. In terms of relative positions, a deprivation conscious societal appraisal ($\kappa = 3$), will judge the performance of the North and Algarve regions as roughly equal, which is in marked contrast to the inequality indifferent evaluation. The reason that, as κ is increased, the health of the Algarve population is judged less disadvantageous is, of course, that income related health inequality is lowest in that region (see Table A7.3 in the Appendix). Similarly, the relative decline in the performance of the Lisbon region is due its having the most extreme socio-economic health differences.

The estimates show that when there is a concern for the economic distribution of health, social judgements on the relative position of different populations can change. This result has implications for regional resource allocations. For instance, suppose that health care resources were apportioned on a *per capita* basis and that in future the policy maker wished to compensate regions with lower levels of health by attributing them higher levels of health care resources. Given the present findings, he would be driven to substantially increase the share of public funds to the Algarve region. However, if health inequity were a fundamental social concern (as it appears to be in most countries), then the allocation mechanism should logically be less generous to the Algarve.

Clearly, we have a long way to go before such implications can be put into practice. There is obvious scope for improving the estimates by using more refined indicators of population health. Once these are available the informational significance of health social welfare measurements will be greatly increased. There is also a need to clarify societal preferences with regard to the trade-off between health output and inequity in its distribution (*ie.* the appropriate value of κ). The empirical illustration nevertheless shows that conventional health output evaluations can be informatively weighted by ethically-founded distributional details.

8.5 The distribution of health social welfare by income groups

Given that the focus of the thesis is on equity, it is important to know how health social welfare is distributed across groups of the population with differential command over resources. This section extends the analysis developed in 8.3 so as to permit measurement of this issue. In effect, it shows how a society that values both the AIP and ASP principles, will gauge the health situation of different income groups.

Suppose that a society is most concerned about the health of the poorest $100p_1$ percent of the population, perhaps on the basis that these persons fall below some accepted poverty line and are judged as having relatively less opportunity to improve their health status than their economically better-off peers. Using equation (8.12), the average level of health welfare of this poorest group is given by

$$W(p_1) = \frac{1}{p_1} \int_0^{p_1} \mu_h [G_h'(p)\,(1+\kappa-\kappa p) - \kappa + \kappa G_h(p)]\,dp \qquad (8.17)$$

which on evaluating the integrals leads to[17]

$$W(p_1) = \frac{\mu_h}{p_1} [G_h(p_1) + \kappa G_h(p_1) - \kappa p_1 G_h(p_1) - \kappa p_1 + 2\kappa \int_0^{p_1} G_h(p)\,dp] \,. \qquad (8.18)$$

Since $G_h(p_1)$ is the proportion of total health held by the poorest p_1 decimal fraction of the population, it follows that the mean level of health status of this group will be:

$$\mu_{hp_1} = \frac{\mu_h G_h(p_1)}{p_1} \qquad (8.19)$$

and that the health concentration coefficient, in the population formed by these individuals, can be written as

---

[17] See Appendix 6 for detailed mathematical steps.

$$C_{hp_1} = 1 - \frac{2}{p_1 G_h(p_1)} \int_0^{p_1} G_h(p)\, dp \ . \tag{8.20}$$

Using both these results, (8.18) simplifies straightforwardly to

$$W(p_1) = \mu_{hp_1} - \kappa[(\mu_h - \mu_{hp_1}) + p_1 \mu_{hp_1} C_{hp_1}] \ . \tag{8.21}$$

This equation is easily computed from survey data as a measure of the average level of health social welfare enjoyed by the poorest group in society. Furthermore, the result generalizes to other percentile groups of the population ordered by income. Note that pW(p) is the area under the welfare curve up to the 100pth percentile. It follows that the area under this curve between the $100p_1$ to $100p_2$ percentiles will be

$$p_2 W(p_2) - p_1 W(p_1) \ . \tag{8.22}$$

Therefore, the average health social welfare of individuals whose income is situated between the $100p_1$ and $100p_2$ percentiles is given by

$$W(p_1, p_2) = \frac{1}{(p_2 - p_1)} [p_2 W(p_2) - p_1 W(p_1)] \ . \tag{8.23}$$

This equation permits the computation of health social welfare enjoyed by any income group. For instance, the average welfare of the second income decile is obtained when $p_1 = 0.1$ and $p_2 = 0.2$; for the third decile, when $p_1 = 0.2$ and $p_2 = 0.3$; and so on. Given that health surveys generally record income variables in discrete intervals, this result has some practical relevance.

The analysis set out here plainly shows that measuring health welfare by income group is distinct from measuring income-related inequality in the distribution of health. For example, suppose that we adopted a shares approach to equity measurement and concentrated on the poorest group in society. Under the inequality method, standard concentration analysis would quantify the ratio $p_1\mu_{hp1} / p\mu_h$, thereby establishing the share of total health held by the poor. The corresponding share of total welfare, on the other hand, is given by $p_1W(p_1) / pW(p)$. Equations (8.21) shows that the two

measurements will only coincide if society does not value socially related health deprivation (*ie.* when $\kappa = 0$). If a concern exists, however, the observed health of the poor has to be corrected by two κ-weighted factors in order to arrive at their average level of welfare. The first of these is the absolute difference between the average health status of the poor and that of the population as a whole. A *ceteris paribus* increase (decrease) in this health gap will naturally decrease (increase) the welfare of the poorest group. The other factor which affects health welfare among the class of poor persons is shown by the third term in equation (8.21). It implies, quite logically, that *ceteris paribus* increases in the extent of health inequality *within* the group will also decrease their level of welfare.

8.6 A further empirical illustration

This section provides an empirical illustration of the measures developed in section 8.5. The objective of the analysis is to establish levels of health social welfare for income ranked percentiles of the portuguese population. This is an important issue in the light of the normative meaning given to the term inequity throughout the thesis. We have been essentially concerned with measuring inequality in *achievement* and *advantage* attributes that is systematically associated with inequality in the extent of choice over commodity bundles. Measurement of the distribution of social welfare by income groups may be seen to provide a more accurate portrayal of economic differences in health *achievement*, than the more common endeavour of measuring income related health inequality. Like the overall social welfare measure W_κ, the disaggregated income group indices provide a balance between aggregative and distributional policy goals.

The data base and variable definitions are the same used in the earlier illustration. Attention is focused on measuring the distribution of health social welfare by quintiles of equivalent income. In order to remove the confounding influence of intervening demographic variables the distributions were standardized for age and sex by the direct method. The procedure is basically the same used in Chapter 5, with the distinction that it is carried out on income ranked subsamples of the INS data rather than on the sample

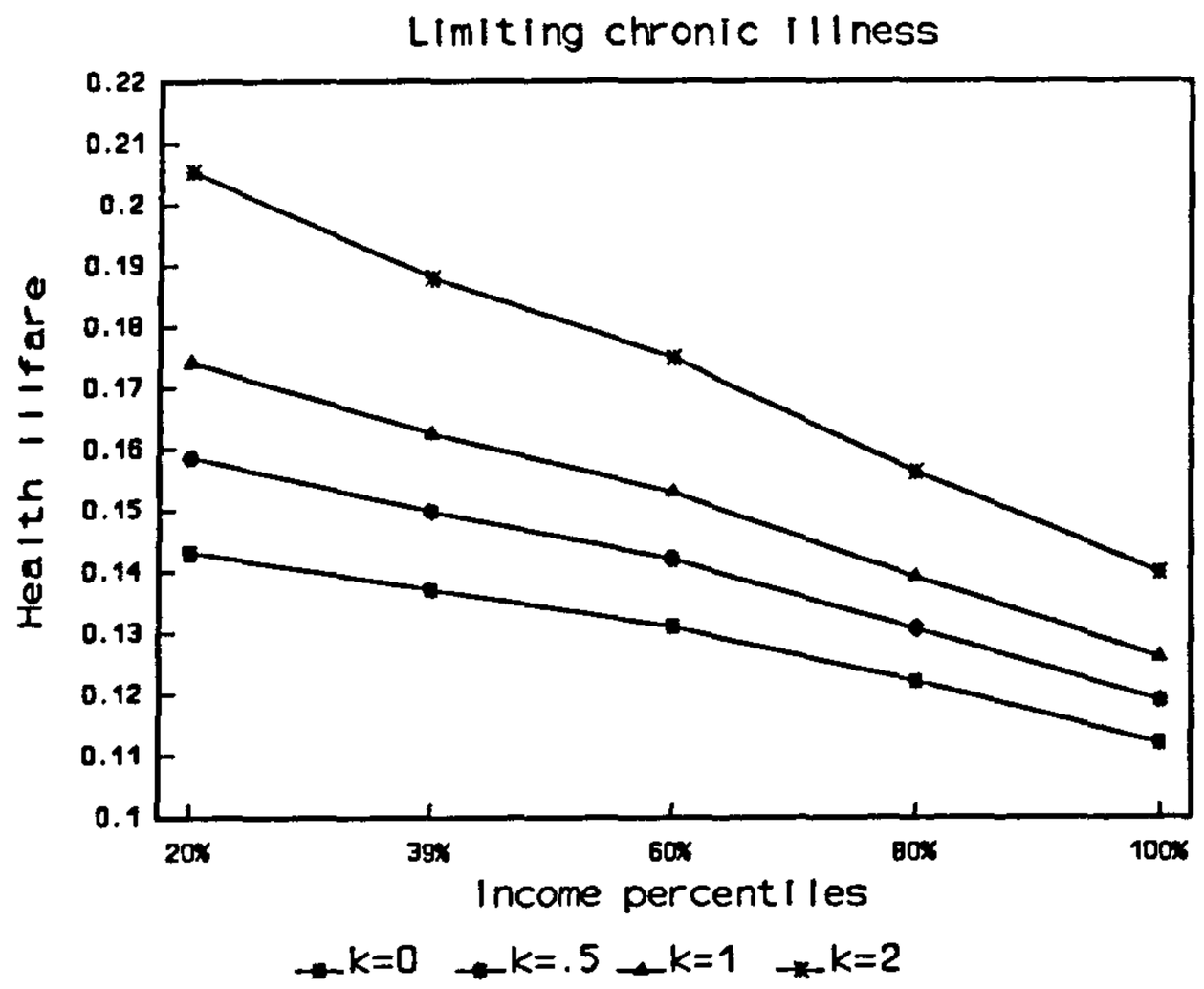
as a whole. Thus, for the poorest 20 per cent of the population, standardized average levels of morbidity for income segments of that population were calculated on the assumption that each has the age-sex distribution of the group as a whole. Likewise for the poorest 40, 60, 80 and 100 per cent of the population. The segments within the percentile groups were given by the income vingtile divisions for the sample as a whole. The resulting information allows computation of (i) standardized average levels of illness for each of the five percentile groups, and (ii) the respective illness concentration indices (estimated using the linear approximation method). These are the key inputs to estimation of the health social welfare indices implied by equations (8.21) and (8.23). The empirical analysis uses measures defined for each of four values of the inequality aversion parameter, in the range $\kappa = 0$ to $\kappa = 2$.

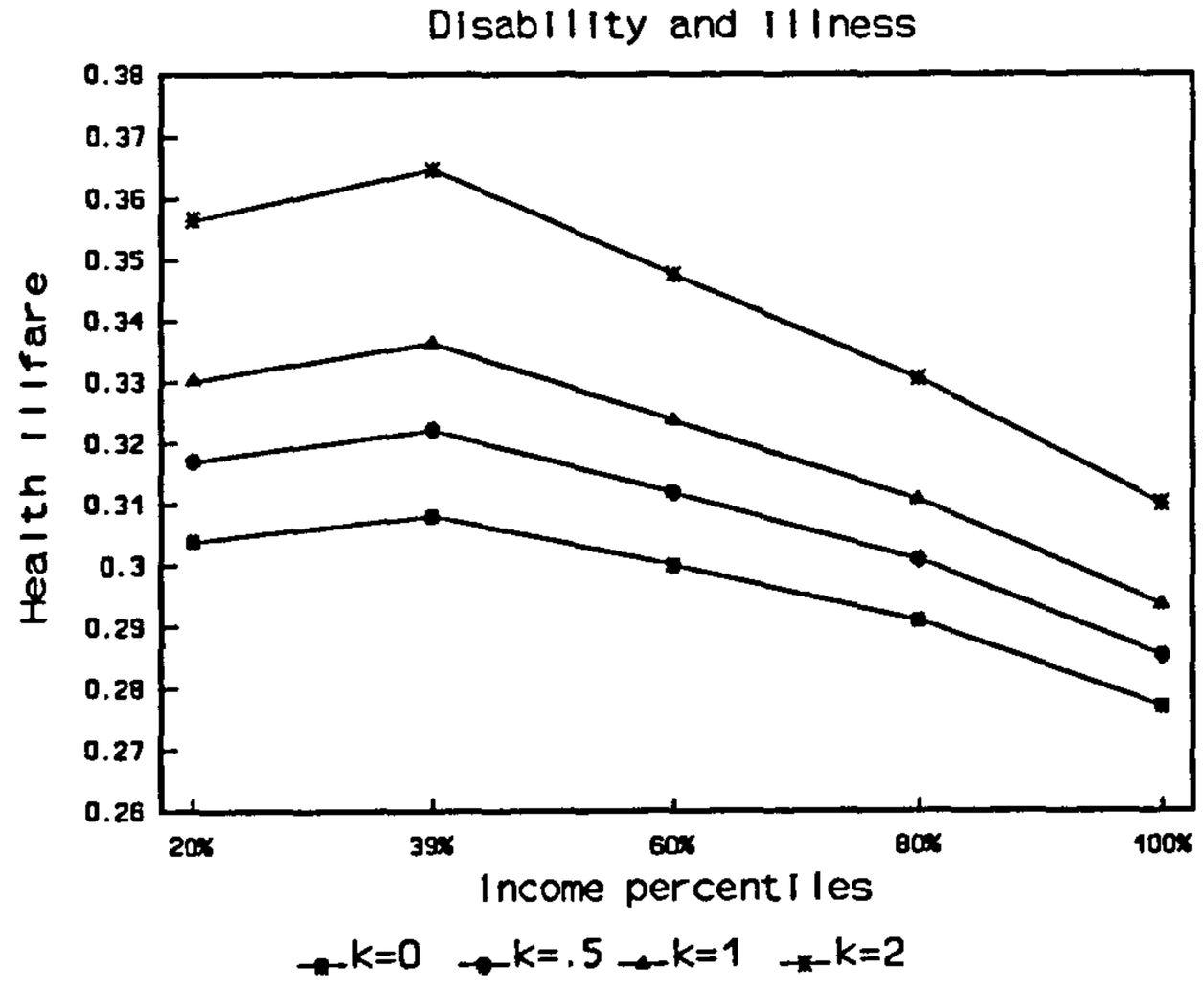
The results are presented in Figures 8.3 and 8.4.[18] The first of these shows average levels of social illfare associated with the morbidity distribution for cumulative percentiles of the income ranked population. In other words, the graphs show the social illfare curves implied by equation (8.21). Figure 8.4 is based on estimation of equation (8.23). It shows the variation in average levels of health social illfare enjoyed by equivalent income quintiles, relative to the population average.

The evidence is easily summarized. With one exception, health social welfare increases (illfare decreases) with income for all three health indicators and at all values of κ. The exception is for the disability and illness indicator where the second quintile has a lower level of welfare than the first. Generally, it may be concluded that the social welfare associated with the health of richer persons is higher than for the poorest, no matter how much normative weight is given income related health inequality. Unsurprisingly the welfare difference between poor and rich is greater the higher the value of κ assumed. For example, for the sickdays variable, at $\kappa = 0$ the health social illfare of the poorest quintile is twice that of the richest. At $\kappa = 2$ it is almost three times as great. If society values both the ASP and AIP principle, these results indicate that distributional comparisons of health status levels tend to underestimate the true degree of inequality.

---

[18] The computed numerical values are shown respectively in Tables A7.4 and A7.5.





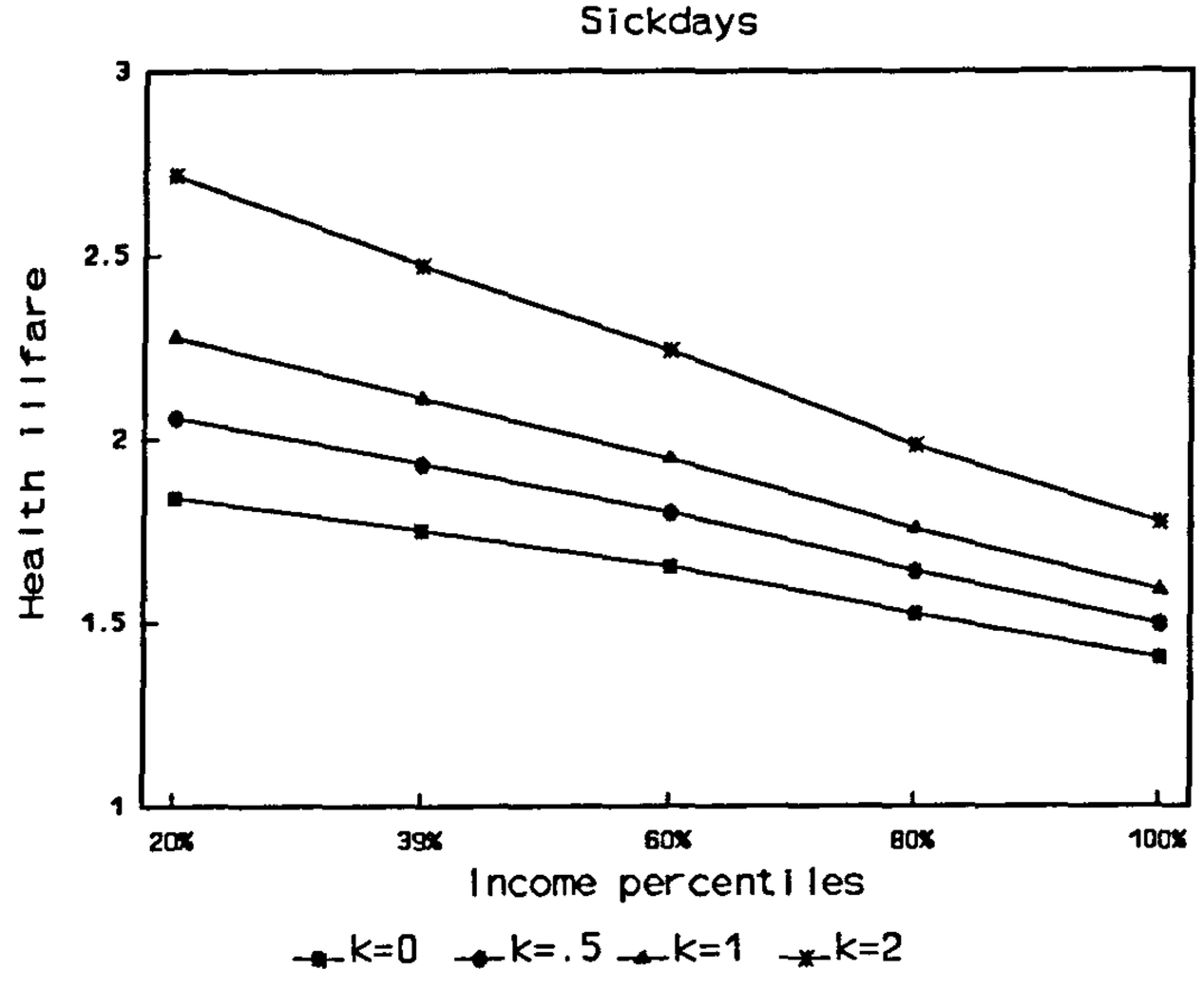


Figure 8.3: Social illfare curves based on three morbidity indicators. Age and sex-standardized. Values of $\kappa = 0, 0.5, 1$ and 2

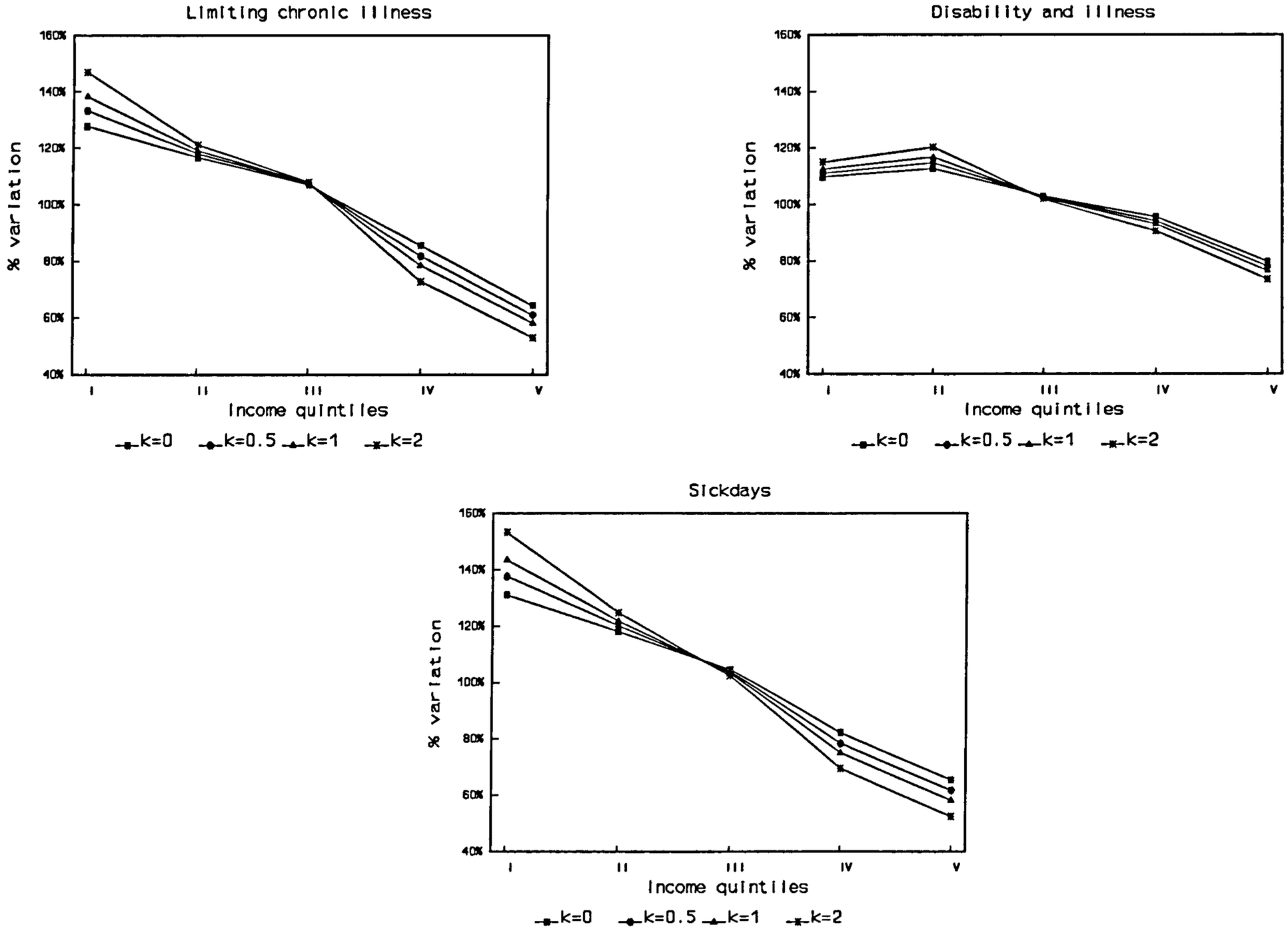


Figure 8.4: Variation in the social illfare associated with health by income quintiles. Age and sex-standardized.

8.7 Conclusions

This chapter was motivated by an implicit requirement of the capabilities approach to health-equity analysis: that of making an overall assessment of society's *achievement.* This problem is equivalent to measuring the social well-being associated with the health distribution, an area of analysis that has received insufficient attention in the health economics literature. It is possible to shed light on this question by drawing on developments in other areas of economic science, particularly work that investigates the problem of income distribution in the context of social welfare. It was argued that the approach adopted by Wagstaff (1991), which draws on Atkinson's (1970) well-known contribution, suffers from a number of limitations. An alternative approach, also based on work in the income inequality literature, was presented. It has three crucial distinctions in relation to Wagstaff's method: (i) it assumes that health social welfare evaluation involves two separate efficiency and equity considerations, and that people may attach different weight to these; (ii) the equity concern is motivated by the observed systematic association between health and income inequalities and not by health differences *per se*; and (iii) the interdependence of individual well-being is explicitly recognized and explained through the notion of income related health deprivation. All of these seem reasonable if a coherent ethically-founded evaluation of health social welfare is sought.

The principal result of the Chapter is a rationalization of a family of welfare indices that allow for a trade-off between equity and efficiency considerations (equation 8.15). This result, which is based on a similar rationale for the Gini coefficient provided by Kakwani (1986), is important in a number of respects. First of all, the indices provide assessments of *achievement* on the health scale, that take due account of the size of health output *and* the way this output is distributed among different income groups. Consequently, they provide a means of integrating equity and efficiency analysis through a single summary measure. Added to the fact that the indices are easily interpreted, this should mean that they have a potentially promising role to play in policy evaluation. Secondly, it has been shown that concentration coefficients, recently the most important indices used by health economists to study inequity, are admissable summary measures of inequality in the context of social welfare. Finally,

it should be noted that the rationale provided could be generalized to encompass other variables whose systematic association with income levels is a source of social concern. Therefore, the present analysis also has implications beyond the health economics sphere. However, it does not provide a complete answer to the underlying problem of the policy-maker: that of trading the nation's health against other social goals. The SWF on which the analysis is based merely represents that part of the overall well-being which is associated with the population's health.

The empirical illustrations show that the proposed indices can be used to examine questions of evident normative importance. The comparison of health social welfare across geographical entities shows that the Centre and Algarve regions have consistently worse levels than other parts of Portugal. However, if there is a concern for the economic distribution of health, the position of the Algarve becomes less disadvantageous. The analysis may easily be applied to cross-country or longitudinal comparisons. The former raises the problem of cultural variability in survey reporting of health status levels (*eg.* for a given objective level of illness, the populations of southern europe might reveal a higher propensity to declare than their more stoic neighbours to the north). However, such differences, if they exist, are not necessarily any greater than those between persons from different socio-economic groups. Comparisons between rich and poor, or individuals belonging to particular occupational groups, are after all the basis of most health domain inequity evaluations. What the approach of the present chapter does imply is that more refined measures of health status are required. Once these are available, the type of comparisons shown in section 8.6 may come to be seen as a more useful exercise than simply measuring relative inequality in health status levels. Future analyses should consider how the overall index W_κ may be further decomposed to show levels of social welfare associated with the health of specific population groups (*eg.* those defined by age, sex, education, employment status, and so on).

Chapter 9

Measurement of Inequity in the Delivery of Health Care

"As with any pioneering effort, observers can find issues that are left hanging or that are subject to challenge. Indeed, the essence of seminal contributions is that they spawn additional research to correct, modify and extend the initial effort"
Aaron (1992, p. 467)

9.1 Introduction

This chapter discusses indices of concentration aimed at measuring vertical and horizontal inequity in the delivery of health care. In common with Chapter 8, it seeks to provide conceptual groundwork to future empirical research, whether in Portugal or in other countries.

Recall from the discussion in Chapter 3, that the normative basis for examining inequity in the delivery of care derives from a fundamental concern for the distribution of health capability. Health care is important because of its instrumentality as an agent for the improvement of health (or the minimization of ill-health). Hence, inequity in the distribution of health care needs to be analysed because of its contribution to inequity in the distribution of health. This line of reasoning suggests that the focus of (inequity in delivery) empirical analysis ought properly to be on the care *received* by individuals; and not simply on their *access* to the health care system. As a corollary, portuguese policy statements (see Chapter 2) are given the following interpretation: an equitable health care delivery system is that in which the care received by individuals is distributed according to need, rather than economic and social circumstances.

An innovative approach to measuring economic related violations of distribution of health care according to need is that of the ECuity group [see, *eg.*, Wagstaff *et al*

(1989), Wagstaff *et al* (1991a,b) and van Doorslaer *et al* (1992)]. The present analysis aims to review, appraise and extend the concentration curve-based tools used by these authors when examining inequity in the delivery of care. An important feature of the discussion is the use of the continuous random variable framework employed in preceding chapters. The method allows general properties of the various ECuity indices to be conveniently determined and helps to distinguish conceptual from data-specific features, something which is not always clear in the descriptions provided by the ECuity researchers. Indeed, it serves to highlight an inconsistency in Wagstaff *et al's* analysis and leads subsequently to its reformulation. It is argued that the ECuity delivery-side analysis can only be accepted as measuring horizontal inequity under quite restrictive data assumptions, and that it is best reinterpreted as measurement of vertical inequity. Accordingly, a new index of horizontal inequity is developed, which brings the ECuity analysis more in line with methods currently employed in other areas of economics. The new measure, which is based on the correspondence of rank positions in the utilization and health status distributions, is more general than those previously available to health economists.

It should be noted that the analysis does not deal explicitly with empirical specification. The issue is only raised as a means of highlighting specific conceptual attributes of the measures under discussion. For reasons largely related to data availability, the ECuity work has proxied 'health care utilization' by the value of health care expenditures and 'need' by initial or presenting health status. Proxying 'need' by health status is a common approach in the health economics and health services research literature (*eg.* Le Grand, 1978; Aday *et al*, 1980; Puffer, 1986). The use of expenditures to measure 'utilization' is less common but nevertheless necessary if one wishes to aggregate heterogeneous forms of health care consumption. There is, however, some controversy surrounding the validity of these simplifications [see, *eg.*, Culyer (1991)]. It is assumed for present purposes that they are in fact useful for studying inequity in the delivery of care. An implication for the discussion which follows is that the term 'expenditure' is used as a synonym for 'utilization' or 'treatment'. Similarly, 'illness', 'ill-health' and 'morbidity' are used as synonyms for 'need'.

The chapter is structured as follows. Section 9.2 describes the ECuity approach

to measuring inequity in the delivery of health care. It reinterprets this analysis as measurement of vertical inequity and proposes a number of new indices based on the analysis undertaken in Chapter 4 (*eg.* a Suits index of inequity in health care and *generalized* versions of the ECuity indices). In Section 9.3, an alternative measure of horizontal inequity in delivery is developed. The final section summarizes and concludes the analysis.

9.2 The ECuity approach to inequity in health care delivery

The ECuity approach to measurement of inequity in the delivery of care is conceptually similar to that of the finance side analysis (see Chapter 4). Distribution of health care according to need (health status) can also be interpreted in terms of both *horizontal equity* (the requirement that persons with equal presenting health situations be treated equally) and *vertical equity* (the requirement that persons with unequal health status be treated in an appropriately dissimilar way). The ECuity research ostensibly focuses on the former of these, searching for violations of the horizontal principle that are related to income [Wagstaff *et al* (1991b, pp.169-70); van Doorslaer *et al* (1992, p. 390)]. It will be argued that the concentration-based indices suggested by these authors can only be interpreted as measures of horizontal inequity under quite restrictive data-related assumptions. In general, they actually measure departures from proportionality of the utilization distribution *vis-à-vis* the distribution of 'need' (when both are indexed to income levels), and are therefore, best seen as indices of vertical inequity.

The first ECuity index was initially suggested by Wagstaff *et al* (1989) as an extension to Le Grand's (1978) pioneering approach to measurement of equity in delivery. Le Grand used 1972 GHS data to compare the shares of NHS medical expenditures received by socioeconomic groups (SEG's), with their respective shares of ill-health. Two results presented in the study became the focus of much public debate on the equity of the NHS. They were, (i) that the top two SEG's received 40% more health care expenditure per person ill than the bottom two (the use-need ratio); and (ii) that the top two SEG's received 16.8% of NHS expenditure but contained only 13.9% of persons reporting illness, whereas the bottom two SEG's received 27.3% of

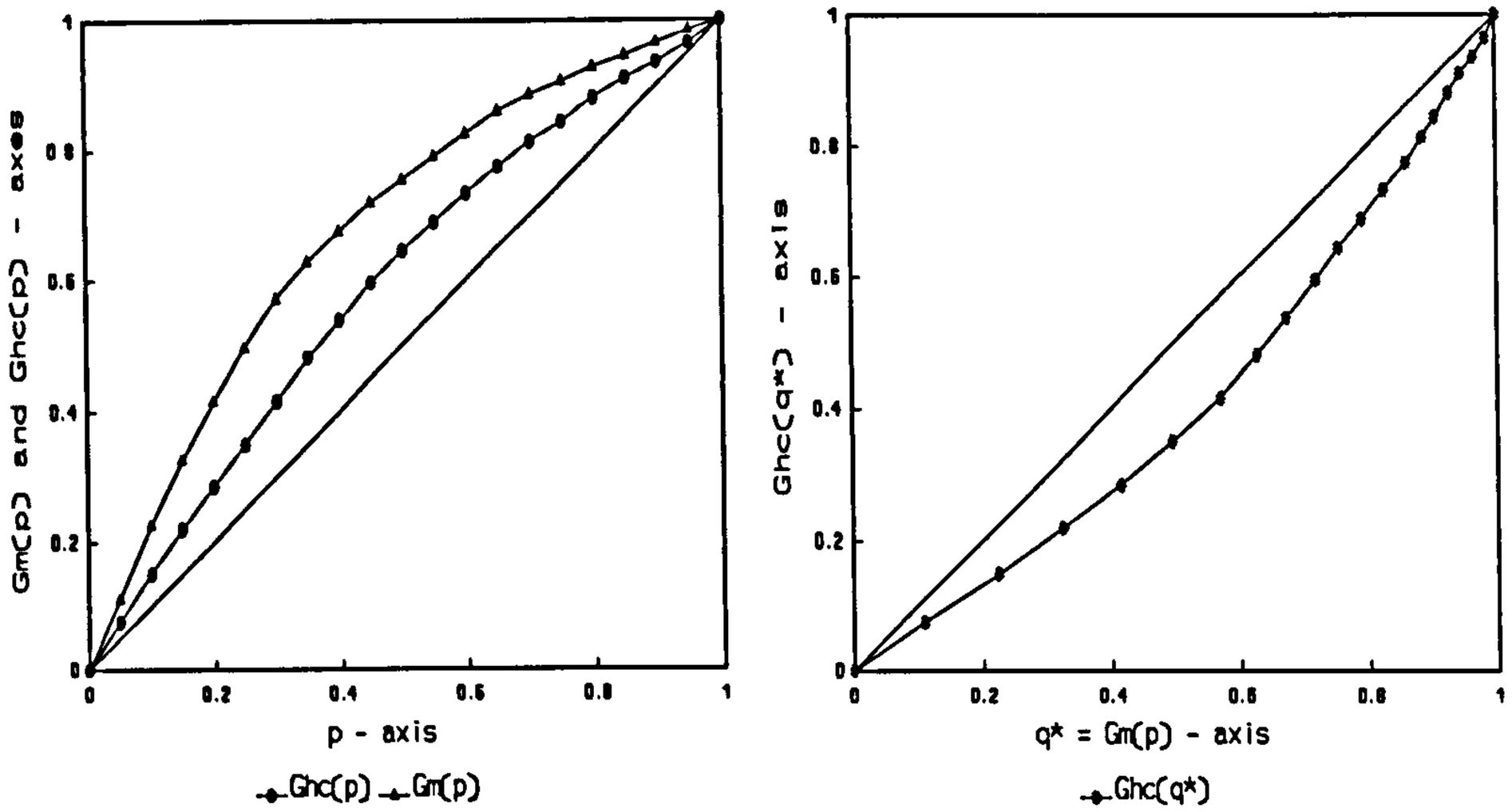


Figure 9.1 (a)-(b): Illness and utilization concentration curves - a Le Grand-type approach

expenditure but contained 31.9% of persons ill. It was the second of these calculations which led to the Wagstaff *et al* (1989) index of inequity in delivery.

In Figure 9.1(a), the curve $G_m(p)$ is an illness concentration curve of the type defined in Chapter 4. It provides the benchmark against which to assess the fairness of the utilization distribution. Consider next the variable HC, representing utilization of health care. Assume that, like M, it is related to income by a continuous function, say $h(Y)$. It follows that a *utilization concentration curve* can be generated by plotting the cumulative proportions of the income ranked population, $p = F(y)$, against the proportions of care received, $G_{hc}(p) = F_1[h(y)]$. If $h'(y) > 0$ for all $y \geq 0$, $G_{hc}(p)$ is convex to the p-axis and falls below the diagonal, implying that higher income groups receive a higher share of health care relative to their population share; if $h'(y) < 0$, then $G_{hc}(p)$ is concave to the p-axis and lies above the diagonal (as shown), signifying that poorer income groups are more intensive users of health care.

Following a Le Grand-type approach, income related inequities in the delivery of care can be examined by comparing $G_{hc}(p)$ with $G_m(p)$. Thus,

(i) if $G_{hc}(p)$ coincides with $G_m(p)$, it is implied that health care is distributed across income groups in exact proportion to illness, signifying an equitable distribution in delivery;

(ii) if $G_{hc}(p)$ lies everywhere below $G_m(p)$, higher income groups receive proportionally more medical care in relation to their levels of illness, in which case there is inequity favouring the rich [the case shown in Figure 9.1(a)]; and,

(iii) if $G_{hc}(p)$ lies everywhere above $G_m(p)$, higher income groups receive proportionally less medical care in relation to their levels of illness, and there is inequity favouring the poor.

Furthermore, the extent of inequity can be measured as twice the area between $G_{hc}(p)$ and $G_m(p)$, which is equivalent to the numerical difference between the concentration indices for utilization and morbidity:

$$I_{K\text{-}LG} = 2\int_0^1 [G_m(p) - G_{hc}(p)]\,dp \qquad (9.1)$$
$$= C_{hc} - C_m$$

where C_{hc} is a *utilization concentration index* defined analogously to C_m (see Chapter 4).[1] $I_{K\text{-}LG}$ measures disproportionality of the health care distribution in relation to the distribution of illness, when both are indexed to income. If $I_{K\text{-}LG} > 0$, there is inequity favouring the rich. Maximal pro-rich inequity of the delivery system is given by $1\text{-}C_m$ (= +2.0, when only the poorest person is sick but the richest uses all health care provided). If $I_{K\text{-}LG} < 0$, the delivery system favours poorer individuals, with maximal pro-poor inequity given by $-1\text{-}C_m$ (= -2.0, when only the richest person is sick but all health care expenditures go to the poorest). The situation where $I_{K\text{-}LG} = 0$ signifies that there are no systematic departures from proportionality, in which case the distribution of health care across income groups relative to their need, can be described as equitable.

It should be apparent that the $I_{K\text{-}LG}$ index is the delivery side analogue of the Kakwani index of payments progressivity presented in Chapter 4. It follows that the ECuity researchers could equally have proposed a Suits index of inequity in delivery:

---

[1] All indices described in this chapter are summarized in Appendix 3.

$$I_{S\text{-}LG} = 2 \int_0^1 [q^* - G_{hc}(q^*)]\, dq^* \qquad (9.2)$$

where $q^* = G_m(p)$. Like its finance-side counterpart, $I_{S\text{-}LG}$ varies from -1.0 (maximal inequity favouring the poor) to +1.0 (maximal inequity favouring the rich). The graphical representation of this index is shown in Figure 9.1(b), which uses the same information content of Figure 9.1(a) to generate a *relative health care concentration curve* $G_{hc}(q^*)$.

Which of the two indices is the more appropriate for measuring Le Grand-type inequity? In terms of graphical representation, $I_{K\text{-}LG}$ is likely to be preferred since it reveals inequality in the underlying distribution of need as well as the departure from proportionality of the utilization distribution in relation to need. However, the main criterion of choice should rest on the normative judgements implied by each measure. $I_{K\text{-}LG}$ assigns weights according to household rank in the income distribution, which implies that greater weight is given to changes in health care utilization and morbidity ocurring around the mode of the income distribution. $I_{S\text{-}LG}$ has a different weighting scheme. It assigns weights according to cumulated shares of total morbidity ranked by income. If poorer income groups reveal a higher propensity to ill-health (as we found, in Chapter 5, to be the case for Portugal), maximum significance is given to changes in health care utilization ocurring at levels of income below the mode. If there is particular concern for inequity in health care affecting lower income groups, the $I_{S\text{-}LG}$ measure appears superior to $I_{K\text{-}LG}$.

Empirical application of equations (9.1) and (9.2) is likely to provide a biased picture of inequity in the distribution of health care if the population is heterogeneous in relevant demographic characteristics along the income distribution. Le Grand (1978) faced up to this problem by standardizing the SEG related utilization and morbidity distributions by age and sex (*ie.* the same procedure used with respect to morbidity in Chapters 5 and 8). In the volume by Van Doorslaer *et al* (1993) the results of such calculations are only discussed in a few of the country reports, and then only briefly. Much more emphasis is given to the unstandardized values of $I_{K\text{-}LG}$.

Wagstaff, van Dooorslaer and Paci (1991b) later argued that, even if the population shares of M and HC are standardized, the $I_{K\text{-}LG}$ index is still a biased measure of the degree of inequity.[2] They based their critique on three shortcomings: "(i) a failure to allow for the possibility that the non-sick may be consumers of health care as well as the sick, (ii) a failure to allow for the possibility that the chronically and acutely sick may ... receive different amounts of health care, and (iii) the use of an inappropriate procedure for standardizing for demographic factors" [Wagstaff *et al* (1991b, p. 189)]. Building on the empirical approaches of Collins and Klein (1980) and Puffer (1986), which do not suffer from these drawbacks, they then suggested a new concentration measure of inequity in the delivery of care.[3] This index - denoted here by I_{WVP} - differs from $I_{K\text{-}LG}$ in taking the income ordered population shares as the reference distribution against which to judge the fairness of the delivery system. For this procedure to be consistent with the normative notion of utilization according to need it is necessary to standardize the expenditures shares by variables that proxy differences in need (*eg.* age, sex *and*, of course, morbidity). The standardized figures, can be interpreted as the expenditure each income unit would receive if it had the age, sex and morbidity of the population as a whole.

Assuming that standardized expenditure HC^* is a continuous random variable related to income by means of the function $h^*(Y)$, and has analogous properties to those defined for M and HC, it is possible to define a *standardized utilization concentration curve*, $G_{hc}^*(p)$. This curve, shown in Figure 9.2, is interpreted as the proportion of HC^* going to the lowest pth fraction of income units when these are arranged in ascending order of incomes. If there is inequity favouring the rich, $G_{hc}^*(p)$ lies below the diagonal, whereas the opposite is true if there is inequity favouring the poor.

The index I_{WVP} is defined as an area measure between $G_{hc}^*(p)$ and the 45 degree line:

---

[2] See also O'Donnell and Propper (1991a).

[3] The critique also led them to a useful regression based test for inequity in the distribution of health care consumption [Wagstaff *et al* (1991b, pp. 189-91); Van Doorslaer *et al* (1992, pp. 395-397)]. Given that the focus of this chapter is on index number approaches, the test is not discussed.

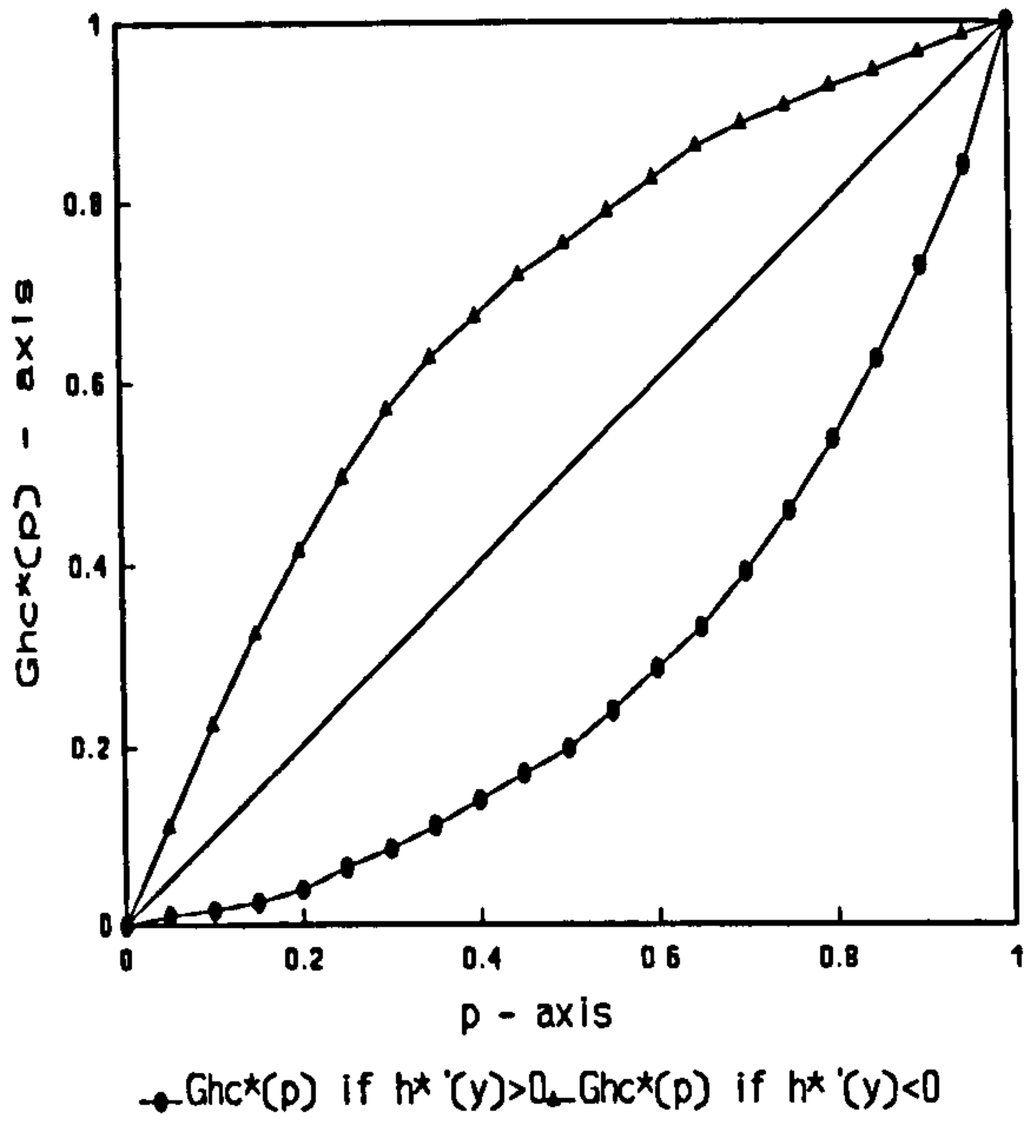


Figure 9.2: Standardized concentration curves - the Wagstaff, Van Doorslaer and Paci approach

$$I_{WVP} = 1 - 2\int_0^1 G_{hc}^{*}(p)\,dp \tag{9.3}$$

It is bounded by -1.0 (corresponding to maximal inequity favouring the poor) and +1.0 (maximal inequity favouring the rich). It takes a value of zero when there is no systematic association between standardized expenditures and a person's rank in the income distribution. The normative judgements implied by the measure are straightforward: greater weight is given to changes in standardized health care consumption around the mode of the income distribution.[4]

Based on the analysis developed in Chapter 4, one may also define parameterized versions of each of the three indices discussed above:

---

[4] The $I_{K\text{-}LG}$ and I_{WVP} indices have been used by Pereira (1992) and Pereira and Pinto (1993) to examine health care inequity in Portugal. The analyses draw on the 1987 National Health Survey used in Chapters 5 and 8 of this thesis. Using the $I_{K\text{-}LG}$ index shows that health care delivery is unambiguously favourable to the better-off. If the I_{WVP} measure is used the results are less conclusive but they still point to a slight pro-rich bias.

(i) *The family of generalized Le Grand-type Kakwani indices*

$$I_{K\text{-}LG}(\delta) = \delta(\delta-1)\int_0^1 (1-p)^{\delta-2}\,[G_m(p) - G_{hc}(p)]\,dp \qquad (9.4)$$
$$= C_{hc}(\delta) - C_m(\delta)$$

where $C_{hc}(\delta)$ is a generalized health care concentration coefficient.

(ii) *The family of generalized Le Grand-type Suits indices*

$$I_{S\text{-}LG}(\delta) = \delta(\delta-1)\int_0^1 (1-q^*)^{\delta-2}\,[q^* - G_{hc}(q^*)]\,dq^* \qquad (9.5)$$

(iii) *The family of generalized WVP indices*

$$I_{WVP}(\delta) = 1 - \delta(\delta-1)\int_0^1 (1-p)^{\delta-2}\,G^*_{hc}(p)]\,dp \qquad (9.6)$$

Once again, the sign properties of these indices are identical to their standard analogues. Values of $\delta = 2$ represent the respective standard indices. Measures of vertical inequity in delivery that are sensitive to the position of poorer persons are obtained if $\delta > 2$, and indifference to inequity is implied by measures that are computed with values of $\delta \to 1$.

The relative merits of the approaches underlying the ECuity delivery-side indices have been the subject of vigorous debate [*eg.* Le Grand (1991), O'Donnell and Propper (1991a, b), Wagstaff *et al* (1991b, c)]. Curiously, this discussion provides limited guidance as to which of the measures is generally more appropriate, since it deals largely with empirical matters. In particular, the debate concentrates on the implications of using categorical illness data as the measure of need.[5] At the same time, basic conceptual issues have been left unattended to. Thus it appears, from the debate, that the I_{WVP} index provides a conceptual advantage over the Le Grand-type indices since it

[5] A possible explanation for this 'bias' in the exchange is that virtually all important work on equity in the delivery of care has been carried out in the U.K., drawing on data from the General Household Survey. Categorical illness data are the only morbidity measures available in that survey.

addresses the issue of the non-sick also being consumers of health care. Ultimately, however, this edge is the result of specificities in health survey data. In a world where ratio-scale measures of illness were readily available, the three indices would provide virtually identical results when applied to whole populations. Even if these measures are not available but one is able to operate with traditional health survey indicators that characterize the severity of illness rather than its prevalence (*eg.* the number of disability days), much of the difference between the results provided by I_{WVP} and the $I_{K\text{-}LG}$ and $I_{S\text{-}LG}$ indices vanishes, since both the sick and non-sick population are used in determining the reference distribution of need.[6]

In all publications deriving from the ECuity research, $I_{K\text{-}LG}$ and I_{WVP} are taken to be measures of the degree of *horizontal* inequity (*ie.* income related violations of equal treatment for equal need). This interpretation is not, however, particularly compelling. Generally, the indices actually measure departures from proportionality of the income-ranked utilization distribution (standardized for equity relevant characteristics or not) in relation to some reference distribution (income-ranked illness or population shares). In each case, there is an implicit assumption that inequality in the distribution of health care is fair if it is proportionate to inequality in the reference distribution. This is usually accepted to be a vertical equity principle, and indeed in the finance side analysis analogous indices are used to measure that same principle.[7]

Of course, vertical equity, if interpreted in the Aristotlean sense of unequal treatment *in proportion* to the unequal situation, has a logical corollary in the principle of equal treatment of equals. In empirical applications, this association sometimes leads to a thin data-related dividing line between measurement of vertical inequity and its horizontal counterpart. This seems to be the case with the ECuity research. It is

---

[6] This contention requires that the Le Grand-type indices are standardized for demographic differences along the income distribution. The point seems to be implicitly accepted in the later publications of the ECuity group [*eg.* Van Doorslaer *et al* (1993, pp. 78-80)].

[7] It is interesting to note that in Le Grand's (1978) article there is no claim, explicit or implicit, that horizontal inequity is being measured. Such an assertion is, however, made in his later book [Le Grand (1982, p. 46)]. Perhaps due to this or simply because of the overwhelming emphasis given to the horizontal equity principle in British health policy documents, subsequent research has always taken for granted that Le Grand's initial analysis was directed at that issue.

implicitly assumed that indices such as $I_{K\text{-}LG}$ and I_{WVP} can measure the degree of horizontal inequity, provided that 'equals' can be identified. This raises the question of how 'equals' are to be distinguished in empirical analysis? The answer provided by the ECuity researchers is to assume that any two persons share the same level of need for health care if they answer positively to a survey question of the type "Do you have any long-standing health problem or chronic illness?" However, this is a particularly strong assumption which is ultimately contradicted by the results reported in the study. The analysis for Denmark, for example, shows that if the number of chronic conditions are used (rather than relying on information of whether the individual suffered from a chronic illness) the values of the I_{WVP} index are significantly affected. More detailed information on illness leads to index values which show inequity to be less favourable to the poor [Van Doorslaer *et al* (1992, pp. 404-5)].[8] The main reason for this effect appears to be that *within* crude morbidity categories there are income related variations in the severity of illness which generally favour the rich.

A possible answer to this problem - using the the same widely available health survey data - would be to disaggregate the analysis, either by illness-type or by number of conditions or illness days, thereby identifying persons in 'equal need' more accurately. However, there are at least three reasons why it might not be advisable to measure horizontal inequity in this particular manner. One, is that the level of disaggregation required to make the analysis meaningful will tend to overburden empirical work. The number of what might be accepted as 'equal health situations' is immense and the researcher could conceivably end up carrying out a separate analysis for each. Related to this is the second point, that it may be difficult to obtain agreement on what exactly constitutes an 'equal health situation'. For instance, how would one decide if a person suffering from a stomach ulcer is equal or not to one suffering from chronic bronchitis? Arguably, given the present state of knowledge regarding such prioritization of 'needs', it would be wiser for economists to opt for a notion of horizontal equity that is less demanding in terms of information requirements. Finally, even admitting that disaggregations were accomplished in an acceptable manner, it

---

[8] The country chapters from the Netherlands, U.K. and Portugal in van Doorslaer *et al* (1993) also report information from other studies which cast doubt on the assumption; as indeed does the research reported in Chapter 5 of this thesis.

would still leave open the question of how vertical inequity in delivery ought to be measured. If an index number approach was sought, the most suitable candidates would, quite probably, be $I_{K\text{-}LG}$, $I_{S\text{-}LG}$ and I_{WVP}, which would suggest that there is no conceptual difference between horizontal and vertical equity, merely an empirical one.[9]

In conclusion then, (i) the difference between LG- and WVP-type inequity ultimately rests on the nature of categorical survey data; and (ii) each of the indices described can only serve as measures of horizontal inequity if persons with 'equal levels of health status' can be unambiguously identified. In this chapter it is assumed that the second proposition is unlikely to be attained with traditional health survey data, and that it is therefore more appropriate to adopt the indices' more obvious interpretation as measures of vertical inequity. Their relative merits then depend simply on the descriptive properties and the normative judgements which they imply. Given, however, that policy statements often make reference to income related violations of the horizontal equity principle, the problem can hardly be left unexamined. Consequently, attention is next turned to an alternative method for measuring horizontal inequity in the delivery of health care.

9.3 A new index of horizontal inequity

When studying the redistributive effects of tax and transfer policy, modern approaches in Public Economics make a crucial distinction between horizontal and vertical equity principles. Plotnick (1981), for example, argues that "the principle of horizontal equity addresses the fairness of a *process* of redistribution, a measure does not, and should not attempt to compare the actual extent of redistribution or change in inequality to some exogenous criterion" (p. 283). Vertical equity, on the other hand - he goes on to argue - is an end-state principle: it compares the observed distribution

---

[9] It is important to note that none of these arguments signify that the ECuity approach is irrelevant to measurement of horizontal inequity. The logic of direct evaluation of whether 'equals' are treated equally is itself appealing, and has strong traditions in the taxation literature (*eg.* Johnson and Mayer, 1962; Berliant and Strauss, 1985). What is being questioned is the over-reliance of the approach on restrictive data-related assumptions.

of an attribute to an ideal distribution. This distinction has important implications for measurement and has led to the term horizontal equity (HE) being assigned a precise meaning in the context of redistributional analysis. HE is said to hold if the redistribution brought about by a tax or transfer *does not alter the rank order of units by utility level.* This idea, initially proposed by Feldstein (1976), forms the basis of some well-known summary indices of horizontal inequity: namely the Atkinson (1980)-Plotnick (1981) concentration index and the measures suggested by King (1983) and Cowell (1985), which draw respectively on the normative and axiomatic approaches to inequality measurement. In this section the underlying method used by Atkinson and Plotnick is applied, in a partial symmetry framework, to the issue of horizontal equity in health care delivery.

It could be argued that the notion of horizontal equity as absence of utility rank reversals is irrelevant to health care delivery. It is conceivable, for instance, that public intervention in health care carries an implication to the contrary. Policy makers might view subsidised consumption of health services as some sort of 'social wage' which raises the utility levels of the poor beyond that which the economic sphere provides them. However, it should be noted that the application provided here is based, like the rest of the thesis, on the 'extra-welfare' approach which underpins much of modern applied health economics (Culyer, 1989). It bypasses the utility criterion and concentrates simply on the relationship between utilization and need, which are taken to be the entities of ethical interest. In common with the ECuity approach, it is assumed that society aims for a distribution of resources in health care delivery that reflects the distribution of need (measured by presenting health status).

Consider, therefore, the following definition:

D 9.1. *A horizontally equitable health delivery system is one that preserves the same rank order of units in the utilization distribution as in the underlying distribution of need (proxied by individual morbidity levels).*

This conception encompasses the classic definition of equal treatment of equals. Logically, any two persons that share the same rank in the illness distribution will also

have the same level of health. They will be 'treated equally' if, and only if, their rank positions are identical in the utilization distribution (*ie.* enjoy the same level of expenditure). However, the definition is also more general and conceptually complete. As King (1983) observes: "In practice, ... no two individuals are ever identical, and the principle of equal treatment of equals has little empirical significance unless it can be usefully extended to include *and unequals treated accordingly*" (p. 101). This view might raise the objection that HE refers only to equal treatment of those with equal need and that non-correspondence of ranks is really a vertical equity principle. But this conclusion does not follow. Vertical inequity comparisons measure the distance between an observed distribution and an ideal one. They ignore the question of rank reversals. Therefore, the above specification *does* relate to the question of horizontal equity, but it is a 'stronger' form of HE than that implied by the narrow 'equals treated equally' view (Jenkins, 1988b).

Definition 9.1 says nothing, however, of income related violations of the HE principle, which have been the primary focus of studies by health economists (and of policy concerns in Portugal). In order to tackle this problem consider also:

D 9.2. *For any mutually exclusive and equity relevant K partitions of the population, inter group differences in the achievement of HE are given by comparing the respective K measures of within group horizontal inequity.*

If the K partitions are income ranked percentiles of the population, it is possible to compare the degree to which HE is being achieved in each income grouping. Thus, the question of income related violations *can* be addressed, but there is no requirement that persons in 'equal need' be identified through arbitrary categorization based on survey data. The approach shifts the issue of inter-personal comparability from the variable 'morbidity' to the variable 'income', where a broader base of agreement about partitions (*eg.* percentiles) is likely to be forthcoming. In this sense at least, the method should facilitate applied work. Note, however, that it requires access to ordinal morbidity indicators with a reasonable number of classes or a transformation that assumes an underlying latent morbidity variable (see Wagstaff and van Doorslaer, 1994), so that the ranking of need can be established.

Following Cowell's (1980) discussion of comparability in inequality measurement the approach is labelled one of 'partial symmetry'. Individuals within each income group are treated anonymously by the analysis, but differentially *between* groups, given that income position is seen as relevant to horizontal inequity calculations. Thus, partially symmetric summary measures are computed - one for each income partition - and then compared to those of other groups.[10]

Assume that the population can be divided into K (k = 1, ..., K) mutually exclusive income partitions. To devise a measure of horizontal inequity in the *k*th group, we proceed very much as before drawing on the continuous variable framework for expositional reasons. Allow that illness M is a continuous random variable with p.d.f. $f(M)$, distribution function F(M) and mean μ_m.[11] In contrast to Section 9.2, no strict relationship between income and health is conjectured. The proportion of individuals having an illness level less than or equal to m is given by $r = F(m) = \int_0^m f(M)dM$, while the proportional share of total morbidity experienced by units having an illness level less than or equal to m is given by $F_1(m) = 1/\mu_m \int_0^m Mf(M)dM$. Both these functions vary from 0 to 1. The relationship between them defines the Lorenz curve for morbidity, shown in Figure 9.3 as $L_m(r)$.

Next assume that health care utilization, HC, is related to morbidity by the continuous function $hc = g^+(m)$, and that $g^+(M) \geq 0$ for all $M \geq 0$. It follows that the Lorenz curve for the health care distribution, $L_{hc}(s)$ in Figure 9.3, is obtained from the relationship between $s = F[g^+(m)] = F(hc) = \int_0^{hc} f(HC)dHC$ and $F_1[g^+(m)] = F_1(hc) = 1/\mu_{hc} \int_0^{hc} HCf(HC)dHC$. Note that both $L_{hc}(s)$ and $L_m(r)$ are standard Lorenz curves: in each case, persons are ordered on the abcissa by ascending levels of an attribute, the corresponding shares of which are shown on the ordinate. The relative configurations shown in Figure 9.3, indicate that there is less inequality in the distribution of health care, from whence it could be argued that there is vertical inequity in delivery *within*

---

[10] A similar approach has been suggested by Jenkins (1988b) for measuring horizontal inequity in taxation. He describes his method as taking a position intermediate between the utility reranking approach and the traditional 'equals treated equally' view.

[11] In order to facilitate notation subscripts *k* are not presented, but it is important to recognize that the analysis pertains only to the *k*th group.

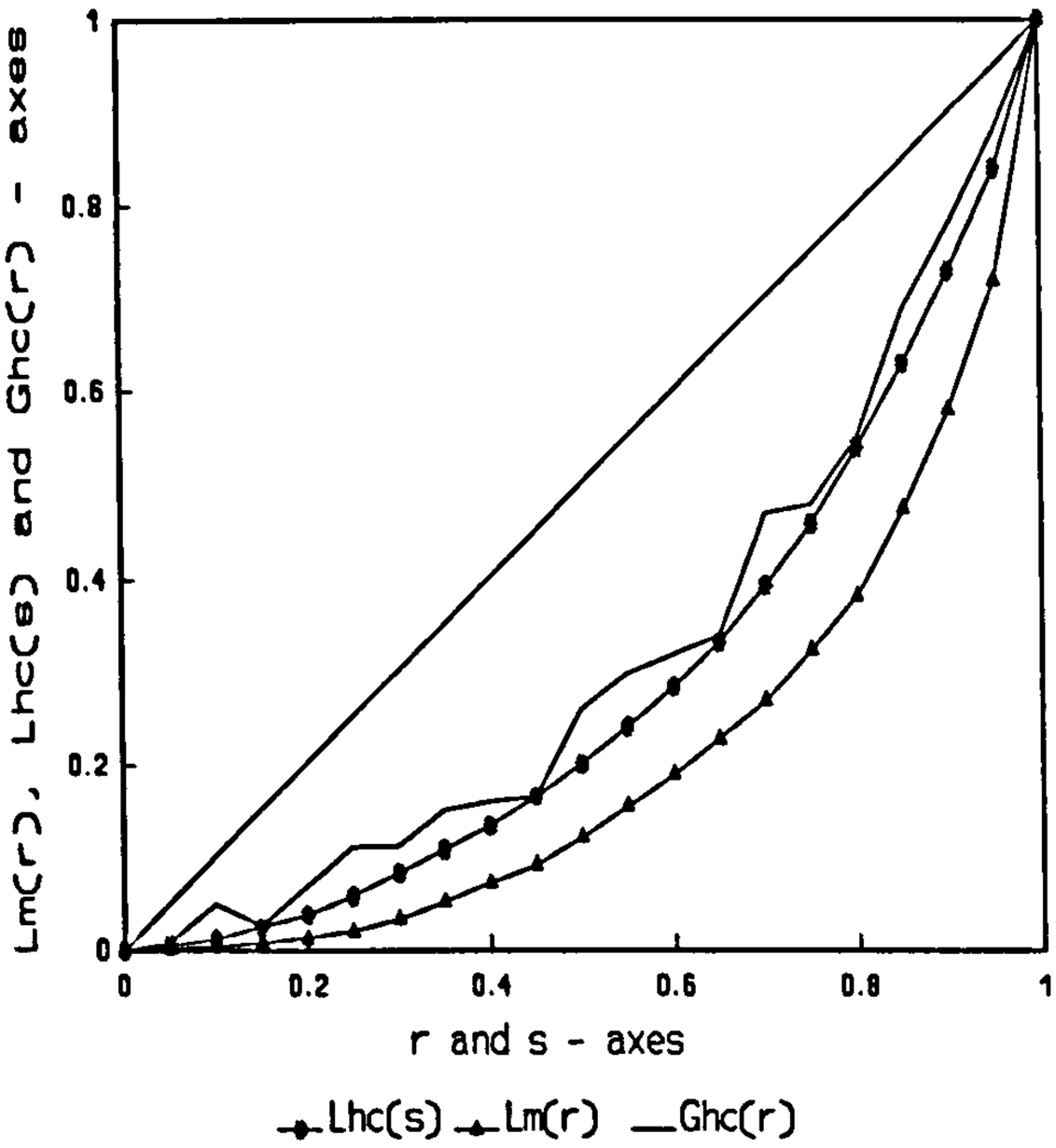


Figure 9.3: Standard Lorenz curves for morbidity and health care and a concentration curve for health care ordered by morbidity

the *k*th group. Comparisons of $L_m(r)$ and $L_{hc}(s)$, however, reveal nothing about horizontal inequity in the sense defined above, since a given s fraction of individuals may not be the same persons that are included in the equivalent *r*th decimal fraction.

Consider a new curve $G_{hc}(r)$, which shows the relationship between F(m) and $F_1(hc) = F_1[g^+(m)]$. In other words, this is the concentration curve for health care utilization when units are arranged in ascending order of illness levels. $G_{hc}(r)$ has the following properties[12] :

(i) $G_{hc}(r)$ lies below (above) the diagonal *iff* $\eta_{g+}(m) = g^{+\prime}(m) \cdot m \,/\, g^+(m)$, the elasticity of health care utilization with respect to health, is positive (negative) everywhere;

(ii) $G_{hc}(r)$ lies below, on or above $L_m(r)$ *iff* $\eta_{g+}(m)$ is greater than, equal to or less

---

[12] See Kakwani (1980b, Chapter 8) for general proofs of the first three properties and Atkinson (1980) for a proof of the fourth.

than unity at all values of m;

(iii) if the function $g^+(m)$ has a continuous derivative $g^{+\prime}(m)$ strictly positive for all $m \geq 0$, $G_{hc}(r)$ coincides with $L_{hc}(s)$;

(iv) for all values of m, $G_{hc}(r)$ lies everywhere on or above $L_{hc}(s)$.

An important implication of these properties is that, if the ranking of all units in the morbidity distribution corresponds to the ranking in the health care distribution then $G_{hc}(r) = L_{hc}(s)$ for all *r* and *s*. This suggests that the degree of horizontal inequity can be indexed by measuring the area between $L_{hc}(s)$ and $G_{hc}(r)$.

For each of the curves exemplified above, area measures can be calculated. Call $Z_{m/k}$ the Gini for morbidity, $Z_{hc/k}$ the Gini for health care utilization, and $C_{hc/k}^+$ the concentration index for health care. As with any Gini measure $Z_{m/k}$ and $Z_{hc/k}$ each vary from 0 to 1. $C_{hc/k}^+$, on the other hand, is bounded by $-Z_{hc/k}$ and $Z_{hc/k}$.[13]

The extent of horizontal inequity within the *k*th income group can be measured as:

$$HI_k = \frac{Z_{hc/k} - C_{hc/k}^+}{2\, Z_{hc/k}} \qquad (9.7)$$

which is a variant of the Atkinson/Plotnick index of horizontal inequity. The numerator measures the extent of non-correspondence in ranks between the health and health care distributions within the *k*th group. Note that non-correspondence of rank positions in other income groups has no effect on a given index of within group horizontal inequity.

---

[13] A proof of this proposition can be obtained by substituting the relevant values in equation (4.9) of Chapter 4 so that it reads:

$$C_{hc/k}^+ = \frac{R[g^+(m), r(m)]}{R[g^+(m), r(g^+(m))]}\, Z_{hc/k}$$

If $g^{+\prime}(m) \geq 0$ for all m, m and $g^+(m)$ will have exactly the same ranking, in which case the correlation between $g^+(m)$ and $r(g^+(m))$ will be equal to the correlation between $g^+(m)$ and r(m). It follows that $C_{hc/k}^+$ is always positive and equal to $Z_{hc/k}$. If $g^{+\prime}(m) < 0$ for all m, m and $g^+(m)$ will have exactly opposite ranking and the correlations $R[g^+(m), r(m)]$ and $R[g^+(m), r(g^+(m))]$ will be of the same magnitude but opposite sign. In this case, $C_{hc/k}^+$ is equal to minus one times $Z_{hc/k}$. If $g^+(m)$ is not a monotonic function, $-Z_{hc/k} \leq C_{hc/k}^+ \leq Z_{hc/k}$.

The denominator, which is equivalent to the maximum possible area between $L_{hc}(s)$ and $G_{hc}(r)$, provides a normalization which ensures that HI_k is bounded by zero and one. If $HI_k = 0$, the ranking of units by morbidity corresponds to the ranking by health care utilization, in which case horizontal equity is said to hold. There is *equal treatment for equal need* in the sense that each person attains the *same* ranking in the health care distribution as he/she holds in the distribution of illness. Values of $HI_k > 0$ reveal that one or more persons do not have equivalent rankings. With the limiting case of $HI_k = 1$, it is implied that the person with the lowest ranking in the morbidity distribution is also the one who receives the highest value of health care utilization or expenditure.

The partial symmetry approach emphasises that information disaggregated by equity relevant characteristics is important. Therefore, differences in the attainment of HE *across* income groups are obtained by comparing the K measures of within-group horizontal inequity. If, for example, there are two income groups - poor and rich - and $HI_p < HI_r$, then it is implied that horizontal inequities in health care delivery favour the poor. In some situations, however, the overall level of horizontal inequity may also be a factor in social decisions. For these, the computation of a population index through straightforward weighted aggregation is suggested:

$$HI_K = \sum_{1}^{K} \omega_k HI_k \qquad (9.8)$$

where ω_k is group k's population share and $\sum_k \omega_k = 1$. Such an index can also be used as the denominator when measuring relative dispersion in HI_k values.

Table 9.1 exemplifies some hypothetical horizontal inequity comparisons of income quintiles ordered from poorest to richest. In distribution A, the HI_k indices for the better-off reveal lower values than those of poorer groups. This implies that the rich are more successful in obtaining health care resources that are consistent with their levels of need, which would point to horizontal inequity favouring the rich. Distribution B shows the reverse situation. In distribution C there are no differences between the HI_k indices leading to the conclusion that horizontal inequities are not related to income position. Suppose, however, that a decision maker places some value on the level of overall horizontal inequity, then it is possible that distributions A and B would be

Table 9.1: Comparison of horizontal inequity measures for different income groups

Income Quintiles	Hypothetical distributions			
	A	B	C	D
1	0.5	0.1	0.5	0.2
2	0.4	0.2	0.5	0.2
3	0.3	0.3	0.5	0.15
4	0.2	0.4	0.5	0.15
5	0.1	0.5	0.5	0.15
Overall	0.3	0.3	0.5	0.17

preferred to C, given that overall $HI_K^C > HI_K^A = HI_K^B$. Such a possibility is all the more transparent if distribution C is compared to distribution D. In the latter, there is once again horizontal inequity favouring the rich but of a much reduced degree in relation to A. It could be argued that the small absolute and relative differences between the HI_k's do not warrant concern and that distribution D is preferable to C. Thus, the approach allows for the fact that the overall level of horizontal inequity is also important for social decisions.

The 'partial-symmetry rank-correspondence' approach outlined here has a number of advantages over previous efforts by health economists to index horizontal inequity in delivery. One such advantage is given by its inherent flexibility. The weights ω_k in equation (9.8) could be viewed as 'discrimination coefficients', with those groups recognized as more deserving receiving greater weight in assessments of the overall level of horizontal inequity. Furthermore, the approach can be readily extended to examine any ethically relevant violations of the HE principle. It is simply a matter of identifying non-trivial partitions of the population (*eg.* regions, occupations, age groups) and calculating partially symmetric indices for each group. A certain flexibility is also implied by the possibility of *value* trade-offs between group and population inequity, as exemplified in Table 9.1. This might appear to provide a somewhat imprecise measurement of inter-group inequities, but in fact it highlights a further strength of the

approach, which is the richness of the information content. Not only are we able to observe the degree of horizontal inequity across groups but also its relation to the overall level.

However, the main advantages of the approach are that it provides a conceptual distinction between measurement of horizontal and vertical inequity and facilitates empirical work on system-wide violations of the HE principle. Under the ECuity method, the difference between horizontal and vertical inequity is not specified but appears to be given simply by the nature of the morbidity data being used. In the present approach, HE is given by correspondence of rank positions in the health and health care distributions (the difference between $Z_{hc/k}$ and $C_{hc/k}^{+}$). Vertical inequity, on the other hand, represents deviations of the health care distribution in relation to an ideal allocation (*eg.* the difference between $Z_{hc/k}$ and $Z_{m/k}$). The ECuity method appears, at first hand, to be less demanding in terms of information requirements. However, the procedure can only produce meaningful evaluations of 'horizontal inequity' if (i) health status measures that allow for unambiguous identification of persons in 'equal need' are available (which is generally not the case); or (ii) the information on morbidity is highly disaggregated by pathology (in which case empirical work becomes cumbersome). For system-wide assessments of horizontal inequity (related to income or other ethically relevant categories) the approach suggested here is more appropriate. It changes the nature of the assumptions regarding interpersonal comparability in the morbidity distribution. There is no requirement that persons in 'equal levels of health status' be distinguished directly from the empirical base. By substituting the concept 'utilization equivalence within health status groupings' with the more amenable concept of 'rank equivalence', the approach therefore facilitates empirical work.

9.4 Conclusion

The ECuity approach to studying inequities in the delivery of health care constitutes a significant development on earlier work by Le Grand (1978), Collins and Klein (1980) and others. The major contribution appears to have been the introduction of index numbers that allow the extent of inequity to be quantified. All ECuity research

to date claims to have to have measured the attribute of horizontal inequity. However, this interpretation is only valid under quite restrictive data assumptions. The present chapter has argued that the ECuity indices, which measure departures from proportionality of the utilization distribution *vis-à-vis* the distribution of 'need' (when both are indexed to income levels), are best seen as measures of vertical inequity. This interpretation is consistent with other uses of the concentration curve methodology in other areas of economics, including the ECuity approach to finance-side inequity (see Chapters 4 and 7). Other indices of vertical inequity in delivery were also proposed: namely, a Suits index of Le Grand-type inequity that attaches greater weight to inequalities at the bottom end of the income distribution; and generalized versions of the $I_{K\text{-}LG}$, $I_{S\text{-}LG}$ and I_{WVP} indices that permit preferences regarding the degree of equity to be achieved to be made explicit. Finally, the chapter proposed a new index of horizontal inequity in delivery. This measure is based on the correspondence of rank positions in the utilization and health status distributions. Given that it bypasses the need to empirically identify persons with 'equal levels of health status', it should facilitate applied work. However, it does require the researcher to have access to morbidity indicators with a reasonable number of classes (or transformations that assume an underlying latent morbidity variable), so that the ranking of health status may be established.

Part 6

CONCLUSIONS

Chapter 10

Equity, Health and Health Care in Portugal: Summary and Implications

Following the publication of the Black Report, a vigorous scientific debate has developed on the nature, extent and causes of inequity in the sphere of health. It has spanned a great many countries with particular incidence on the european continent. Most of the research has been undertaken by epidemiologists, demographers and social scientists. Economists initially took a back-seat in the polemic but, more recently, have begun to uncover the immense corpus of knowledge in the discipline that may be used to study the problem. In the policy arena, despite some hesitation by national governments, reduction of health-inequity has remained a priority issue. Most, if not all, developed countries profess to safeguard the position of economically disadvantaged individuals, and many have devised policies aimed at abating what are perceived as unjust health and health care distributions.

This study has been primarily concerned with measuring the structure and time path of health-inequity in Portugal, a country where despite the evident concern of policy-makers, detailed empirical knowledge of the problem is scarce. It has also sought to examine and develop economic measurement procedures that are able to shed light on the question, whether in Portugal or abroad. By way of conclusion, the present chapter provides a summary of the thesis and identifies the main implications for policy initiatives and economic analysis of equity in the domain of health.

The thesis began with an overview of issues relevant to the study of health-inequity in Portugal. Basic features of the health and health care systems were described and policy statements and previous research were examined. Given the

somewhat opaque nature of policy objectives, the study proceeded to examine various distribution rules that might shed light on the concerns expressed by portuguese policy makers and be used as a means of organizing empirical research. An alternative rule based on Sen's notion of equality of capabilities was proposed, and from this analysis an agenda of positive investigation was developed. The key features of the approach are a focus on (i) the actual levels of the functioning 'good health' that individuals with differential command over resources obtain, and (ii) their experience in the space of commodities that are instrumental to attainment of the desired functioning. Whilst this approach fails to capture the complete essence of Sen's framework, it has the advantage of converging on specific issues identified in policy statements and of producing results that are comparable with previous research.

The study then set out the means by which inequity in health and health care was to be measured. It adopted the concentration index approach recently used by other authors. The normative and statistical properties of health inequality and health care finance progressivity indices were clarified and new measures that take due account of alternative social judgements concerning the degree of equality preference were introduced. The analysis also laid the basis for further theoretical innovations in the latter part of the thesis: namely, a rationalization for a family of indices that measure social welfare in the domain of health (defined over the twin policy attributes of equity and efficiency); and a new index of horizontal inequity in the delivery of care, based on the correspondence of rank positions in the utilization and health status distributions. Each of the new developments provides measurement tools in areas where previous research was either silent or debatable.

Empirical analysis was directed at four issues. First, the economic distribution of morbidity. Strong evidence was found showing that the burden of illness in Portugal is generally unfavourable to poorer income groups. In particular, these groups tend to suffer from more severe conditions and reveal slower recovery periods. When the data were disaggregated by age, it was found that poor middle-aged individuals are particularly disadvantaged. In contrast, there appears to be no economic differentiation in the health experience of children and adolescents (although this evidence may be unreliable due to the nature of available morbidity indicators). The second issue

addressed was the time-trend of inequity in infant mortality. The findings suggest that the degree of pro-rich bias is currently much lower than in any period since the early 1970's. There are, however, important counteracting trends. Close-to-birth mortality inequity seems to have been largely eradicated; but the distribution of post-neonatal deaths is once again becoming less favourable to the poor, after showing the greatest improvements during the 1970's.

The third empirical analysis was aimed at measuring the degree of progressivity of the health care financing system in the period 1980-1990. The main finding was that the system evolved from being overall progressive to overall regressive. The burden of payments shifted to middle income groups, with the main beneficiaries being households situated in the richest quintile of the income distribution. The intertemporal change resulted from two major factors: reduced progression of the tax system and an increase in the share of revenues raised directly from consumers. It was also shown that alternative revenue raising sources have distinct progressivity characteristics. Portuguese taxation is moderately progressive and out-of-pocket payments are highly regressive. Social and private insurance contributions were found to be highly progressive, but the result is largely attributable to selective coverage of better-off households.

A fourth issue - distribution of social welfare associated with the health distribution - was examined in less detail, given that the purpose was simply to illustrate the new measures proposed in the study. Nevertheless, two findings are worthy of note. Geographically, the Centre and Algarve regions have consistently worse levels of health social welfare than other parts of the country. However, if there is a concern for inequity the position of the Algarve becomes noticeably less detrimental. This result has obvious implications for regional resource allocation mechanisms that aim to compensate areas with lower levels of health by attributing them higher amounts of health care resources. Disaggregation of health social welfare by income groups showed that it is highest for the rich irrespective of the social preference for equity or efficiency. The results also suggest that, if society is concerned about income related health deprivation, distributional comparisons of health status levels tend to underestimate the true degree of inequality.

Given that the focus of the thesis has been on measurement rather than explanation, the scope for making policy recommendations is limited. Nevertheless, some general implications can be drawn. The morbidity results suggest that there is a need to enhance the health investment opportunities of the poor. Possible interventions include food subsidies; improvements in housing and working conditions; better access to education; and more generally, income redistribution policies through the tax and benefit system. If health care is deemed instrumentally important in attaining health potential, then effective health care services should be made equally available to all. This may require adjustments to payment mechanisms; training of medical personnel to improve their awareness of the particular health difficulties faced by the less well-off; or prevention and rehabilitation measures specially targeted at the economically disadvantaged. Which of these or related policies is the most appropriate can only be determined by explanatory analysis. The results do suggest, however, that priority should be given to alleviating the unequal burden of illness supported by middle-aged poor individuals. The large differential has inevitable resource consequences in the form of lost production and increased use of public health and social services. It also imposes high costs on poorer families if breadwinners have to withdraw from the labour market.

The infant mortality results do not signify that equitable distribution is no longer an issue in this area. The recent downturn in post-neonatal inequity and the position of the very poorest are two specific questions that need to be monitored carefully. Although many of the social and economic circumstances which lead to infant deaths no longer apply in Portugal, some deaths could probably be avoided. This might best be achieved by targeting social support at persons in extreme poverty. Equally, the results imply that it may be useful to target non-economic factors known to affect the survival chances of infants (*eg.* very young mothers or the spatial distribution of medical technology).

The progressivity results indicate that a greater emphasis on out-of-pocket payments, as recently advocated by the Minister of Health, is likely to involve poorer income groups supporting a higher share of the health care financing burden. Such an outcome might be avoided if NHS co-payments are related to ability to pay. The results show that expenditure on pharmaceuticals - the main area where the NHS fails to

differentiate between consumers on the basis of income - is the principal factor affecting the highly regressive structure of direct payments. It is unclear from the results if a shift towards more insurance financing would make the system more or less progressive. However, evidence from other countries suggests that increased population coverage would tend to lower progressivity levels and even make the distribution of insurance payments regressive. Despite a noticeable fall in the progressivity of the tax system, tax revenues still appear to be the choice instrument for lessening the burden of health care payments on the poor. Perhaps the most important feature of the results, however, is that solidarity in financing has been diminishing at a rapid pace. This observation needs to be borne in mind when considering reform the health care financing system.

Although this study has gone some way to uncovering the health-equity profile of the portuguese population, many questions have been left unanswered. The empirical strategy was deliberately aimed at measuring the extent of inequity, and even then at a general level, given that previous research is so limited. There is still a case for pursuing this type of research in future. Within the next two years new data from national health interview and budget surveys will become available. It would be useful to replicate the present analyses on those data and more generally to continue to monitor developments in inequity levels. Similarly, the methods discussed in the latter part of the thesis should be subjected to detailed empirical application. Research efforts should, however, be mainly directed at examining the causes and policy implications of the broad level relationships identified in the study. Two basic strategies appear the most promising. First, using the same measurement procedures after further data disaggregation. For example, the progressivity analysis might be carried out on payments to the NHS rather than to the system as a whole; or one might examine the impact of tax deductibility of health care expenditures on the overall level of progressivity. The finance side techniques could also be used to measure the redistributive impact of reform proposals through simulation analysis. Ultimately, however, the evidence produced in the thesis calls for causal analysis based on behavioural models. This should allow us to verify, for instance, the precise mechanisms whereby poorer individuals end up with lower levels of health than those that are better off; and eventually to design policies that effectively remedy the situation.

Much of the thesis has been concerned less with the portuguese situation and more

with developing economic procedures for analysing inequity in the domain of health. It is useful to also review the implications in this respect. A first point is that the ethical framework adopted for the thesis suggests that the common empirical focus on issues like socio-economic inequalities in health status and health care utilization is not as misconceived as some authors have suggested. Such research may be seen as a practical compromise (imposed by data limitations) aimed at assessing unequal opportunity to realize health potential. It does not provide all the necessary information but nonetheless offers useful insights into accomplishment of the underlying goal.

Second, the thesis suggests an important future role for health-inequity measures that take due account of alternative social judgements. Researchers have typically used indices that entail (possibly) unappealing assumptions about the weight attached to inequity at different points of the reference distribution. Such measures also presume a degree of agreement about equity judgements, which is rarely found in practice. Income inequality analysts long ago recognized these problems. In future, health economists should adopt a similar posture when measuring the extent of health domain inequity by employing parametric measures. The concentration based indices applied in this study are one possibility but another is the family of generalized entropy indices which have the advantage of being additively decomposable. This property is particularly useful if there is an interest in the relationship between variables that contribute to health and health care differences and observed levels of inequality. Indirectly, the use of parametric indices also points to another desirable development. In the past health economists have been essentially concerned with global inequity levels. It may be instructive in future to focus on specific parts of the income distribution. For example, by comparing health status levels of persons above and below some accepted poverty line or by computing measures of local progression when assessing progressivity of the health care financing system.

Third, the thesis has proposed a clear demarcation between measurement of horizontal and vertical inequity in the delivery of care. The approach, which involves considering horizontal violations as non-correspondence of rank positions in the health care utilization and health status distributions, brings current methods employed by health economists more in line with those used in other areas of economics. Future research should provide empirical applications of the new indices in order to better

evaluate their potential. These investigations may require more refined health status measures than are usually employed, possibly involving statistical transformations of categorical morbidity variables.

Another avenue opened up by the present research is integrated measurement of the equity and efficiency attributes of health distributions, described here as assessment of health social welfare. In the past, the performance of health systems with regard to distributional and aggregative goals has been examined in isolation; largely, one suspects, because of the underdevelopment of techniques for joint measurement. This study has shown that health concentration indices have an ethically founded role as measures of inequity. The new indices of health social welfare allow for an explicit trade-off between income related inequity and health maximization. They have a considerable number of potential applications which suggests that they may find a niche in the assemblage of empirical tools used by health economists. Obviously, far more detailed practical application is required before a final judgement is pronounced. There is also a need for further conceptual analysis aimed at devising alternative measures and for research that elicits societal values of the equity-efficiency trade-off parameter, in this case possibly drawing on experimental methods.

Finally, the thesis has shown that far greater attention should be given to the impact of measurement procedures on the degree of measured inequity. Distributional analysis in the field of health invariably involves the adoption of methods over which there is controversy (*eg.* the procedure for aggregating inequality, choice of equivalence scale, incidence assumptions, and so on). It may be that, as was found for the intertemporal progressivity analysis, choice of methods has no noticeable impact on the conclusions that are drawn. However, it was also shown that health inequity estimates involving income as the reference distribution are particularly sensitive to equivalence scale relativities. This feature can lead to conflicting results when levels of inequity are compared across countries. Future studies should seek to replicate the sensitivity analyses carried out in order to better evaluate which type of methodological choices, and in what contexts, systematically affect inequity measurements. If it is found that some issues have a confounding influence on results, then there is a strong case for general adoption of the eclectic approach to measurement espoused in the thesis.

APPENDICES

Appendix 1

An Exploration into the Modelling of Health Related Behaviour

A1.1 Introduction

The thesis has been essentially concerned with examining inequalities in health and health care variables and their systematic association with inequality in the extent of choice over commodity bundles. These relationships have been examined in isolation because of their importance in establishing the degree to which health-equity has been attained. However, if the capabilities approach is adopted as the underlying ethical theory, measurement of health domain inequity must also consider the process whereby goods (*eg.* health care, education) are transformed into health functioning. This chapter explores how economic theory may help in that task. The research moves us beyond measurement of inequality in the distribution of particular attributes to explanation of how those distributions come about. It is important to note, however, that it is an integral part of the indicative approach to evaluating inequality in health capability sets adopted in the thesis (see Chapter 3).

It was argued in Chapter 3 that the Sen framework has similarities with Grossman's (1972) model of the demand for health. Various authors have indeed suggested that the model may be used to analyse the persistence of health inequality, notably Culyer (1976), Maynard (1983), Williams (1984), Muurinen and Le Grand (1985) and Wagstaff (1986a). This appendix examines the suitability of Grossman's model for health-equity analysis and proposes an alternative specification that draws on suggestions by Muurinen (1982a,b) and Williams (1984).

The organization is as follows. The economic approach to modelling health demand, associated with Grossman (1972), is presented in A1.2. The following section

discusses various reasons why the model in its original form is not a suitable basis for considering the most important questions raised in the health inequality debate (A1.3). Section A1.4 then goes on to present a model which is arguably more suited to the problem at hand.

A1.2 Grossman's household production model

Michael Grossman's (1972) model of the demand for health is often considered to be the major theoretical innovation to have emerged from the sub-discipline of health economics (*eg.* Culyer, 1981). Prior to Grossman's study there were no satisfactory explanations of the production of health at the individual level; nor indeed of how important variables such as health status or education (seen as forming part of the exogenously determined "taste matrix" in traditional demand analysis) affected the demand for health care. Clearly, if health economics is to serve a useful purpose in evaluating policy alternatives the effects of shifts in variables other than price and income have to be predicted within a general health-decision framework.

Grossman's model is built on the theory of household production pioneered by Becker (1965), Lancaster (1966) and Muth (1966). In this approach, individuals *produce* fundamental commodities such as recreation, travel or health by combining inputs of market goods and own-time within a full-wealth constraint of wage and non-wage income and the monetary value of non-market time. Health care is one of the inputs into the production of health and hence the demand for that good is essentially derived from the demand for health itself.

Table A1.1 sets out the fundamental equations of Grossman's theoretical structure which may be outlined as follows. The individual derives utility from being in good health (h_t) as well as from a composite of all other fundamental commodities (Z_t). Health is demanded for two reasons: it yields direct increases in utility (*consumption benefits*, such as avoidance of pain and discomfort) and indirect increases, through more healthy time being available for activities such as consumption, working and leisure (*investment benefits*). The amount of healthy time produced is a function of the

Table A1.1: Basic equations of the Grossman model

1. Intertemporal Utility Function	$U = U(\phi_0 H_0, \ldots, \phi_T H_T, Z_0, \ldots, Z_T)$, where $\phi_t H_t = h_t$
2. Health Capital Production Function	$I_t = I_t(M_t, TH_t; E_t)$
3. Other Commodities Production Function	$Z_t = Z_t(X_t, T_t; E_t)$
4. Health Capital Stock	$H_{t+1} - H_t = I_t - \delta H_t$
5. Goods Budget Constraint	$\frac{\Sigma P_t M_t + F_t X_t}{(1+r)^t} = \frac{\Sigma W_t TW_t}{(1+r)^t} + A_0$
6. Time constraint	$TW_t + TH_t + T_t + TL_t = \Omega$
7. Full Wealth Constraint	$\frac{\Sigma P_t M_t + F_t X_t + (TH_t + T_t + TL_t)}{(1+r)^t} = \frac{W_t \Omega}{(1+r)^t} + A_0 = R$

Definitions:

All subscripts refer to time periods

ϕ = the flow of health per unit stock
H = health stock
Z = production of non-health fundamental commodities
h = healthy time
I = gross investment in health
M = market goods input in health capital production (assumed to be simply medical care)
TH = time input in health capital production
E = stock of human capital (an efficiency factor in production) production
X = market goods input in other commodities' production
T = time input in other commodities' production
δ = rate of depreciation in health capital
P = price per unit of goods input M
F = price per unit of goods input
r = rate of interest
W = wage rate
A = non-wage income
TW = time spent working in the market
TL = sick time
Ω = total time available
R = full wealth

individual's health stock (H_t), which he is endowed with at birth and depreciates with age. Additions to this stock can be produced using time and purchased goods and services (*eg.* medical care) as inputs, as well as through investment in human capital (*eg.* education) which is assumed to improve the efficiency of the production process in the non-market sector of the economy. The production of fundamental commodities other than health is also achieved in a similar fashion. Net investment in any particular period is that newly produced minus that which has depreciated from the previous period. The stock of health is, therefore, partially endogenous, leading to the oft cited argument that in Grossman's framework the individual chooses his time of death (*ie.* at time T when $H_t = H_{min}$). Life-time utility is maximized subject to the constraint of the present value of the individual's full wealth, which is a composite of the discounted money and time value of the resources available to the individual over his life-time.

The optimality conditions for the total amount of health newly produced in each period imply that the present value of the marginal cost of producing health equals the present value of marginal benefits. This is shown in the following equation:

$$\frac{\partial h_t}{\partial H_t}\left[W_t + \left(\frac{Uh_t}{\lambda}\right)(1+r)^t\right] = N_{t-1}(r - \tilde{N}_{t-1} + \delta) \qquad (1.1)$$

where $\partial h_t / \partial H_t$ is the marginal product of the stock of health in the production of healthy time; Uh_t is the marginal utility of healthy time; λ is the marginal utility of wealth; N_t is the marginal cost of gross investment ($= P_tM_t + W_tTH_t$); and $\tilde{N}_t$, the percentage rate of change in marginal cost.

The marginal benefits - shown on the left-hand-side of (1.1) - are composed of the two components corresponding to the investment and consumption motives, which in turn lead Grossman to formulate two competing sub-models to be tested. What each of these implies can be seen by dividing both sides of (1.1) by N_{t-1}:

$$\frac{\dfrac{\partial h_t}{\partial H_t} W_t}{N_{t-1}} + \frac{\dfrac{\partial h_t}{\partial H_t}\dfrac{Uh_t}{\lambda}(1+r)^t}{N_{t-1}} = r - \tilde{N}_{t-1} + \delta \qquad (1.2)$$

When the first expression on the LHS is assumed to equal zero the pure consumption

model holds. By the same token, the pure investment model, which forms the basis of Grossman's empirical derivations, is based on the second expression on the LHS - the consumption benefits - being absent.

Grossman goes on to derive the comparative static predictions separately for the two sub-models. One important difference which arises is that in the pure investment model wealth effects are not relevant since an increase in wealth, without analogous changes in the interest rate and the rate of depreciation does not alter the equality between the cost of capital and the rate of return on an investment in health. Despite this result, Grossman opts to base empirical analysis on the pure-investment model arguing that it generates equally powerful predictions from simple analysis and inocuous assumptions [Grossman (1972, p. 39).

The effects of exogenous changes in the theoretical variables under the investment model are summarized in Table A1.2. Although the predicted effect of wealth is zero it was included by Grossman in his empirical analysis in order to evaluate the predictive power of the two sub-models. Wealth was observed to have a positive effect on the demand for health care and a negative effect on the demand for health. These results were argued to be due to the correlation between high income and insalubrious health habits (*eg.* excessive eating or smoking) which lower the stock and flow of health and increase the need for medical treatment.

The wage rate was included in Grossman's demand equations for two reasons: because it determines the value of the marginal product of health and because it enters into the calculation of the total cost of a unit of health capital. The predicted net effect of wages was to increase the demand for both health and medical care. In the demand for health, the wage rate operated to raise the marginal product of the health stock more than it increased the marginal cost of health investment, because time costs were only a portion of the total cost of health capital. This argument was supported by empirical evidence. In the derived demand for medical care, the argument was similar with one important addition - medical care and time devoted to health were assumed to be substitutes. That is, medical care is a less time-intensive method of producing health, *vis-à-vis* other methods such as adequate sleep, exercise and so on. The empirical

Table A1.2: The Grossman model: Predicted and actual signs of the effects of exogenous variables on the demand for health and medical care

	Wealth	Wage Rate	Education	Age
Demand for health				
Predicted sign	0	+	+	-
Observed sign	-	+	+	-
Demand for medical care				
Predicted sign	?	+	-	+
Observed sign	+	-	+	+

Source: Muurinen (1982a, p.14).

results did not show the expected positive wage effect; instead it was negative (although statistically insignificant), a result which Grossman attributes to measurement error.

Rising education levels are postulated to increase the individual's efficiency in producing new health. Less time is required per unit production of health capital and therefore the marginal cost of health capital production falls with marginal benefits held constant. This enables identical increases in health to be produced from fewer inputs of both time and medical care than was hitherto the case. Therefore, more health and less health care are demanded in order to maintain equilibrium. Empirical evidence was found for the the first of these effects but not for the second, which should only hold if the elasticity of the marginal efficiency of health capital is less than unity.

Finally, Grossman also considered the effects of age, assuming that as it increases so does the rate of depreciation on health. As a result, the marginal cost of producing healthy time increases, leading the production of health (and hence demand for health) to fall, in order to maintain equilibrium. Providing that the price elasticity of the demand for health is less than unity, individuals will offset part of the reduction in health capital by increasing gross investment. Since medical care is one of the inputs

in the health investment function the demand for health care will tend to increase. The actual signs revealed by empirical analysis tended to substantiate the theoretical predictions.

A1.3 An appraisal of the Grossman model as a guide to health-equity research

The model of health investment summarized above has undergone various extensions and generated a considerable number of refined specifications and estimations of the demand equations.[1] Generally, this work has retained the underlying structure of the initial model and introduced only slight alterations. In this context, it makes sense to pose the question as to whether the original Grossman framework is indeed suited for studying the health-equity problem, as so many authors have claimed.

An obvious starting point is to recognize that the model was not designed explicitly for analysing socio-economic differentials in health and health care, and much less for answering whether a particular distribution is equitable or not. Grossman's interest lay in understanding the relationship between health status, health care utilization, wage rates, education and age. Why then, one could ask, should the analysis be extended to study questions raised by the health inequity debate? It seems that two factors have inspired suggestions along such lines. Firstly, developments in the human capital literature. Early work in this area was directed at explaining how schooling choices affected labour supply and economic growth, but it was later applied with considerable appeal to the study of interpersonal distribution of income (*cf.* Rosen, 1987; Becker and Tomes, 1978). This development has in all likelihood led health economists to accept that a similar course might be taken by models of the demand for health in the human capital tradition.

By far the more important reason, however, has been the lack of rigorous and persuasive alternatives. Richard Berk (1987) has noted that human capital models are

---

[1] See Grossman (1982) for a succint review up to that date, and Birch (1987: Chapter 1) for a more recent and detailed critique.

"effectively directed at the soft underbellies of other social-science disciplines" (p. 130). As the previous section showed, Grossman's model provides a coherent method of explaining individual behaviour with regard to health and health care. It yields predictions on the effects of independent variables such as the wage rate and education - which may be seen as components of a person's socio-economic position - thereby shedding light on the process by which differentials are generated.[2] Though it may seem a heavy-handed way of drawing rather unremarkable conclusions it has the advantage of being explanatory in character, thus lending itself to empirical verification or falsification to a greater extent than other approaches to the problem. Such a systematic analytical framework may be precisely what has been missing from the type of research reviewed by the Black Report (DHSS, 1980).

Remarkably, however, virtually no empirical analyses of the health-equity relationship have sought to draw on Grossman's economic theory as a guide. It would be too easy to ascribe this state of affairs to communication barriers in the intellectual division of labour. Health economists, after all, have generally not based their own empirical studies of the equity problem on the model. That this is in part due to the unavailability of suitable data goes without question. But it may also reflect a certain uneasiness with regard to the assumptions and mechanisms underlying the theoretical model.

All this goes to suggest that, whilst the Grossman framework reveals a comparative advantage over the largely *ad hoc* conceptualizations employed in health-equity research, attention should be paid to developing those aspects of the model that appear to preclude a satisfactory analysis of key questions raised in the debate. Three distinct disadvantages are suggested here. The aim is not to provide a systematic critique of Grossman's work, but rather to bring to the fore possible alterations to the basic model which render it more suitable for the task at hand.

The first criticism has already been alluded to. It is that the pure investment

---

[2] See Muurinen and Le Grand (1985) and Wagstaff (1986a) for informal expositions of how the framework provides an account of the causal mechanisms underlying health differences.

model is not particularly informative of the relationship between command over resources and health and health care, since individuals' choices about health investment at each moment in time are independent of their initial and current stocks of wealth (Muurinen, 1982a; Dardadoni and Wagstaff, 1987). Obviously, one could simply base empirical analysis on the *pure consumption* model, but this would mean abstracting from a major insight of the capabilities framework: that people combine the characteristics of goods fundamentally to produce human functionings (the investment benefits of health). If for any individual these functionings depend on inputs of own and goods characteristics, which are observable, then one is able to make equity comparisons. The direct utility consequences, on the other hand, are by no means straightforward, depending on a variety of emotional responses. Moreover, the *pure investment* model has dominated discussion of the Grossman framework, due to its clarity as a teaching aid and ability to yield sharp predictions from simple assumptions.

There have been attempts in the literature to overcome the wealth neutrality inherent in the investment model. Dardanoni and Wagstaff (1987) suggest introducing uncertainty. They show that if individuals display decreasing absolute risk aversion, the wealthier among them will invest more in health than those with lower initial stocks of financial assets. Essentially, they do so because their larger stocks of initial wealth capital provide an incentive to invest in the riskier asset - health. The argument depends crucially on considering simply the uncertainty surrounding the flow from the health stock (ignoring, therefore, uncertainty on the return to financial capital as well as that surrounding the effectiveness of health inputs) and on the assumption of decreasing absolute risk aversion.

Dardanoni and Wagstaff's analysis raises the spectre of complete certainty which has long been argued to be a major drawback in the Grossman model (*eg.* Dowie, 1975), even though Grossman himself acknowledged that this assumption should be dropped in future work.[3] However, in order to resolve the wealth neutrality problem one need not introduce uncertainty. Another possibility (accepted by Dardanoni and

---

[3] Several authors later took up the challenge, notably Phelps (1973) and Cropper (1977) on the uncertainty regarding future illness, and Dowie (1975) and Keeler *et al* (1977) on that concerning the impact of health inputs.

Wagstaff) is to assume imperfect markets, so that the rate of depreciation and/or the price of health investments are decreasing functions of health investment. The marginal cost of health investment would then be decreasing in wealth, so that the better-off would *ceteris paribus* choose higher values of health investment.

There is, however a more straightforward approach which allows wealth to affect health decisions in a generalized investment framework while retaining the complete certainty construct. Muurinen (1982b) views the investment and consumption benefits of health as explicitly complementary, rather than as substitutes as Grossman's analysis seems to imply. Analytically, the separation of benefit types is retained in the alternative specification but they are treated as being produced from the same addition to the stock of wealth. Intuitively, this alteration seems correct since health is demanded simultaneously for its utility consequences (enjoying good health) and for what it allows in terms of functional capacity (better performance of social and economic tasks). Moreover, Muurinen's specification implies a richer interpretation of the investment benefits which is closer to Sen's capabilities analysis and contrasts sharply with the narrow view implied by the human capital model, where investment benefits are seen simply as increases in healthy time available for activities valued at the exogenous wage rate. The investment benefits of health are now "seen as increased capacity to perform one's tasks in the role of a producer (either inside the home or in the labour force), and the value of this as dependent on the characteristics of this participation, the individual in question, and his or her circumstances more generally" (Muurinen, 1982b, p. 7).

A second drawback of Grossman's formulation is its handling of variables such as education, and so-called environmental variables like housing or work conditions. Consider first the case of education. In the Grossman model better education simply enhances the technical efficiency of household production. This implies that its role is dependent on the use of that specific approach to analysing health behaviour and, that the precise nature of the education-health relationship remains relatively unspecified. It overlooks, for example, the role of education as improving allocative efficiency in terms of choices between production processes. In the inequality in health debate it is this aspect which has attracted the greater attention. The better educated are seen as having greater awareness of the adverse health effects of particular activities and hence

choose more suitable health promoting behaviour.

More recent studies have modelled the relationship in a different light, explicitly taking into account the allocative benefits of education. Muurinen (1982b), for example, defines the rate of depreciation on health as a function of the use of related variables. Education is seen as one such variable which redirects the choice of different production processes towards a life-style that is less use-intensive in terms of health. Whilst this approach has some appeal, it sketches over the precise relationship between health, education and other important variables such as work environment (Kemna, 1985) or cigarette consumption (Farrel and Fuchs, 1982). Education is not simply one amongst many factors which affect health depreciation, but itself largely determines consumption of those other goods and activities. Furthermore, it is likely to affect an individual's command over resources which in turn affects health decisions. Thus a more refined framework would view education as an investment activity involving opportunity costs in return for future benefits, very much in the same way as health.

Muurinen's use-related depreciation concept does, however, seem pertinent for handling variables such as diet, smoking, exercise or even housing or work environment. The latter are commonly correlated with health and health care as evidence of inequality, without a rigorous explanation of the mechanisms by which differentials arise (*cf.*, *eg.*, the research reviewed by Whitehead, 1987). Viewing their effect as operating through use related depreciation on health provides such a rationale which is also altogether more appealing than Grossman's approach. He suggests (Grossman, 1972, Chapter 6) that these variables might enter the model as health production function inputs, in addition to medical care and time. This would lead, however, to highly aggregated functions and make the analysis of activities undertaken specifically to produce health (*eg.* medical care) more difficult. This is unfortunate given that public pronouncements on equity usually stress access to health care rather than other health-producing goods. It would also raise theoretical problems. On the one hand, the model would have to consider joint production for which theory has so far not produced adequate answers; and on the other, the assumption of positive health investment in all periods could not be sustained, unless benefical inputs were always greater than those that are harmful to health.

This brings us to the final major drawback of the Grossman model as a basis for undertaking explanatory health-equity research: its location within a household production framework and treatment of the wage rate as the value of time. Household production theory has drawn a number of significant criticisms in the literature, among them the unsatisfactory treatment of joint production, the reliance on constant returns to scale in the production functions, the disregard for transaction and adjustment costs, and the unfortunate feature of key variables being typically, or even in principle, unobservable (see, *eg.* Berk, 1987). Each of these is applicable to Grossman's specification. For instance, the additive nature of the time constraint ignores joint production in the use of household time, implying that sick time is totally unproductive, so that it may not even be used for utilizing health care. Yet health economists have generally been want to cast doubt on the household production construct itself. The vast majority of 'post-Grossman' demand for health models have retained the underlying structure and simply introduced slight alterations, the effects of which are then tested against data.

On a practical level, Grossman's use of the wage rate as a universal measure of the marginal valuation of time creates various difficulties. In particular, it may overestimate the opportunity cost of work time lost due to illness/use of medical care, if institutional arrangements exist whereby employees suffer no loss in earnings during such periods. As Le Grand (1982) has argued, poorer income groups, who are often paid on an hourly rather than salary basis, are the more likely to experience reductions in income, and thus face higher opportunity costs of time. If this is the case, applying each persons wage rate as the measure of time would lead to biases in measurement of access cost and indirectly of other variables included in the model.

The unsuitability of Grossman's approach is further compounded once it is realized that adequate empirical estimates of the wage rate are not readily available for the whole population. All those outside the labour force (*eg.* the unemployed, housewives, and pensioners) are effectively excluded from empirical analysis, even though there is evidence that they may experience higher rates of sickness and consume health care over-proportionately (*eg.* Andersen et al, 1975). Since these are also groups which generally reveal lower command over resources, it makes equity analysis based

on the Grossman model rather uninformative. Calculation of a reservation wage would go some way to resolving these problems but it is likely to involve new data collection and prove cumbersome (see, *eg.*, Coffey, 1983). Moreover, most health data sources do not include, even for those in work, an adequate estimate of the wage rate, so that unreliable proxies have to be used.[4] Given these drawbacks, the use of a model based upon household production theory can only be defended if there are strictly no suitable alternatives.

In the following section an economic model of health behaviour is presented that may be seen as an extension of Grossman's investment framework, but which dispenses with the more problematic features identified above. The new specification draws on work by Muurinen (1982a, 1982b) and in particular on a model of health behaviour sketched informally by Williams (1988a). Since Williams' analysis reveals how many of the drawbacks in the Grossman formulation may be improved, it is described here as a prelude to mathematical specification. Its point of departure is that individuals are considered to hold, at any point in time, three distinct stocks of capital: namely health, wealth and "wisdom", conforming to the proverb "early to bed and early to rise, makes a man healthy, wealthy and wise." Health behaviour, therefore, does not simply imply optimizing the flow from the health stock, but simultaneous optimization of the services arising from the three individual assets.

Figure A1.1 provides a schematic view of the Williams framework.[5] The three stocks are capable of being augmented by suitable investment processes (row 1) and are each subject to time and use-related depreciation (rows 2 and 3). Health capital, for example, may be added to by use of health care or activities such as adequate sleep, exercise or shelter. It depreciates in time due to the natural ageing process (time-related depreciation) and the effects of individuals undertaking harmful activities, such as smoking, excessive drinking and working in a stressful environment (use-related depreciation). The wealth stock is augmented by financial (net saving) and

---

[4] Birch (1987) for example, uses occupational status as a proxy for the wage rate.

[5] In Williams' (1988a) article a slightly more elaborate diagram is presented, highlighting the various interrelationships and possibilities for substitution between the stocks.

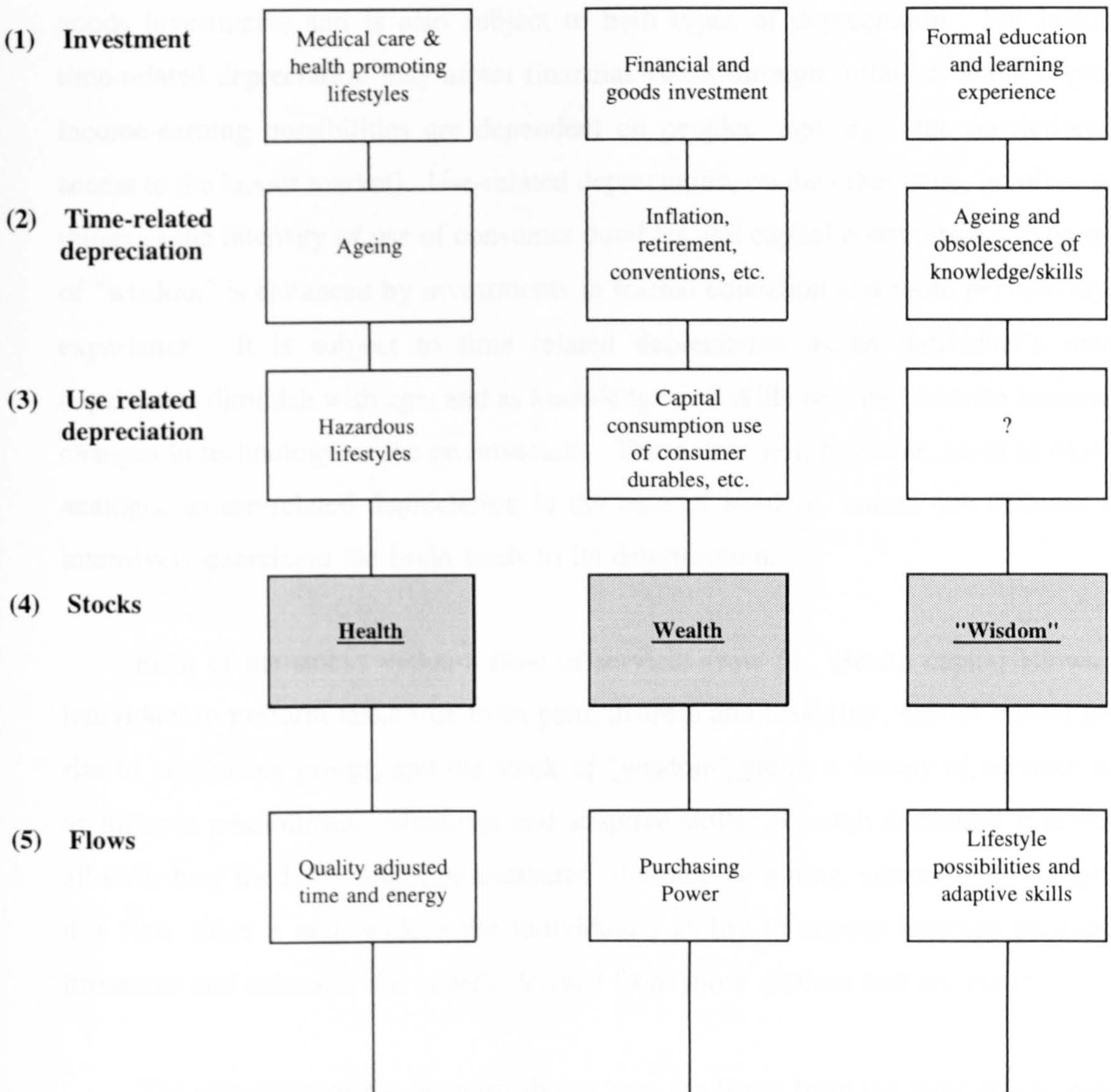


Figure A1.1: Williams' schematic view of health behaviour

goods investments and is also subject to both types of depreciation. For instance, time-related depreciation may affect financial assets through inflation, while expected income-earning possibilities are dependent on peoples' age (*eg.* age restrictions on access to the labour market). Use-related depreciation, on the other hand, involves such things as the intensity of use of consumer durables and capital consumption. The stock of "wisdom" is enhanced by investments in formal education and more pervasively by experience. It is subject to time related depreciation as an individual's mental capabilities diminish with age, and as knowledge and skills become obsolete because of changes in technology or the environment. There does not, however, seem to exist an analogue to use-related depreciation in the case of wisdom, unless one believes that intensively exercising the brain leads to its deterioration.

Each of the stocks yields a flow of services (row 5). Health capital allows the individual to perform tasks free from pain, distress and disability; wealth capital gives rise to purchasing power; and the stock of "wisdom" yields a variety of services such as lifestyle possibilities, valuations and adaptive skills. Though at present it is not at all clear how the latter might be measured, it would be wrong, conceptually, to ignore this flow since it both widens the individual's ability to choose between production processes and enhances the benefit derived from those choices that are made.

The remainder of the diagram shows how the flows from the three stocks enable individuals to engage in various activities, which then have consequences for the level of health capital they are able to achieve in the following period. For instance, quality adjusted time and energy allows the individual to participate in the labour market, household production and consumption activities, which in themselves may add or subtract from her health. But the other flows permit similar activities so that the individual may wish to trade-off investments in a particular stock with those in another. In this sense, health behaviour is explained by the simultaneous management of the three individual assets.

In reality, this task is complicated by at least two factors. First of all, there is likely to be a time lag between activities and their effects on the capital stocks; and

secondly, all decisions are subject to risk and uncertainty.[6] Nevertheless, there are important lessons to be derived from the Williams framework even within a complete certainty framework. From the point of view of equity it suggests that those with higher stocks of knowledge and wealth are able to offset deleterious effects upon their health stocks to a greater extent than other individuals. Furthermore, if the stocks are interchangeable, then it is reasonable to assume that all individuals will have *some* possibility to minimize the effects of price and opportunity constraints, so that the levels of their initial capital assets are the only constraints that are strictly beyond control. Since inequality in the initial stocks of health and "wisdom" are likely to be less than in the inherited stock of wealth, it would appear that initial holdings of wealth may have a determining role in the individual's life-time profile of health capital. The model developed below aims to explore these suppositions.

A1.4 An Alternative specification

This section presents a formal model that could be used as a guide to empirical analysis. The model is set up with a specific research topic in mind - the relationship between socio-economic inequalities, health and health care. It aims to identify in a rigorous fashion the mechanisms by which these categories are related and, adjacently, to provide a bridge from Sen's normative analysis of equity to positive verification. Therefore, the framework should also provide guidance as to how equal opportunities for improvement of the functioning *health* might be achieved.

At the same time, the model constitutes a development on previous economic specifications of health-related behaviour. In essence, all such models - including Grossman's and others that have drawn on the conceptual tools of the demand for health - are simplified versions of Williams' informal framework. Muurinen's (1982a,b) is perhaps the more general specification of the relationships suggested by Williams.[7] The

---

[6] Zweifel (1989) has drawn particular attention to these aspects showing at the same time that there is likely to be positive correlation of stochastic shocks among all of the three assets. This results in a reinforcement of exposure to risk and may explain the prevalence of social - as opposed to private - insurance in the health care market.

[7] Its essential features were also later adopted by Wagstaff (1985).

present model goes a step further by introducing an independent stock of "wisdom", simplified to denote *education capital.* Whilst this refinement does not produce a fundamentally different equilibrium condition for health capital it permits a clearer understanding of the relationship between wealth and knowledge and of their impact on health and health care. Other alterations to existing models are minor and essentially determined by an attempt to facilitate the mathematical derivations.

A strict adherence to Sen's capabilities framework would suggest that individuals should be represented as maximizing *functionings* such as 'healthy time'. In order not to depart from previous specifications of the demand for health, however, individuals in this model are assumed to be life-time utility maximizers. This assumption does not signify that the Sen and Grossman approaches cannot be combined, it merely states that the link from functionings to utility is viewed as unproblematic.

A1.4.1. Assumptions and equilibrium conditions

The representative individual is assumed to derive utility at each point in time according to

$$U = U[h(t), Q(t)] \tag{1.3}$$

where $h(t)$ is the functioning *healthy time*, produced from the stock of health, $H(t)$, by the following relationship

$$h(t) = \phi[H(t)], \qquad \phi' > 0, \quad \phi'' < 0. \tag{1.4}$$

$Q(t)$ is a composite functioning, that is, an aggregate of all other functionings from which utility is derived. The first partials of utility with respect to $h(t)$ and $Q(t)$ are assumed to be positive (*ie.* $\partial U/\partial h(t) > 0$ and $\partial U/\partial Q(t) > 0$).

At time 0 the individual inherits an initial stock of health capital, H_0, which thereafter evolves according to the relationship:

$$\dot{H}(t) = I(t) - \delta[t, k(t), X(t)]\, H(t) \tag{1.5}$$

This states that net investment in health, $\dot{H}(t)$, is equal to gross investment, $I(t)$, less $\delta[.]$,

the rate at which health capital depreciates. This depreciation is an explicit function of age *t*, the services of the individual's stock of education *k(t)* - defined below - and the vector *X(t)*, representing all other relevant variables such as work environment. The latter are postulated to act as lifestyle selection factors so that those who work in safe environments experience lower rates of depreciation and hence choose healthier lifestyles. Conversely, individuals who smoke, eat excessively or hold hazardous jobs reveal a greater deterioration of their health stock.

Gross investment in health, *I(t)*, is assumed in this model to be produced solely from activities which are undertaken primarily for health reasons. Institutional medical care is the obvious example, but other things such as self-care could just as well be included in this category. All those other activities which give rise to improvements in health enter the model via a negative effect on the rate of depreciation, thereby appreciating the existing stock of health capital. Further, it is postulated that the effects of *I(t)* on health are always positive. Whilst these are somewhat restrictive assumptions they are justified by the fact that the main object of the model is to analyse how the distribution of opportunities in achieving good health is conditioned by investment in the good which policy makers and the public in general associate with better health, namely health care. Furthermore, people make such investments because they generally believe it will do them some good.

The services of the education stock are produced according to the following relationship

$$k(t) = \mu [E(t)], \qquad \mu' > 0, \quad \mu'' < 0. \tag{1.6}$$

where *E(t)* is the stock of education, and *k(t)* may be seen as the lifetime possibilities and adaptive skills enjoyed by the individual. The education stock, itself, evolves from H_0 at time 0 according to:

$$\dot{E}(t) = L(t) - d\,[t]\; E(t) \tag{1.7}$$

where, *L(t)* is gross investment in education and *d[.]*, the rate of depreciation on education. Thus, the forces determining the path of the education stock are similar to those postulated for the stock of health, with the exception that use-related depreciation

is not important in this case.

The individual is also assumed to inherit an initial stock of wealth denoted by A_0. Thereafter it evolves according to:

$$\begin{aligned} \dot{A}(t) = \; & rA(t) + Y[h(t), I(t), k(t), Z(t)] \\ & - [P_Q(t)\,Q(t) + P_I(t)\,I(t) + P_L(t)\,L(t)] \end{aligned} \tag{1.8}$$

where r is a time invariant rate of interest; Y earned income, which is a function of the flow of health $h(t)$, use of health-specific investment goods $I(t)$ (given that income is lost in order to undertake such activities), the services of the education stock $k(t)$, and other relevant variables, $Z(t)$. $P_Q(t)$, $P_I(t)$ and $P_L(t)$ are the exogenous prices of $Q(t)$, $I(t)$ and $L(t)$ respectively.

Given these expressions the individual's objective is to maximize discounted lifetime utility

$$Max \quad \int_0^T e^{-rt}\, U[\phi(H(t)), Q(t)]\, dt \tag{1.9}$$

subject to (1.5), (1.7) and (1.8) and the required boundary and negativity conditions

$$\begin{aligned} A(0) &= A_0, \\ E(0) &= E_0, \\ H(0) &= H_0, \\ A(T) &\geq 0, \\ E(T) &\geq 0, \\ H(T) &\geq \bar{H}(t) \end{aligned} \tag{1.10}$$

where $\bar{H}(t)$ is the level of the health stock at which death occurs. In equation (1.9), e^{-rt} is a time discounting factor where the rate of time preference is assumed to be equal to the rate of interest, r. It is also assumed that $H(t)$, $Q(t)$, $I(t)$, $P_I(t)$ and $P_Q(t)$ are

non-negative for all t.

The problem has been set up in a form which permits its solution and analysis through the maximum principle in optimal control theory. Thus it is assumed that individuals are able to manipulate $I(t)$, $Q(t)$ and $L(t)$ - the *control* variables - choosing optimal paths for the stocks $H(t)$, $A(t)$ and $E(t)$ - the *state* variables.

Maximization of the objective function subject to all the constraints gives the following *Hamiltonian*:

$$\begin{aligned}\Lambda = {} & e^{-rt}\, U[\phi(H(t)), Q(t)] + \lambda_H(t)\,\{I(t) - \delta[t, \mu(E(t)), X(t)]\, H(t)\} \\ & + \lambda_E(t)\,\{L(t) - d[t]\, E(t)\} + \lambda_A(t)\,\{rA(t) + Y[\phi(H(t)), I(t), \mu E(t), Z(t)] \\ & - [P_Q(t)\, Q(t) + P_I(t)\, I(t) + P_L(t)\, L(t)]\}\end{aligned} \tag{1.11}$$

where $\lambda_H(t)$, $\lambda_E(t)$ and $\lambda_A(t)$ are the costate variables of $H(t)$, $E(t)$ and $A(t)$ respectively.

A maximum requires that the folowing conditions hold:

$$(i) \qquad \frac{\partial \Lambda}{\partial I(t)} = \lambda_H(t) + \lambda_A(t)\left[\frac{\partial Y}{\partial I(t)} - P_I(t)\right] = 0 \tag{1.12}$$

$$(ii) \qquad \frac{\partial \Lambda}{\partial Q(t)} = e^{-rt}\,\frac{\partial U}{\partial Q(t)} - \lambda_A(t)\, P_Q(t) = 0 \tag{1.13}$$

$$(iii) \qquad \frac{\partial \Lambda}{\partial L(t)} = \lambda_E(t) - \lambda_A(t)\, P_L(t) = 0 \tag{1.14}$$

$$\begin{aligned}(iv) \qquad \frac{\partial \Lambda}{\partial H(t)} &= -\dot{\lambda}_H(t) \\ &= e^{-rt}\left(\frac{\partial U}{\partial \phi}\,\frac{\partial \phi}{\partial H(t)}\right) - \lambda_H(t)\,\delta[t, \mu(E(t)), X(t)] \\ &\quad + \lambda_A(t)\left(\frac{\partial Y}{\partial \phi}\,\frac{\partial \phi}{\partial H(t)}\right)\end{aligned} \tag{1.15}$$

(v) $$\frac{\partial \Lambda}{\partial E(t)} = -\dot{\lambda}_E(t)$$

$$= \lambda_H(t)\, H(t) \left(\frac{\partial \delta}{\partial \mu} \frac{\partial \mu}{\partial E(t)} \right) - \lambda_E(t)\, d[t] + \lambda_A(t) \left(\frac{\partial Y}{\partial \mu} \frac{\partial \mu}{\partial E(t)} \right) \tag{1.16}$$

(vi) $$\frac{\partial \Lambda}{\partial A(t)} = -\dot{\lambda}_A(t)$$

$$= r\, \lambda_A(t) \tag{1.17}$$

plus (1.5), (1.7), (1.8) and the boundary conditions.

In order to derive the optimal health investment strategy proceed as follows. Given that (1.17) is a first-order differential equation it follows that

$$\lambda_A(t) = \lambda_A(0)\, e^{-rt}.$$

From (1.12)

$$\lambda_H(t) = \left[\lambda_A(t) \left(\frac{\partial Y(t)}{\partial I(t)} - P_I(t) \right) \right] \tag{1.19}$$

Differentiating this expression with respect to time and defining $C(t) = [P_I(t) - \partial Y/\partial I(t)]$ gives

$$\dot{\lambda}_H(t) = \dot{\lambda}_A(t)\, C(t) + \lambda_A(t)\, \dot{C}(t) \tag{1.20}$$

Finally, substituting (1.19) and (1.20) into (1.15) and taking into account (1.17) and (1.18) one arrives at

$$\left[\frac{\frac{\partial U}{\partial \phi} \frac{\partial \phi}{\partial H(t)}}{\lambda_A(0)} \right] + \left[\frac{\partial Y}{\partial \phi} \frac{\partial \phi}{\partial H(t)} \right] = \left\{ \delta\,[t,\, \mu(E(t)),\, X(t)] + r - \frac{\dot{C}(t)}{C(t)} \right\} C(t). \tag{1.21}$$

Equation (1.21) is the equilibrium condition for health capital, stating that the marginal consumption and benefits of health must equal the user cost of health capital. The terms in this equation have the following economic interpretation:

$\left[\frac{\frac{\partial U}{\partial \phi}\frac{\partial \phi}{\partial H(t)}}{\lambda_A(0)}\right]$ is the marginal consumption benefit of health;

$\left[\frac{\partial Y}{\partial \phi}\frac{\partial \phi}{\partial H(t)}\right]$ is the marginal production benefit of health;

$\lambda_A(0)$ is the marginal utility of initial wealth; and,

$C(t) = \left[P_I(t) - \frac{\partial Y}{\partial I(t)}\right]$ is the effective marginal cost of new health investment (*ie.* money price and opportunity cost, in the form of lost income, of investment in goods aquired primarily for health reasons).

The optimal wealth and education investment strategies are not analysed here, given that the model seeks only to describe health behaviour. Nevertheless, it is worth looking a little closer at condition (1.17), which provides the basis for the optimal allocation of lifetime wealth. Its solution, (1.18), suggests that wealth should be allocated so as to maintain the compounded value of the marginal utility of wealth constant over the lifecycle and equal to the number $\lambda(0)$. Under conditions of certainty, the individual regards the level of assets from previous decisions as given and effectively remaximizes the remaining lifetime utility at each point in time, thereby continuing on the same path as that selected at time 0 (see Heckman, 1976). Therefore, $\lambda(0)$ may be regarded as indicating the individuals permanent real wealth.

A1.4.2. Predictions of the model

From the point of view of the health inequality debate there are three independent variables whose changes are of particular interest: namely wealth, education services and work environment. The procedure generally followed in models of health investment that adopt a dynamic optimization approach is to consider what effectively are parametric shifts in the lifecycle profile of the exogenous variable in question (see Cropper, 1977; Muurinen, 1982b, Wagstaff, 1985). Whilst this has the effect of abstracting from the essential dynamics of the relationships (see Forster, 1989) it permits the deduction of unambiguous predictions.

The mathematical derivation of the predictions is not undertaken in this study. However, it should be noted that despite the introduction of the independent stock of education as a constraint in the model, the equilibrium condition for health capital (equation 1.21) is virtually the same as that derived by Muurinen (1982a,b). In both models the education variable influences the optimal health investment strategy through its impact on the depreciation of the health stock. If one assumes that the services of the education stock in the present model are positive at all points in time, then their effect is analogous to Muurinen's education variable, the only difference arising in empirical analysis. Hence, a more refined measure of what education does for people (possibly taking into account length of time since finishing formal education or the degree to which the work setting provides opportunities for enhancing the education stock), rather than simply "years of schooling", is warranted by the model suggested here. The wealth and work environment variables have an identical interpretation to that suggested in the Muurinen model.

Consequently, it seems valid to infer Muurinen's results for the present framework. With regard to the three exogenous variables of interest, their impact on the demand for health and that for health-specific investment goods is as follows:

(i) An increase in *wealth* increases the demand for health. This result is derived on the assumption that the initial stock of wealth is - in a complete certainty framework - a suitable indication of the individual's permanent real wealth. The size of the wealth

elasticity will depend on the composition of health benefits. The impact on the demand for health-specific investment goods (*eg.* health care) is also predicted to be positive.

(ii) Larger *education* benefits arising from that capital stock also increase the demand for health, but reduce the demand for activities undertaken for health reasons.

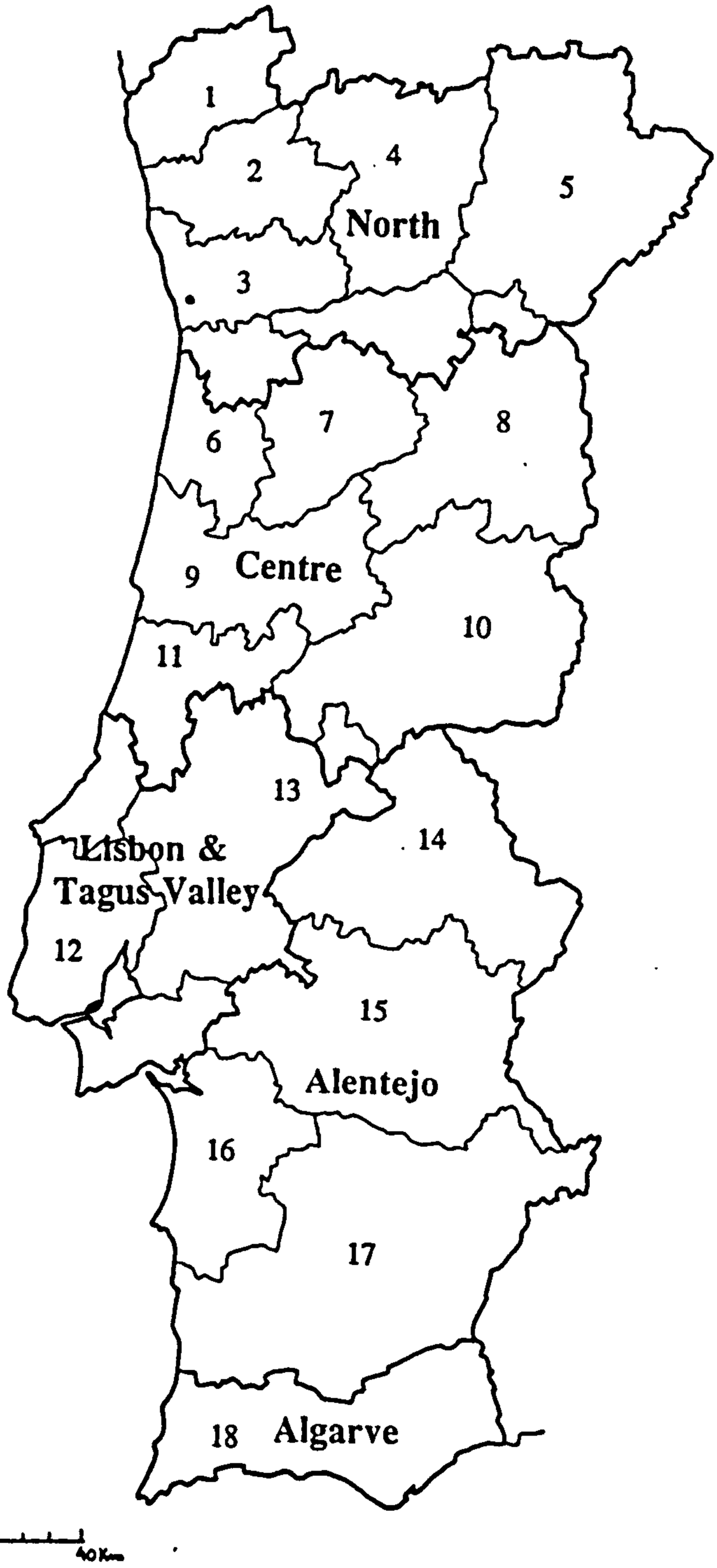
(iii) The impact of *work environment* is identical to that for education given that it enters the model as part of the X profile, similarly postulated to affect the rate of depreciation on health. Thus, those who work in safe environments demand more health and less health care goods and activities than those who work in hazardous conditions.

A1.5 Conclusion

Clearly, further conceptual groundwork is required before the alternative model can be used as a suitable guide to explaining unequal distributions in the domain of health. However, the exploratory analysis presented above does seem a promising line of research for this purpose. Demand for health models have obvious similarities with the normative approach adopted in Chapter 3, which it was argued presents the most suitable means currently available for understanding health-inequity. They are potentially useful in organizing and analysing hypotheses on differences in health states and health care utilization. The particular structure set out in A1.4 may be seen to remedy weaknesses in Grossman's specification which make it unattractive when studying the health-inequity issue. Ultimately, however, the predictions of the model must be tested against suitable data to examine the extent to which they are consistent with the evidence. Empirical work on explaining unequal levels of health status and use of services would indeed appear to be the most important future complement to the evidence on the extent of inequity produced in this thesis.

Appendix 2

Geographical Delimitations in Portugal



Districts

1. V. Castelo
2. Braga
3. Porto
4. Vila Real
5. Bragança
6. Aveiro
7. Viseu
8. Guarda
9. Coimbra
10. C. Branco
11. Leiria
12. Lisboa
13. Santarém
14. Portalegre
15. Évora
16. Setúbal
17. Beja
18. Faro

Appendix 3

Summary of Indices used in the Thesis

ATTRIBUTE	MEASURE	NOTATION	DEFINITION	SIGN PROPERTIES
Health	Illness concentration index	C_m	$C_m = 1 - 2\int_0^1 G_m(p)\,dp$	-1, 0, +1
	Generalized illness concentration index	$C_m(\delta)$	$G_m(\delta) = 1-\delta(\delta-1)\int_0^1 (1-p)^{\delta-2}\, G_m(p)\,dp$	-1, 0, +1
Health Care Finance	Kakwani progressivity index	π_K	$\pi_K = 2\int_0^1 [L_y(p) - G_{hf}(p)]\,dp$ $= C_{hf} - Z_y$	-1-Z_y, 0, 1- Z_y
	Suits progressivity index	π_S	$\pi_S = 2\int_0^1 [q - G_{hf}(q)]\,dq$	-1, 0, +1
	Generalized Kakwani progressivity index	$\pi_K(\delta)$	$\pi_K(\delta) = \delta(\delta-1)\int_0^1 (1-p)^{\delta-2}\,[L_y(p) - G_{hf}(p)]\,dp$ $= C_{hf}(\delta) - Z_y(\delta)$	-1-$Z_y(\delta)$, 0, 1- $Z_y(\delta)$

	Generalized Suits progressivity index	$\pi_S(\delta)$	$\pi_S(\delta) = \delta(\delta-1)\int_0^1 (1-q)^{\delta-2}[q - G_{hc}(q)]dq$	-1, 0, +1
Health Care Delivery	Le Grand-type Kakwani index of vertical inequity	$I_{K\text{-}LG}$	$I_{K\text{-}LG} = 2\int_0^1 [G_m(p) - G_{hc}(p)]dp$ $= C_{hc} - C_m$	$1\text{-}C_m$, 0, $C_m\text{-}1$
	Le Grand-type Suits index of vertical inequity	$I_{S\text{-}LG}$	$I_{S\text{-}LG} = 2\int_0^1 [q^* - G_{hc}(q^*)]dq^*$	-1, 0, +1
	Wagstaff, van Doorslaer and Paci index of vertical inequity	I_{WVP}	$I_{WVP} = 1 - 2\int_0^1 G^*_{hc}(p)dp$	-1, 0, +1
	Generalized Le Grand-type Kakwani index of vertical inequity	$I_{K\text{-}LG}(\delta)$	$I_{K\text{-}LG}(\delta) = \delta(\delta-1)\int_0^1 (1-p)^{\delta-2}[G_m(p) - G_{hc}(p)]dp$ $= C_{hc}(\delta) - C_m(\delta)$	$1\text{-}C_m(\delta)$, 0, $C_m\text{-}1(\delta)$

	Generalized Le Grand-type Suits index of vertical inequity	$I_{S\text{-}LG}(\delta)$	$I_{S\text{-}LG}(\delta) = \delta(\delta-1)\int_0^1 (1-q^*)^{\delta-2}\,[q^* - G_{hc}(q^*)]\,dq^*$	-1, 0, +1
	Generalized WVP index of vertical inequity	$I_{WVP}(\delta)$	$I_{WVP}(\delta) = 1-\delta(\delta-1)\int_0^1 (1-p)^{\delta-2}\,G^*_{hc}(p)\,dp$	-1, 0, +1
	Atkinson/Plotnick-type 'partial symmetry' indices of horizontal inequity	HI_k	$HI_k = \dfrac{Z_{hc/k} - C^+_{hc/k}}{2\,Z_{hc/k}}$	0, +1
Health Social Welfare	*Health social welfare index*	W_κ	$W_\kappa = \mu_h(1 - \kappa C_h)\,, \qquad \kappa \geq 0$	Measured on the health scale
	Index of health social welfare by income groups	$W(p_1, p_2)$	$W(p_1,p_2) = \dfrac{1}{(p_2 - p_1)}\,[p_2 W(p_2) - p_1 W(p_1)]$	Measured on the health scale

Notation:

$G_m(p)$ = illness concentration curve
p = cumulative proportion of income units
δ = distributional judgement parameter
$L_y(p)$ = Lorenz curve for income
$G_{hf}(p)$ = health care finance concentration curve
C_{hf} = health care finance concentration index
Z_y = Gini index for income
q = cumulative proportion of income
$G_{hf}(q)$ = relative health care finance concentration curve
$C_{hf}(\delta)$ = generalized health care finance concentration index
$Z_y(\delta)$ = generalized Gini index for income
$G_{hc}(p)$ = health care utilization concentration curve
C_{hc} = health care utilization concentration index
q^* = cumulative proportion of illness experienced by units ranked by income
$G_{hc}(q^*)$ = relative health care utilization concentration curve
$G^*_{hc}(p)$ = standardized health care utilization concentration curve
$C_{hc}(\delta)$ = generalized health care utilization index
$Z_{hc/k}$ = Gini index for health care utilization in the kth income partition
$C^+_{hc/k}$ = concentration curve for health care utilization with units arranged in order of illness level
μ_h = mean level of health status
κ = income related health deprivation parameter
C_h = health concentration index
p_1 = poorest p_1 decimal fraction of the population
p_2 = poorest p_2 decimal fraction of the population
$W(p_1)$ = average level of health welfare enjoyed by poorest p_1 decimal fraction of the population
$W(p_2)$ = average level of health welfare enjoyed by poorest p_2 decimal fraction of the population

Note: The new indices of health domain inequity and health social welfare suggested in this study are shown in *italic*.

Appendix 4

Inequity in Health: Further Results

Table A4.1: Distribution of morbidity by equivalent income quintile
Actual distribution

Indicator	Approximate Income Quintiles					
	I	II	III	IV	V	Total
Beddays	24.2%	36.9%	16.8%	13.0%	9.1%	100.0%
Off-work days	24.7%	23.8%	24.2%	15.8%	11.5%	100.0%
Restricted activity days	24.2%	30.8%	18.2%	15.3%	11.5%	100.0%
Bedridden	24.3%	32.3%	17.8%	14.6%	11.1%	100.0%
Offwork	23.6%	23.3%	22.5%	17.4%	13.2%	100.0%
Restricted activity	23.6%	27.2%	19.1%	16.4%	13.7%	100.0%
Illness	17.4%	30.5%	18.6%	18.9%	14.6%	100.0%
Limiting Chronic	23.6%	27.9%	19.0%	16.9%	12.5%	100.0%
Acute	20.4%	19.8%	22.0%	19.1%	18.7%	100.0%
Chronic	19.6%	31.3%	18.8%	16.8%	13.5%	100.0%
Circulatory	18.7%	35.3%	17.4%	16.0%	12.5%	100.0%
Respiratory	18.2%	22.3%	22.5%	19.5%	17.5%	100.0%
Digestive	19.9%	23.7%	20.5%	19.1%	16.8%	100.0%
Muscoskeletal	22.5%	31.1%	17.9%	16.4%	12.1%	100.0%
Injury and Poisoning	23.2%	21.6%	22.6%	19.7%	12.9%	100.0%
Acute respiratory	18.4%	19.4%	23.6%	20.0%	18.6%	100.0%
Chronic respiratory	16.2%	31.4%	20.2%	17.4%	14.9%	100.0%
High blood pressure	19.7%	30.3%	17.4%	17.7%	14.8%	100.0%
Population	19.9%	18.8%	20.9%	20.4%	20.0%	100.0%
Average age (years)	35.4	49.6	34.6	35.9	37.2	38.3
Sex distribution M	47.9%	44.0%	49.5%	48.1%	48.6%	47.7%
F	52.1%	56.0%	50.5%	51.9%	51.4%	52.3%

Table A4.2: Distribution of morbidity by equivalent income quintile
Age and sex standardized

Indicator	Approximate Income Quintiles					
	I	II	III	IV	V	Total
Beddays	28.4%	25.5%	20.9%	14.7%	10.6%	100.0%
Off-work days	25.3%	21.0%	24.5%	17.0%	12.3%	100.0%
Restricted activity days	25.5%	21.2%	21.3%	17.3%	14.7%	100.0%
Bedridden	27.1%	24.3%	20.9%	15.6%	12.1%	100.0%
Offwork	23.3%	20.1%	22.7%	19.3%	14.5%	100.0%
Restricted activity	24.1%	20.5%	21.5%	17.9%	16.0%	100.0%
Illness	18.5%	20.9%	21.5%	21.5%	17.6%	100.0%
Limiting Chronic	25.3%	22.1%	22.0%	17.7%	12.9%	100.0%
Acute	20.9%	18.6%	22.2%	18.9%	19.4%	100.0%
Chronic	21.7%	22.5%	21.9%	18.7%	15.2%	100.0%
Circulatory	21.8%	22.3%	22.0%	18.7%	15.2%	100.0%
Respiratory	18.9%	20.6%	19.4%	21.9%	20.3%	100.0%
Digestive	21.2%	19.4%	21.9%	20.3%	17.2%	100.0%
Muscoskeletal	24.3%	23.2%	21.0%	17.6%	13.8%	100.0%
Injury and Poisoning	24.0%	20.2%	23.0%	18.4%	14.4%	100.0%
Acute respiratory	18.8%	18.5%	23.8%	19.4%	19.5%	100.0%
Chronic respiratory	20.0%	25.6%	20.6%	19.1%	14.8%	100.0%
High blood pressure	20.7%	20.6%	21.5%	20.2%	17.0%	100.0%
Population	19.9%	18.8%	20.9%	20.4%	20.0%	100.0%

Table A4.3: Morbidity concentration indices by age group

Indicator	Age Groups				
	< 18	18-34	35-44	45-64	65 +
Bed days	-0.027	-0.179	-0.294	-0.205	-0.125
Off-work days	-0.088	-0.173	-0.198	-0.193	-0.068
Restr. activ. days	-0.055	-0.141	-0.170	-0.176	-0.102
Illness	0.061	-0.039	-0.047	-0.061	-0.052
Limiting chronic	-0.035	-0.094	-0.136	-0.139	-0.109
Acute	0.038	-0.036	-0.016	-0.045	-0.047
Chronic	0.035	-0.046	-0.137	-0.092	-0.050
Circulatory	0.388	-0.114	-0.332	-0.065	-0.044
Respiratory	0.036	0.014	0.054	-0.044	-0.064
Digestive	0.027	-0.029	-0.066	-0.058	-0.051
Muscoskeletal	0.062	-0.189	-0.130	-0.148	-0.062
Injury & poisoning	-0.012	-0.159	-0.212	-0.102	-0.041
Sample size	9312	7833	4456	9410	5991

Table A4.4: Statistical significance of the relationship between morbidity and equivalent income quintiles by age group.
Chi-squared Tests

Indicator	Age Groups				
	< 18	18-34	35-44	45-64	65 +
Bed days	-	****	***	****	***
Off-work days	-	***	****	****	*
Restr. activ. days	-	****	****	****	****
Illness	-	-	-	***	***
Limiting chronic	-	***	****	****	****
Acute	-	-	-	-	-
Chronic	-	-	****	****	****
Circulatory	-	-	****	****	*
Respiratory	-	-	*	*	*
Digestive	-	-	-	-	-
Muscoskeletal	-	**	*	****	****
Injury & poisoning	-	*	**	**	-

\- Not significant at acceptable levels
* Significant at p = 0.05 ** Significant at p = 0.01
*** Significant at p = 0.005 **** Significant at p = 0.001

Table A4.5: Equivalence scales used in portuguese research and in the ECuity study. Description and estimated family size elasticities

Scale	Scale relativities	Source	Portuguese or health studies where scale used	Family size elasticity	R-square
1. OECD	First adult = 1.00 Other adults = 0.70 Child aged < 14 = 0.5	OECD (1982)	Teekens (1991) Rodrigues (1993) Lachaud and Rochaix (1993) Christiansen (1993)	0.769	0.996
2. ILO	Adult male aged 14-59 = 1.00 Adult female aged 14-59 = 0.80 Adult aged ≥ 60 = 0.80 Child aged 12-13 = 0.80 Child aged 10-11 = 0.70 Child aged 6-9 = 0.50 Child aged 4-5 = 0.40 Child aged 2-3 = 0.30 Child aged < 2 = 0.20	INE (1990)	Used generally by Portuguese Statistical Office (INE)	0.823	0.974
3. Santos I	First adult = 1.00 Spouse of head = 0.77 *Other adults = 0.77* Child aged 14-17 = 0.59 Child aged 5-13 = 0.50 Child aged < 5 = 0.39	Santos (1984)		0.783	0.993
4. Santos II	First adult = 1.00 Spouse of head = 0.89 *Other adults = 0.89* Child aged 14-17 = 0.74 Child aged 5-13 = 0.58 Child aged < 5 = 0.46	Santos (1984)	Pinto (1988)	0.864	0.994

cont.

Table A4.5: Equivalence scales used in portuguese research and in the ECuity study.
Description and estimated family size elasticities
(cont.)

Scale	Scale relativities	Source	Portuguese or health studies where scale used	Family size elasticity	R-square
5. CSO-DSS	First adult = 1.00 Spouse of head = 0.64 Other second adult = 0.79 Third adult = 0.69 Each subsequent adult = 0.59 Child aged 16-17 = 0.59 Child aged 13-15 = 0.44 Child aged 11-12 = 0.41 Child aged 8-10 = 0.38 Child aged 5-7 = 0.34 Child aged 2-4 = 0.3 Child aged < 2 = 0.15	McClements (1978)	O'Donnell and Propper (1991) Propper and Upward (1992) O'Donnell *et al* (1993)	0.690	0.986
6. Ireland	First adult = 1.00 Each extra adult = 0.66 Each child = 0.33	Nolan (1993), based on relativities implicit in social welfare payments	Nolan (1993)	0.677	0.984
7. Netherlands	First person = 1.00 Second person = 0.40 Third person = 0.28 Each extra person = 0.17	Schiepers (1992)	Van Doorslaer, Wagstaff and Janssen (1993)	0.454	0.969
8. Switzerland	First adult = 1.00 Second adult = 0.5 *Each additional adult = 0.5* First child = 0.39 Second child = 0.22 Each additional child = 0.25	Leu and Gerfin (1993)	Leu and Gerfin (1993)	0.590	0.986

Table A4.6: ECuity study estimates of health inequality in nine countries
Standard concentration indices for unstandardized distributions

Morbidity variable	Denmark 1982	France 1980	Ireland 1987	Italy 1985	Netherlands 1981	Portugal 1987	Spain 1987	Switzerland 1981	U.K. 1985
Chronic illness	-0.086	-0.009	-0.148	-0.075	-0.040	-	-0.058	-0.013	-0.140
	0.08	*0.00*	*0.14*	*0.07*	*0.03*	-	*0.05*	*0.00*	*0.13*
Limiting chronic illness	-	-	-	-0.041	-	-0.111	-0.104	-0.202	-0.243
	-	-	-	*0.00*	-	*0.07*	*0.06*	*0.16*	*0.20*
Self-assessed health	-0.237	-	-	-0.166	-0.115	-	-0.220	-0.156	-0.317
	0.12	-	-	*0.05*	*0.00*	-	*0.11*	*0.04*	*0.20*

Notes: The numbers directly below each concentration index value represent the absolute difference in relation to the lowest observed value.
Source: Van Doorslaer *et al* (1993, p.87).

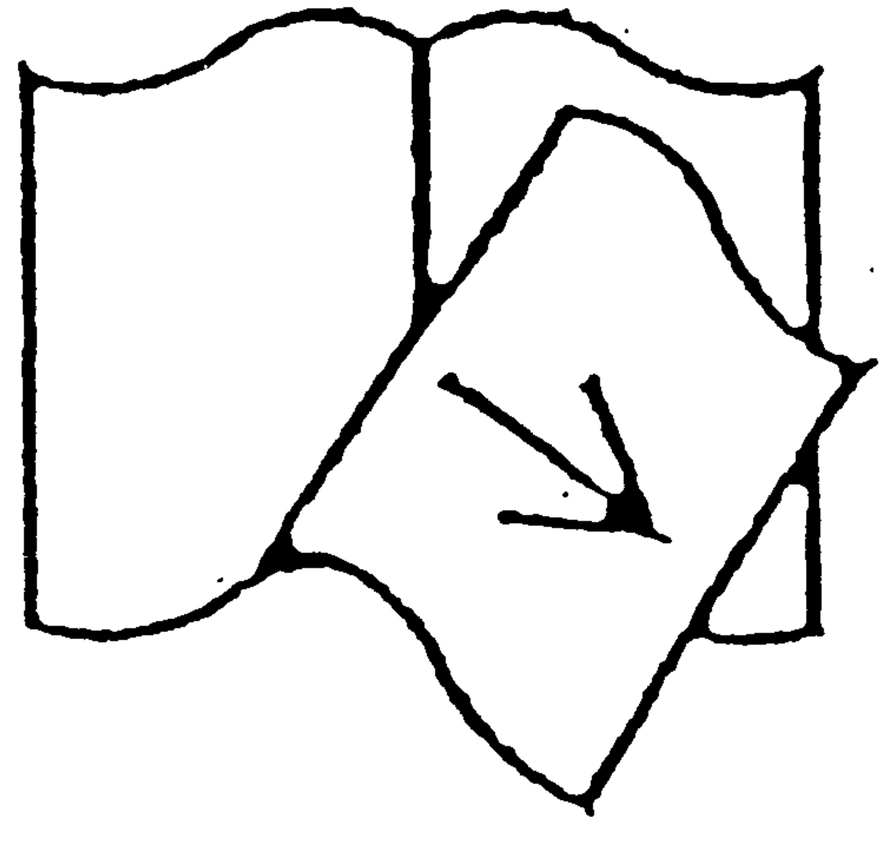
Notes to Table A4.5

1. The sources for the Swiss and Santos scales do not indicate a weight to be attributed to third and further adult family members. In applying these scales to the INS data I have assigned such persons the same weight as the second adult (shown in italics in column 2).
2. The Dutch scale is a summary of the more detailed CBS scale used by van Doorslaer, Wagstaff and Janssen (1993). The summary which I have used is provided by Schiepers (1992), who is one of the original scale's authors.
3. It is not clear from the U.S. country report whether or which equivalence scale was applied. I have therefore not included that country in the calculations.
4. The results in the two final columns were obtained by estimating equation (5.5) using a family aggregated INS data file.

Appendix 5

Inequity in Health Care Finance: Further Results

PAGE(S) MISSING
NOT AVAILABLE



p 301 - 309

Appendix 6

Health Social Welfare: Mathematical Derivations

Population health welfare index

In this appendix the terms h_i and h_j are substituted by their equivalent expressions $g(y_i)$ and $g(y_j)$. Equation (8.8) gives the expected health welfare enjoyed by the representative individual with income y_i

$$E[HW \mid g(y_i)] = \int_0^{y_i} g(y_i)\, f(y_j)\, dy_j + \int_{y_i}^{\infty} [g(y_i) - \kappa[g(y_j) - g(y_i)]]\, f(y_j)\, dy_j \quad (8.8)$$

the solution of which is:

$$= g(y_i) \int_0^{y_i} f(y_j)\, dy_j + g(y_i) \int_{y_i}^{\infty} f(y_j)\, dy_j - \kappa \int_{y_i}^{\infty} g(y_j) f(y_j)\, dy_j + \kappa\, g(y_i) \int_{y_i}^{\infty} f(y_j)\, dy_j$$

$$= g(y_i) - \kappa \int_{y_i}^{\infty} g(y_j) f(y_j)\, dy_j + \kappa\, g(y_i)\, [F(\infty) - F(y_i)]$$

$$= g(y_i) - \kappa \int_{y_i}^{\infty} g(y_j) f(y_j)\, dy_j + \kappa\, g(y_i)\, [1 - F(y_i)].$$

From probability theory:

$$\int_0^{\infty} g(y_j)\, f(y_j)\, dy_j = E[g(y_j)] = \mu_h$$

$$\rightarrow \quad \int_0^{y_i} g(y_j)\, f(y_j)\, dy_j + \int_{y_i}^{\infty} g(y_j)\, f(y_j)\, dy_j = \mu_h$$

$$\rightarrow \quad \int_{y_i}^{\infty} g(y_j)\, f(y_j)\, dy_j = \mu_h - \int_0^{y_i} g(y_j)\, f(y_j)\, dy_j.$$

Now substitute back into solution:

$$= g(y_i) - \kappa \left[\mu_h - \int_0^{y_i} g(y_j) f(y_j)\, dy_j\right] + \kappa\, g(y_i)\,[1 - F(y_i)].$$

Given that individuals are arranged in ascending order of their incomes the proportion of health enjoyed by persons with income equal to or below y_i is given by the first-moment distribution function:

$$F_1[g(y_i)] = \frac{1}{E[g(y_i)]} \int_0^{y_i} g(Y)\, f(Y)\, dY \qquad (8.9)$$

$$= \frac{1}{\mu_h} \int_0^{y_i} g(Y)\, f(Y)\, dY.$$

It follows that:

$$\int_0^{y_i} g(y_j)\, f(y_j)\, dy_j = \mu_h\, F_1\,[g(y_i)].$$

Substitution once again will give the following health welfare curve

$$W(h_i) = g(y_i) - \kappa\mu_h \{1 - F_1[g(y_i)]\} + \kappa\, g(y_i)\,[1 - F(y_i)] \qquad (8.10)$$

If income y_i corresponds to the 100pth percentile, then the health concentration curve is represented by $G_h(p) = F_1[g(y_i)]$, where $p = F(y_i)$ and $0 \leq p \leq 1$. The following conditions also hold:

$$(a) \quad \text{if } p = 0, \qquad G_h(p) = 0$$

$$(b) \quad \text{if } p = 1, \qquad G_h(p) = 1$$

$$(c) \quad G_h'(p) = \frac{dF_1[g(y_i)]}{dF(y_i)} = \frac{[g(y_i)]}{\mu_h} \geq 0 \tag{8.11}$$

$$G_h''(p) = \frac{d^2F_1[g(y_i)]}{dF^2(y_i)} = \frac{g'(y_i)}{\mu_h} \cdot \frac{1}{f(y_i)} > 0.$$

Using these relationships in (8.10) yields the average health well-being enjoyed by an individual at the 100pth percentile.

Recall (8.10):

$$W(h_i) = g(y_i) - \kappa\mu_h\{1 - F_1[g(y_i)]\} + \kappa g(y_i)[1 - F(y_i)] \tag{8.10}$$

$$= g(y_i) - \kappa\mu_h[1 - G_h(p)] + \kappa g(y_i)[1 - p]$$

$$\frac{W(h_i)}{\mu_h} = \frac{g(y_i)}{\mu_h} - \kappa + \kappa G_h(p) + \frac{\kappa g(y_i)}{\mu_h} - \frac{\kappa p\, g(y_i)}{\mu_h}$$

$$\frac{W(h_i)}{\mu_h} = G_h'(p) - \kappa + \kappa G_h(p) + \kappa G_h'(p) - \kappa p\, G_h'(p)$$

$$W(h_i) = W(p) = \mu_h[G_h'(p)(1 + \kappa - \kappa p) - \kappa + \kappa G_h(p)] \tag{8.12}$$

The area under this curve can be interpreted as the average health welfare enjoyed by society

$$\int_0^1 W(p)\,dp = \mu_h \int_0^1 [\, G_h'(p)\,(1+\kappa-\kappa p) - \kappa + \kappa G_h(p)\,]\,dp \tag{8.13}$$

which on evaluating the integrals leads to:

$$= \mu_h \int_0^1 G_h'(p)\,(1+\kappa)\,dp - \mu_h \int_0^1 \kappa p\, G_h'(p)\,dp - \mu_h \int_0^1 \kappa\,dp + \mu_h \int_0^1 \kappa G_h(p)\,dp$$

$$= \mu_h\,(1+\kappa)\,G_h(p)\;\Big|_0^1 - \mu_h \kappa \int_0^1 G_h'(p)\,p\,dp - \mu_h\,\kappa\;\Big|_0^1 + \mu_h \kappa \int_0^1 G_h(p)\,dp$$

$$= \mu_h\,(1+\kappa) - \mu_h \kappa\,[\,1 - \int_0^1 G_h(p)\,dp\,] - \mu_h\,\kappa + \mu_h \kappa \int_0^1 G_h(p)\,dp$$

$$= \mu_h\,[\,1 + \kappa \int_0^1 G_h(p)\,dp - \kappa + \kappa \int_0^1 G_h(p)\,dp\,]$$

Therefore:

$$W_h = \mu_h[\,1 - \kappa + 2\kappa \int_0^1 G_h(p)\,dp\,]$$

$$= \mu_h[\,1 - \kappa(1 - 2\int_0^1 G_h(p)\,dp)\,] \tag{8.14}$$

Health welfare by income groups

The average level of health welfare of the poorest $100p_1$ percent of the population is given by [taking into account equation (8.12)]

$$W(p_1) = \frac{1}{p_1} \int_0^{p_1} \mu_h [G_h'(p)\,(1+\kappa-\kappa p) - \kappa + \kappa G_h(p)]\,dp \qquad (8.17)$$

$$= \frac{\mu_h}{p_1} \left[\int_0^{p_1} (1+\kappa)\, G_h'(p)\,dp - \int_0^{p_1} \kappa p\, G_h'(p) - \int_0^{p_1} \kappa\, dp + \int_0^{p_1} \kappa\, G_h(p)\,dp \right]$$

Integration by parts leads to:

$$= \frac{\mu_h}{p_1} \{ (1+\kappa)\, G_h(p) \Big|_0^{p_1} - [\, \kappa p G_h(p) \Big|_0^{p_1} - \int_0^{p_1} G_h(p)\kappa dp\,] - \kappa \Big|_0^{p_1} + \int_0^{p_1} \kappa G_h(p)dp \}$$

$$= \frac{\mu_h}{p_1} \{ (1+\kappa)\, G_h(p_1) - \kappa p_1\, G_h(p_1) + \kappa \int_0^{p_1} G_h(p)\,dp - \kappa\, p_1 + \kappa \int_0^{p_1} \kappa\, G_h(p)\,dp \}$$

$$= \frac{\mu_h}{p_1} \left[G_h(p_1) + \kappa\, G_h(p_1) - \kappa\, p_1\, G_h(p_1) - \kappa p_1 + 2\kappa \int_0^{p_1} \kappa\, G_h(p)\,dp \right] \qquad (8.18)$$

$$= \frac{\mu_h\, G_h(p_1)}{p_1} + \frac{\kappa\, \mu_h\, G_h(p_1)}{p_1} - \frac{\kappa p_1 \mu_h G_h(p_1)}{p_1} - \frac{\kappa p_1 \mu_h}{p_1} + \frac{2\kappa \mu_h}{p_1} \int_0^{p_1} G_h(p)\,dp$$

Given that $G_h(p_1)$ is the proportion of total health held by the poorest p_1 decimal fraction of the population, the mean level of health status of this group is given by:

$$\mu_{hp_1} = \frac{\mu_h\, G_h(p_1)}{p_1} \qquad (8.19)$$

Using this result, we can now write $W(p_1)$ as follows:

$$W(p_1) = \mu_{hp_1} + \kappa\,\mu_{hp_1} - \kappa p_1\,\mu_{hp_1} - \kappa\,\mu_h + \frac{2\kappa\,\mu_h}{p_1}\int_0^{p_1} G_h(p)\,dp$$

$$= \mu_{hp_1} + \kappa\,[\mu_{hp_1} - p_1\,\mu_{hp_1} - \mu_h + \frac{2\,\mu_h}{p_1}\int_0^{p_1} G_h(p)\,dp\,]$$

$$= \mu_{hp_1} - \kappa\,[(\mu_h - \mu_{hp_1}) + p_1\,\mu_{hp_1} - \frac{2\,\mu_h}{p_1}\int_0^{p_1} G_h(p)\,dp\,]$$

$$= \mu_{hp_1} - \kappa\,\{\,(\mu_h - \mu_{hp_1}) + p_1\,\mu_{hp_1}\,[\,1 - \frac{2\,\mu_h}{p_1^2\,\mu_{hp_1}}\int_0^{p_1} G_h(p)\,dp\,]\,\}$$

$$= \mu_{hp_1} - \kappa\,\{\,(\mu_h - \mu_{hp_1}) + p_1\,\mu_{hp_1}\,[\,1 - \frac{2\,\mu_h}{p_1^2\,\mu_h\,\dfrac{G_h(p_1)}{p_1}}\int_0^{p_1} G_h(p)\,dp\,]\,\}$$

$$= \mu_{hp_1} - \kappa\,\{\,(\mu_h - \mu_{hp_1}) + p_1\,\mu_{hp_1}\,[\,1 - \frac{2}{p_1\,G_h(p_1)}\int_0^{p_1} G_h(p)\,dp\,]\,\}$$

From Kakwani (1980, p. 336) it is clear that:

$$C_{hp_1} = 1 - \frac{2}{p_1\,G_h(p_1)}\int_0^{p_1} G_h(p)\,dp \qquad (8.20)$$

is the health concentration coefficient for the poorest $100p_1$ per cent of the population. It follows that:

$$W(p_1) = \mu_{hp_1} - \kappa[(\mu_h - \mu_{hp_1}) + p_1\,\mu_{hp_1}\,C_{hp_1}]\,. \qquad (8.21)$$

Appendix 7

Health Social Welfare: Further Results

Table A7.1: Micro-level statistics for five continental regions
National Health Survey, 1987

Indicator		Portugal	North	Centre	Lisbon	Alentejo	Algarve
No. valid cases		37002	12695	7675	10545	3769	2318
% of sample		100.0%	34.3%	20.7%	28.5%	10.2%	6.3%
Average family size		3.6	4.0	3.5	3.4	3.2	3.0
Average age		38.4	35.7	39.8	38.0	42.0	44.1
Sex	M	47.7%	47.9%	47.1%	47.6%	47.4%	49.0%
	F	52.3%	52.1%	52.9%	52.4%	52.6%	51.0%
% with consultations		44.9%	43.1%	44.6%	47.9%	44.5%	42.9%
Aver. no. consultations		1.13	1.07	1.13	1.22	1.11	1.07
Mean equiv income		18795	17775	15441	22689	17154	20440
Gini coefficient		0.346	0.351	0.335	0.316	0.288	0.310

Notes:
- Doctor consultations in a 3-month reference period.
- Mean equivalent income on a per month basis in escudos.

Table A7.2: Social welfare associated with health in five portuguese regions

Age and sex-standardized. Values of $\kappa = 0, 1, 2$ and 3.

	Portugal	North	Centre	Lisbon	Alentejo	Algarve
Prevalence of limiting chronic illness						
$\kappa = 0.0$	11.5%	11.3%	14.2%	10.0%	8.9%	14.0%
Variation	*100*	*98*	*123*	*87*	*77*	*122*
$\kappa = 1.0$	12.9%	12.5%	15.9%	11.4%	9.7%	14.7%
Variation	*100*	*97*	*123*	*88*	*75*	*114*
$\kappa = 2.0$	14.3%	13.7%	17.7%	12.8%	10.4%	15.5%
Variation	*100*	*95*	*123*	*89*	*72*	*108*
$\kappa = 3.0$	15.8%	14.9%	19.4%	14.2%	11.2%	16.2%
Variation	*100*	*94*	*123*	*90*	*71*	*103*
Prevalence of disability and illness						
$\kappa = 0.0$	28.2%	27.2%	29.3%	27.1%	28.4%	34.5%
Variation	*100*	*97*	*104*	*96*	*101*	*122*
$\kappa = 1.0$	30.0%	29.0%	30.6%	29.3%	30.4%	36.0%
Variation	*100*	*97*	*102*	*98*	*101*	*120*
$\kappa = 2.0$	31.7%	30.7%	32.0%	31.6%	32.4%	37.5%
Variation	*100*	*97*	*101*	*100*	*102*	*118*
$\kappa = 3.0$	33.5%	32.4%	33.4%	33.8%	34.4%	39.1%
Variation	*100*	*97*	*100*	*101*	*103*	*117*
Number of sick days						
$\kappa = 0.0$	1.419	1.457	1.684	1.199	1.152	1.775
Variation	*100*	*103*	*119*	*84*	*81*	*125*
$\kappa = 1.0$	1.603	1.649	1.865	1.367	1.266	1.870
Variation	*100*	*103*	*116*	*85*	*79*	*117*
$\kappa = 2.0$	1.786	1.840	2.046	1.535	1.380	1.966
Variation	*100*	*103*	*115*	*86*	*77*	*110*
$\kappa = 3.0$	1.969	2.032	2.226	1.703	1.494	2.062
Variation	*100*	*103*	*113*	*86*	*76*	*105*

Note: The number directly below each computed value shows the variation around the national average.

Table A7.3: Average levels of morbidity and income indexed health inequality in five portuguese regions

Observed and age-sex standardized distributions

	Portugal	North	Centre	Lisbon	Alentejo	Algarve
Prevalence of limiting chronic illness						
% reporting						
Unstandardised	11.6%	10.5%	14.9%	10.0%	9.8%	16.1%
Standardised	11.5%	11.3%	14.2%	10.0%	8.9%	14.0%
Concentration coefficients						
Unstandardised	-0.124	-0.105	-0.124	-0.137	-0.057	-0.088
Standardised	-0.124	-0.106	-0.123	-0.141	-0.084	-0.053
Prevalence of disability and illness						
% reporting						
Unstandardised	28.5%	26.0%	30.5%	27.0%	31.0%	38.4%
Standardised	28.2%	27.2%	29.3%	27.1%	28.4%	34.5%
Concentration coefficients						
Unstandardised	-0.072	-0.068	-0.048	-0.089	-0.060	-0.081
Standardised	-0.063	-0.064	-0.047	-0.083	-0.071	-0.044
Number of sick days						
Mean						
Unstandardised	1.42	1.35	1.76	1.19	1.24	2.02
Standardised	1.42	1.46	1.68	1.20	1.15	1.77
Concentration coefficients						
Unstandardised	-0.133	-0.130	-0.117	-0.136	-0.070	-0.103
Standardised	-0.129	-0.131	-0.107	-0.140	-0.099	-0.054

Table A7.4: Health social welfare by income percentiles.

Unstandardized and age-sex standardized. Values of κ = 0, 0.5, 1, and 2.

Income ranked population		Limiting chronic illness						Disability and illness						Sick days					
Percentiles	n	Reporting illness	C_w index	Health illfare				Reporting illness	C_w index	Health illfare				Average number	C_w index	Health illfare			
				κ=0	κ=0.5	κ=1	κ=2			κ=0	κ=0.5	κ=1	κ=2			κ=0	κ=0.5	κ=1	κ=2
Unstandardized																			
≤ 20 %	7361	0.137	0.079	0.137	0.147	0.157	0.176	0.294	0.070	0.294	0.297	0.300	0.305	1.727	0.075	1.727	1.868	2.010	2.292
≤ 39 %	14302	0.154	0.116	0.154	0.170	0.185	0.217	0.347	0.132	0.347	0.369	0.390	0.434	1.905	0.103	1.905	2.110	2.314	2.724
≤ 60 %	22051	0.137	-0.019	0.137	0.148	0.159	0.182	0.319	0.009	0.319	0.334	0.350	0.382	1.703	-0.023	1.703	1.856	2.010	2.316
≤ 80 %	29604	0.126	-0.070	0.126	0.135	0.144	0.162	0.303	-0.030	0.303	0.316	0.328	0.354	1.553	-0.082	1.553	1.670	1.788	2.023
≤ 100 %	37002	0.116	-0.124	0.116	0.123	0.130	0.144	0.285	-0.072	0.285	0.295	0.306	0.326	1.419	-0.133	1.419	1.514	1.608	1.797
Standardized																			
≤ 20 %	7361	0.143	-0.005	0.143	0.159	0.174	0.205	0.304	0.012	0.304	0.317	0.330	0.357	1.838	0.007	1.838	2.057	2.277	2.715
≤ 39 %	14302	0.137	-0.008	0.137	0.150	0.162	0.188	0.308	0.023	0.308	0.322	0.336	0.364	1.750	-0.018	1.750	1.930	2.110	2.471
≤ 60 %	22051	0.131	-0.038	0.131	0.142	0.153	0.175	0.300	-0.004	0.300	0.312	0.324	0.348	1.650	-0.048	1.650	1.798	1.945	2.240
≤ 80 %	29604	0.122	-0.073	0.122	0.131	0.139	0.156	0.291	-0.025	0.291	0.301	0.311	0.331	1.523	-0.089	1.523	1.638	1.753	1.982
≤ 100 %	37002	0.112	-0.124	0.112	0.119	0.126	0.140	0.277	-0.060	0.277	0.285	0.294	0.310	1.402	-0.132	1.402	1.494	1.587	1.772

Table A7.5: Health social welfare by income groups.

Unstandardized and age-sex standardized. Values of κ = 0, 0.5, 1 and 2.

Quintiles	Limiting chronic illness								Disability and illness								Sick days							
	Average level of health illfare				% Variation				Average level of health illfare				% Variation				Average level of health illfare				% Variation			
	κ=0	κ=.5	κ=1	κ=2	κ=0	κ=.5	κ=1	κ=2	κ=0	κ=.5	κ=1	κ=2	κ=0	κ=.5	κ=1	κ=2	κ=0	κ=.5	κ=1	κ=2	κ=0	κ=.5	κ=1	κ=2
Unstandardized																								
I	0.137	0.147	0.157	0.176	119%	120%	121%	122%	0.294	0.297	0.300	0.305	103%	101%	98%	93%	1.727	1.868	2.010	2.292	103%	101%	98%	93%
II	0.172	0.194	0.216	0.260	149%	158%	166%	180%	0.402	0.444	0.487	0.572	141%	150%	159%	175%	2.093	2.365	2.638	3.183	141%	150%	159%	175%
III	0.105	0.108	0.111	0.118	91%	88%	86%	82%	0.267	0.272	0.276	0.286	94%	92%	90%	88%	1.330	1.389	1.447	1.563	94%	92%	90%	88%
IV	0.096	0.098	0.100	0.103	83%	80%	77%	72%	0.258	0.261	0.265	0.272	90%	88%	87%	83%	1.113	1.127	1.141	1.169	90%	88%	87%	83%
V	0.072	0.072	0.073	0.073	63%	59%	56%	51%	0.213	0.213	0.214	0.215	75%	72%	70%	66%	0.886	0.887	0.888	0.890	75%	72%	70%	66%
Total	0.116	0.123	0.130	0.144	100%	100%	100%	100%	0.285	0.295	0.306	0.326	100%	100%	100%	100%	1.419	1.514	1.608	1.797	100%	100%	100%	100%
Standardized																								
I	0.143	0.159	0.174	0.205	128%	133%	138%	147%	0.304	0.317	0.330	0.357	110%	111%	113%	115%	1.838	2.057	2.277	2.715	131%	138%	143%	153%
II	0.131	0.140	0.150	0.169	117%	118%	119%	121%	0.312	0.327	0.343	0.373	113%	115%	117%	120%	1.657	1.795	1.934	2.211	118%	120%	122%	125%
III	0.120	0.128	0.135	0.151	107%	107%	108%	108%	0.285	0.293	0.301	0.316	103%	103%	102%	102%	1.465	1.553	1.640	1.814	105%	104%	103%	102%
IV	0.096	0.097	0.099	0.102	85%	82%	78%	73%	0.265	0.269	0.273	0.281	96%	94%	93%	91%	1.152	1.172	1.191	1.230	82%	78%	75%	69%
V	0.072	0.072	0.073	0.074	64%	61%	58%	53%	0.221	0.223	0.224	0.228	80%	78%	76%	74%	0.918	0.921	0.924	0.929	65%	62%	58%	52%
Total	0.112	0.119	0.126	0.140	100%	100%	100%	100%	0.277	0.285	0.294	0.310	100%	100%	100%	100%	1.402	1.494	1.587	1.722	100%	100%	100%	100%

References

Aaron, H. (1992), 'Equity in the finance and delivery of health care' Journal of Health Economics, Vol. 11, pp. 467-471.

Abel-Smith, B. (1984), Cost Containment in Health Care: A Study of 12 European Countries, 1977-83, Bedford Square Press, London.

Abel-Smith, B. (1992), Cost Containment and new Priorities in Health Care: A Study of the European Community. Avebury, Aldershot.

Aday, L.A., R.A. Andersen, and G.V. Fleming (1980), Health Care in the U.S. Equitable for Whom? Sage Publications, Beverly Hills.

Andersen, R., J. Kravits, and O.W. Anderson (1975), Equity in Health Care Services: Empirical Analyses in Social Policy. Ballinger, Cambridge, Mass.

Archibald, G.C., and D. Donaldson (1979), 'Notes on economic equality' Journal of Public Economics, Vol. 12, pp. 205-214.

Arneson, R. (1989), 'Equality and equal opportunity for welfare' Philosophical Studies, Vol. 56, pp. 77-93.

Aronson, J.R., P. Johnson, and P. Lambert (1994), 'Redistributive effect and unequal tax treatment' Economic Journal, Vol. 104, pp. 262-270.

Arrow, K.J. (1973), 'The utilitarian approach to the concept of equality in public expenditure' In: E. Phelps (*ed.*) Economic Justice, Harmondsworth Penguin, London.

Arrow, K. (1974), 'Gifts and exchanges' Philosophy and Public Affairs, Vol. 1, pp. 343-362.

Assembleia da República (1979), Serviço Nacional de Saúde. Direcção dos Serviços de Documentação e Informação Bibliográfica, Lisboa.

Atkinson, A.B. (1970), 'On the measurement of inequality' Journal of Economic Theory, Vol. 2, pp. 244-63.

Atkinson, A.B. (1980), 'Horizontal equity and the tax burden' In: H. Aaron and M.J. Boskin (*eds.*) The Economics of Taxation. Brookings Institution, Washington.

Atkinson, A.B. (1983), The Economics of Inequality, Second Edition, Oxford University

Press, Oxford.

Atkinson, A.B. (1990), 'The distribution of the tax burden: 30 years after the Theory of Public Finance', Welfare State Programme Discussion Paper no. WSP/51, STICERD, London School of Economics.

Atkinson, A.B., and F. Bourguignon (1982), 'The comparison of multi-dimensioned distributions of economic status' Review of Economic Studies, Vol. 49, pp. 183-201.

Atkinson, A.B., and J. Micklewright (1983), 'On the reliability of income data in the Family Expenditure Survey 1970-1977', Journal of the Royal Statistical Society, Vol. 146, pp. 33-61.

Atkinson, A.B., and J. Micklewright (1992), Economic Transformation in Eastern Europe and the Distribution of Income. Cambridge University Press, Cambridge.

Atkinson, A.B., K. Gardiner, V. Lechêne, and H. Sutherland (1993), 'Comparing poverty in France and the United Kingdom' Welfare State Programme, Discussion Paper no. WSP/84, STICERD, London School of Economics.

Atkinson, A.B., and J. Stiglitz (1980), Lectures on Public Economics. McGraw-Hill, New York.

Baptista Pereira, A. (1986), Speech at the closing session of the First National Workshop on Equity and Health, DEPS, Lisbon, 17 December 1986.

Basu, K. (1987), 'Achievements, capabilities and the concept of well-being' Social Choice and Welfare, Vol. 4, pp. 69-76.

Baumol, W.J. (1986), Superfairness: Applications and Theory. The MIT Press, Cambridge, Mass.

Becker, G. (1965), 'A theory of the allocation of time' Economic Journal, Vol. 75, pp. 493-517.

Becker, G., and N. Tomes (1978), 'An equilibrium theory of the distribution of income and intergenerational mobility' Journal of Political Economy, Vol. 87, pp. 1153-1189.

Bell, J.M. and S. Mendus, *eds.* (1988), Philosophy and Medical Welfare. Cambridge University Press, Cambridge.

Bergson, A. (1938), 'A reformulation of certain aspects of welfare economics' Quarterly Journal of Economics, Vol. 52, pp. 310-334.

Berk, R.A. (1987), 'Household production' In: J. Eatwell, M. Milgate, and P. Newman, (*eds.*) The New Palgrave: A Dictionary of Economics. MacMillan, London.

Birch, S. (1987), *The Demand for Health: Extensions to the Human Capital Model.* Unpublished DPhil Thesis. Department of Economics and Related Studies, University of York.

Bishop, J.A., S. Chakraborti, and P.D. Thistle (1990), 'An asymptotically distribution-free test for Sen's welfare index' Oxford Bulletin of Economics and Statistics, Vol. 52, pp. 105-113.

Bishop, J.A., S. Chakraborti, and P.D. Thistle (1991), 'Relative deprivation and economic welfare: a statistical investigation with Gini-based welfare indices' Scandinavian Journal of Economics, Vol. 93, pp. 421-437.

Blackorby, C., and D. Donaldson (1984), 'Ethical social index numbers and the measurement of effective tax / benefit progressivity' Canadian Journal of Economics, Vol. 17, pp. 683-694.

Blaxter, M. (1981), The Health of Children. Heinemann, London.

Blaxter, M. (1989), 'A comparison of measures of inequality in morbidity' In: J. Fox (*ed.*), Health Inequality in European Countries. Gower, Aldershot.

Blaxter, M. (1990), Health and Lifestyles. Tavistock/Routledge, London.

Borooah, V., P. McGregor, and P. McKee (1991), Regional Income Inequality and Poverty in the United Kingdom. Dartmouth, Aldershot.

Bradbury, B. (1989), 'Family size equivalence scales and survey evaluations of income and well-being' Journal of Social Policy, Vol. 18, pp. 383-408.

Bradshaw, J., H. Edwards, D. Lawton, *et al* (1982), 'Area variations in infant mortality' Journal of Epidemiology and Community Health, Vol. 32, pp. 11-16.

Broome, J. (1988), 'Good, fairness and QALY's' In: J.M. Bell; S. Mendus (*eds.*) Philosophy and Medical Welfare. Cambridge University Press, Cambridge.

Buhmann, B., L. Rainwater, G. Schmaus, and T. Smeeding (1988), 'Equivalence scales, well-being, inequality and poverty: sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database' Review of Income and Wealth, Vol. 34, pp. 115-142.

Campos, A.C., *ed.* (1987), A Combinação Público-Privada em Saúde: Privilégios, Estigmas e Ineficiências. Escola Nacional de Saúde Pública, Lisboa.

Campos, A.C. (1990), 'A hora da reforma: linhas gerais para a revisão do Serviço Nacional de Saúde em Portugal', Farmácia Portuguesa, nº 63, pp. 45-49.

Carr-Hill, R. (1987) 'The inequalities in health debate: a critical review of the issues' Journal of Social Policy, Vol. 16, pp. 509-542.

Carr-Hill, R. (1988), 'QuALitY control: a sensitivity analysis of QALY's' Paper presented to the Meeting of the Health Economists Study Group, University of Brunel, July.

Carrilho, M.J. (1985), 'O declínio recente da mortalidade infantil em Portugal' Revista do Centro de Estudos Demográficos (INE), nº. 27.

Chakravarty, S.R. (1990), Ethical Social Index Numbers. Springer-Verlag, Berlin.

Christiansen, T. (1990), 'Equity in the Finance and Delivery of Health Care' Paper presented to the EC-COMAC Workshop on Equity in the Finance and Delivery of Health Care, Bellagio, November.

Christiansen, T. (1993), 'Denmark' In: E. Van Doorslaer, A. Wagstaff and F. Rutten, *eds.*, Equity in the Finance and Delivery of Health Care: An International Perspective. Oxford University Press, Oxford.

Coffey, R.M. (1983), 'The effect of time price on the demand for medical-care services' Journal of Human Resources, Vol.18, pp. 407-424.

Cohen, G.A. (1990), 'Equality of what? On welfare, goods and capabilities' Recherches Economiques de Louvain, Vol. 56, pp. 357-382.

Collins, E., and R. Klein (1980), 'Equity and the NHS: self-reported morbidity, access and primary care' British Medical Journal, Vol. 281, pp. 1111-1115.

Costa, A. B. (1994), 'The measurement of poverty in Portugal' Journal of European Social Policy, Vol. 4, pp. 95-115.

Coulter, F., F. Cowell, and S. Jenkins (1992a), 'Differences in needs and assessment of income distributions' Bulletin of Economic Research, Vol. 44, pp. 77-124.

Coulter, F.A., F. Cowell, and S. Jenkins (1992b), 'Equivalence scale relativities and the extent of inequality and poverty' Economic Journal, Vol. 102, pp. 1067-1082.

Cowell, F. (1980) 'On the structure of additive inequality measures' Review of Economic Studies, Vol. 47, pp. 521-531.

Cowell, F. (1988), 'Inequality decomposition: three bad measures' Bulletin of Economic Research, Vol. 40, pp. 309-312.

Cropper, M.L. (1977), 'Health, investment in health and occupational choice' Journal of Political Economy, Vol. 85, pp. 1273-1294.

Culyer, A.J. (1971), 'Medical care and the economics of giving' Economica, Vol. 38, pp. 295-303.

Culyer, A.J. (1976), Need and the National Health Service. Martin Robertson, London.

Culyer, A.J. (1978), Measuring Health: Lessons for Ontario. University of Toronto Press, Toronto.

Culyer, A.J. (1980), The Political Economy of Social Policy. Martin Robertson, Oxford.

Culyer, A.J. (1988), 'Inequality in health services is, in general, desirable' In: D.J. Green (*ed.*) Acceptable Inequalities? The IEA Health Unit, London.

Culyer, A.J. (1989), 'The normative economics of health care finance and provision' Oxford Review of Economic Policy, Vol. 5, pp. 34-58.

Culyer, A.J. (1990), 'Commodities, characteristics of commodities, characteristics of people, utilities and the quality of life' In: S. Baldwin, C. Godfrey and C. Propper (*eds.*) The Quality of Life: Perspectives and Policies. Routledge, London.

Culyer, A.J. (1991), 'Health, health expenditures and equity' University of York, Centre for Health Economics, Discussion Paper 83.

Culyer, A.J., and A. Wagstaff (1991), 'Need, equality and social justice', Discussion Paper 90, Centre for Health Economics, University of York.

Culyer, A.J., and A. Wagstaff (1992), 'Need, equity and equality in health and health care', Discussion Paper 95, Centre for Health Economics, University of York.

Culyer, A.J., and A. Wagstaff (1993), 'Equity and equality in health and health care' Journal of Health Economics, Vol.12, pp. 431-437.

Culyer, A.J., E. van Doorslaer, and A. Wagstaff (1992), 'Access, utilization and equity: a further comment' Journal of Health Economics, Vol.11, pp. 207-210.

Dalton, H. (1920), 'The measurement of the inequality of incomes' Economic Journal, Vol. 30, pp. 348-361.

Daniels, N. (1981), 'Health care needs and distributive justice' Philosophy and Public Affairs, Vol. 10, pp. 146-179.

Daniels, N. (1982), 'Equity of access to health care: some conceptual and ethical issues' Millbank Memorial Fund Quarterly/ Health and Society, Vol. 60, pp. 51-81.

Daniels, N. (1990), 'Equality of what: welfare, resources or capabilities' Philosophical and Phenomenological Research, Vol. 1 (Supplement), pp. 237-296.

Danziger, S., and M. Taussig (1979), 'The income unit and the anatomy of income distribution' Review of Income and Wealth, Vol. 25, pp. 365-375.

Dardanoni, V., and A. Wagstaff (1987), 'Uncertainty, inequalities in health and the demand for health' Journal of Health Economics, Vol. 6, pp. 283-290.

Dasgupta, P., A. Sen, and D. Starrett (1973), 'Notes on the measurement of inequality' Journal of Economic Theory, Vol. 6, pp. 180-187.

Davies, J., F. St-Hilaire, and J. Whalley (1984), 'Some calculations of lifetime tax incidence' American Economic Review, Vol. 74, pp. 633-649.

Department of Health (1989), Working for Patients, HMSO, London.

Desai, M. (1991), 'Human development: concepts and measurement' European Economic Review, Vol. 35, pp. 350-357.

d'Houtaud, A., and M.G. Field (1984), 'The image of health: variations in the perception by social class in a French population' Sociology of Health and Illness, Vol. 6, pp. 30-60.

DHSS (1980), Inequalities in Health. Report of a Research Working Group Chaired by Sir Douglas Black. DHSS, London.

Diário de Notícias (1993), 'O estado da saúde', Lisbon, 25, 26, 27, 28 January.

Domingues, O., R. Estêves, C. Figueira, A. Martins, and A. Matias (1984), 'Incidência por classe de rendimento do Imposto sobre o Valor Acrescentado' In: Comissão do IVA/IACEP-GEBEI (*eds.*) O Impacto do IVA na Economia Portuguesa. Imprensa Nacional, Lisboa.

Donaldson, D., and J.A. Weymark (1980), 'A single parameter generalization of Gini indices of inequality' Journal of Economic Theory, Vol. 22, pp. 67-86.

Dowie, J. (1975), 'The portfolio approach to health behaviour' Social Science and Medicine, Vol. 9, pp. 619-631.

Drummond, M., G. Stoddart, and G.Torrance (1987), Methods for the Economic Evaluation of Health Care Programmes. Oxford University Press, Oxford.

Dworkin, R. (1981), ' What is equality?' Philosophy and Public Affairs, Vol. 10, pp. 185-246 and pp. 283-345.

Edouard, L. (1985), 'The epidemiology of perinatal mortality' World Health Statistics Quarterly, Vol. 38, pp. 289-301.

Enthoven, A. (1980), Health Plan: The Only Practical Solution to the Soaring Cost of Medical Care. Addison-Wellesley, Reading, Mass.

Falcão, J.M., and J.P. Silva (1983), 'Validade dos dados do Inquérito Piloto: Influência dos não-respondentes' INS, Nº 2, Ministério da Saúde, DEPS, Lisboa.

Farrel, P., and V. Fuchs (1982), 'Schooling and health. The cigarette connection' Journal of Health Economics, Vol. 1, pp. 217-230.

Feldman, A.M. (1987), 'Equity' In: J. Eatwell, M. Milgate, and P. Newman (*eds.*) The New Palgrave: A Dictionary of Economics. MacMillan, London.

Feldstein, M. (1976), 'On the theory of tax reform' Journal of Public Economics, Vol. 6, pp. 77-104.

Ferreira, V., and J.B. Lemos (1987), 'A saúde dos números - análise dos inquéritos de saúde em Portugal' Revista Crítica de Ciências Sociais, no. 23, pp. 115-128.

Ferreira, L.V. (1993), 'Pobreza em Portugal - Variação e decomposição de medidas de pobreza a partir de orçamentos familiares de 1980/81 e 1989/90' CISEP/ISEG, Documento de Trabalho nº 2/93.

Fisher, F. (1987) 'Household equivalence scales and interpersonal comparisons' Review of Economic Studies, Vol. 54, pp. 519-524.

Foley, D. (1967), 'Resource allocation and the public sector' Yale Economic Essays, Vol. 7, pp. 45-98.

Fox, J., *ed.*, (1981), Health Inequality in European Countries. Gower, Aldershot.

Formby, J., W. Smith, and P. Thistle (1991), 'A comparison of two new measures of tax progressivity' Economic Journal, Vol. 91, pp. 1015-1019.

Forster, B.A. (1989), 'Optimal health investment strategies' Bulletin of Economic Research, Vol. 41, pp. 45-57.

Franco, G. (1990), 'A Europa dos pobres' *Público*, 5 de Novembro, (Suplemento 'Economia'), pp. 4-5.

Freixinho, J.C. (1990), 'Subsistemas de Saúde em Portugal: Estudo, características e evolução futura'. Dissertação ao XIX Curso de Administração Hospitalar, mimeo, ENSP.

Friedman, L.S. (1985), Microeconomic Policy Analysis. McGraw Hill Book Company, New York.

Fuchs, V. (1987), 'Health economics' In: J. Eatwell, M. Milgate, and P. Newman, (*eds.*) The New Palgrave: A Dictionary of Economics. MacMillan, London.

Gastwirth, J. (1972), 'The estimation of the Lorenz curve and Gini index' Review of Economics and Statistics, Vol. 54, pp. 306-316.

Giraldes, M.R. (1988a), 'The equity principle in the allocation of health care expenditure on primary health care services in Portugal: the human capital approach' International Journal of Health Planning and Management, Vol. 3, pp. 167-183.

Giraldes, M.R. (1988b), 'Equidade e despesa em saúde' Estudos de Economia, Vol. 8, pp. 365-388.

Gottschalk, P., B. Wolfe, and R. Haveman (1989), 'Health care financing in the U.S., U.K. and Netherlands: distributional consequences' In: A. Chiancone and K. Messere (*eds.*) Changes in Revenue Structures. Wayne State University Press, Detroit.

Gouveia, M., and J. Tavares (1995), 'The distribution of household income and expenditure in Portugal: 1980 and 1990' Review of Income and Wealth, Vol. 41, pp. 1-17.

Green, D. (1988), Acceptable Inequalities? The IEA Health Unit, London.

Griffin, K., and J. Knight, *eds.* (1989), 'Human Development in the 1980's and Beyond' Journal of Development Planning, Vol. 19 (special number).

Grossman, M. (1972), The Demand for Health: A Theoretical and Empirical Investigation.

National Bureau of Economic Research, New York.

Grossman, M. (1982), 'The demand for health after a decade' Journal of Health Economics, Vol. 1, pp. 1-3.

Hahn, F. (1971), 'Equilibrium with transaction costs' Econometrica, Vol. 39.

Harris, J. (1987), 'QALYfying the value of life' Journal of Medical Ethics, Vol. 13.

Harris, J. (1988), 'More and better justice' In: J.M. Bell and S. Mendus (*eds.*) Philosophy and Medical Welfare, Cambridge University Press, Cambridge.

Heckman, J.J. (1976), 'A lifecycle model of earnings, learning and consumption' Journal of Political Economy, Vol. 84, pp. S11-S44.

Hey, J., and P. Lambert (1980), 'Relative deprivation and the Gini coefficient: Comment' Quarterly Journal of Economics, Vol. 95, pp. 567-573.

Hochman, H.H., and Rogers (1969), 'Pareto-optimal redistribution' American Economic Review, Vol. 59, pp. 542-557.

Illsley, R., and J. Le Grand (1987), 'The measurement of inequality in health' In: A. Williams (*ed.*) Health and Economics. MacMillan, London.

Illsley, R., and P.G. Svensson, *eds.* (1986), The Health Burden of Social Inequities. WHO Regional Office for Europe, Copenhagen.

Infarmed (1993), Informação Estatística 1993. Instituto Nacional da Farmácia e do Medicamento, Ministério da Saúde, Lisboa.

Jazairi, N.T. (1976), Approaches to the Development of Health Indicators, OECD, Paris.

Jenkins, S. (1988a) 'Reranking and the analysis of income redistribution' Scottish Journal of Political Economy, Vol. 35, pp. 65-76.

Jenkins, S. (1988b), 'Empirical measurement of horizontal inequity' Journal of Public Economics, Vol. 37, pp. 305-329.

Jenkins, S. (1988c), 'Calculating income distribution indices from micro-data', National Tax Journal, Vol. 41, pp. 139-142.

Jenkins, S. (1991), 'The measurement of income inequality' In: L. Osberg (*ed.*) Economic Inequality and Poverty: International Perspectives. M.E. Sharpe, New York.

Jenkins, S., and M. O'Higgins (1989), 'Inequality measurement using "norm incomes": were Garvy and Paglin onto something after all?' Review of Income and Wealth, Vol. 35, pp. 265-282.

Johnson, S, and T. Mayer (1962), 'An extension of Sidgwick's equity principle' Quartely

Journal of Economics, Vol. 74, pp. 454-463.

Kakwani, N. (1977a), 'Applications of Lorenz curves in economic analysis' Econometrica, Vol. 45, pp. 719-27.

Kakwani, N. (1977b), 'Measurement of tax progressivity: an international comparison' Economic Journal, Vol. 87, pp. 71-80.

Kakwani, N. (1980a), 'On a class of poverty measures' Econometrica, Vol. 48, pp. 437-46.

Kakwani, N. (1980b), Income Inequality and Poverty: Methods of Estimation and Policy Applications. Oxford University Press, New York.

Kakwani, N. (1986), Analysing Redistribution Policies: A Study Using Australian Data Cambridge University Press, London.

Kakwani, N. (1993), 'Performance in living standards: an international comparison' Journal of Development Economics, Vol. 41, pp. 307-336.

Kakwani, N., and N. Podder (1973), 'On the estimation of Lorenz curves from grouped observations' International Economic Review, Vol. 14, pp. 278-291.

Kakwani, N., and N. Podder (1976), 'Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations' Econometrica, Vol. 44, pp. 137-148.

Kannisto, V. (1986) 'Factores associados às diferenças geográficas da mortalidade infantil em Portugal desde 1950' Revista do Centro de Estudos Demográficos (INE), nº. 28, pp. 11-35.

Kars-Marshall, C., Y.N. Spronk-Boon, and M.J. Pollmans (1988), 'National health interview surveys for health care policy' Social Science and Medicine, Vol. 26, pp. 223-233.

Keeler, E., J. Newhouse, and C. Phelps (1977), 'Deductibles and the demand for medical care services: the theory of a consumer facing a variable price schedule under uncertainty' Econometrica, Vol. 45, pp. 641-655.

Kemna, H.J. (1985), *The Relationship between Schooling and Health: Investigation into an Occupational Linkage*. Unpublished Phd Thesis, University of Pittsburgh.

Kemsley, W., R. Redpath, and M. Holmes (1980), Family Expenditure Survey Handbook. HMSO, London.

Kiefer, D. (1985), 'Distributional tax progressivity indexes' National Tax Journal, Vol.37, pp. 497-513.

King, M.A. (1983), 'An index of inequality: with applications to horizontal inequity and social mobility' Econometrica, Vol. 51, pp. 99-115.

Kirscht, J. (1971), 'Social and psychological problems of surveys on health and illness'

Social Science and Medicine, Vol. 5, pp. 519-526.

Kolm, S.-Ch. (1976), 'Unequal inequalities I and II' Journal of Economic Theory, Vol. 12, pp. 416-42 and Vol. 13, pp. 82-111.

Kondor, Y. (1975), 'Value judgements implied by the use of various measures of income inequality' Review of Income and Wealth, Vol. 21, pp. 309-321.

Kravits, J., and J. Schneider (1975), 'Health care need and actual use by age race and income' In: R. Andersen, J. Kravits and O. Anderson (*eds.*) Equity in Health Services: Empirical Analysis in Social Policy Ballinger, Cambridge, Mass.

Lachaud, C., and L. Rochaix (1992), 'Une approche économétrique de la justice distributive en matière de prestacion et de financement de la santé' In: E. Levy and A. Mizrahi (*eds.*) Colloque Européen: De L'analyse Économique aux Politiques de Santé, CREDES/CES, Université Paris Dauphine, Paris, 18-20 Decembre.

Lachaud, C., and L. Rochaix (1993), 'France' In: E. Van Doorslaer, A. Wagstaff and F. Rutten, *eds.*, Equity in the Finance and Delivery of Health Care: An International Perspective. Oxford University Press, Oxford.

Lambert, P. (1985), 'Social welfare and the Gini revisited' Mathematical Social Sciences, Vol. 9, pp. 19-26.

Lambert, P. (1988), 'Net fiscal incidence progressivity: some approaches to measurement' In: W. Eichhorn (*ed.*) Measurement in Economics. Physica-Verlag, Heidelberg.

Lambert, P. (1989), The Distribution and Redistribution of Income: A Mathematical Analysis Basil Blackwell, Oxford.

Lancaster, K.J. (1966), 'A new approach to consumer theory' Journal of Political Economy, vol.74, pp. 132-157.

Layard, R. (1980), 'Human satisfaction and public policy' Economic Policy, Vol. 90, pp. 737-750.

Le Grand, J. (1978), 'The distribution of public expenditure: the case of health care' Economica, Vol. 45, pp. 125-142.

Le Grand, J. (1982), The Strategy of Equality: Redistribution and the Social Services. George Allen & Unwin, London.

Le Grand, J. (1984), 'Equity as an economic objective' Journal of Applied Philosophy, Vol. 1, pp. 39-51.

Le Grand, J. (1987a), 'Equity, health and health care' Social Justice Research, Vol. 1, pp. 257-274.

Le Grand, J. (1987b), 'Inequalities in health: some international comparisons' European Economic Review, Vol. 31, pp. 181-192.

Le Grand, J. (1991a), Equity and Choice. An Essay in Economics and Applied Philosophy. Harper Collins Academic, London.

Le Grand, J. (1991b), 'The distribution of health care revisited: a commentary on Wagstaff, van Doorslaer and Paci, and O'Donnell and Propper' Journal of Health Economics, Vol. 10, pp. 239-249.

Le Grand, J., and M. Rabin (1986), 'Trends in British health inequality, 1931-83' In: A.J. Culyer and B. Jonsson (*eds.*) Public and Private Health Services. Basil Blackwell, Oxford.

Leclerc, A., F. Lert, and C. Fabien (1990), 'Differential mortality: some comparisons between England and Wales, Finland and France, based on inequality measures' International Journal of Epidemiology, Vol. 19, pp. 1001-1010.

Leitão, A. (1987), 'Equidade na saúde', Saúde em Números, Vol. 2, pp. 9-12.

Lerman, R., and S. Yitzhaki (1984), 'A note on the calculation and interpretation of the Gini index' Economic Letters, Vol. 15, pp. 363-368.

Lerman, R., and S. Yitzhaki (1989), 'Improving the accuracy of estimates of Gini coefficients' Journal of Econometrics, Vol. 42, pp. 43-47.

Lindsay, C.M. (1969), 'Medical care and the economics of sharing' Economica, Vol. 36, pp. 351-362.

Lockwood, M. (1988), 'Quality of life and resource allocation' In: J.M. Bell and S. Mendus (*eds.*) Philosophy and Medical Welfare, Cambridge University Press, Cambridge.

Loewy, E.H. (1987), 'Communities, obligations and health care' Social Science and Medicine, Vol. 25, pp. 783-791.

Long, S.H., M.A. Cooke, and J.P. Crozier (1992), 'Income redistribution under national health insurance financing alternatives', In: J. van der Gaag, W.B. Neenan and T. Tsukahara (*eds.*) Economics of Health Care. Praeger, New York.

Lucas, J.S. (1986), 'Portugal: a research note' In: R. Illsley and P.G. Svensson (*eds.*) The Health Burden of Social Inequities, WHO Regional Office for Europe, Copenhagen.

Maasoumi, E. (1986), 'The measurement and decomposition of multi-dimensional inequality' Econometrica, Vol. 54, pp. 991-997.

Macintyre, S., and P. West (1991), 'Lack of class variation in health in adolescence: an artifact of an occupational measure of social class?' Social Science and Medicine, Vol. 32, pp. 395-402.

Mahalanobis, P.L. (1960), 'A method of fractile graphical analysis' Econometrica, Vol. 28, pp. 325-51.

Mantas, A., C. Costa and F.Ramos (1991) 'Equidade e eficiência no financiamento hospitalar' In: J. Pereira, M.R. Giraldes and A.C. Campos (*eds.*) Desigualdade e Saúde em Portugal. IED, Lisboa.

Margolis, H. (1982), Selfishness, Altruism and Rationality. Cambridge University Press, Cambridge.

Maynard, A. (1982), 'The regulation of public and private health care markets' In: G. McLachlan and A. Maynard (*eds.*) The Public-Private Mix for Health, The Nuffield Provincial Hospitals Trust, London.

Maynard, A. (1983), 'The production of health and health care' Journal of Economic Studies, Vol. 10, pp. 31-45.

Maynard, A. (1993), 'Foreword' In: E. Van Doorslaer, A. Wagstaff and F. Rutten, *eds.*, Equity in the Finance and Delivery of Health Care: An International Perspective. Oxford University Press, Oxford.

McClements, L. (1978), The Economics of Social Security. Heinemann, London.

McDowell, I., and C. Newell (1987), Measuring Health. A Guide to Rating Scales and Questionnaires. Oxford University Press, Oxford.

McLachlan, G., and A. Maynard, *eds.* (1982), The Public-Private Mix for Health. Nuffield Provincial Hospitals Trust, London.

Meade, J.E. (1976), The Just Economy: Principles of Political Economy. George Allen & Unwin, London.

Mendo, P. (1993), 'Política de saúde: não se pode fazer contra os médicos, mas não é dos médicos', Noticias Médicas, nº 2173, 17 December.

Mielch, A., and M.R. Giraldes, *eds.* (1993), Inequalities in Health and Health Care: Review of Selected Publications from 18 Western European Countries. Waxmann Mumster, New York.

Ministério da Saúde (1987), Inquérito Nacional de Saúde - 1987. 2º Volume, Metodologia. DEPS, Lisboa.

Mooney, G.H. (1983), 'Equity in health care: confronting the confusion' Effective Health Care, Vol. 1, pp. 179-185.

Mooney, G.H. (1986), Economics, Medicine and Health Care. Wheatsheaf Books, Brighton.

Mooney, G.H. (1987), 'What does equity in health mean?' World Health Statistics Quarterly, Vol. 40, no. 4.

Mooney, G., and A. McGuire (1987), 'Distributive justice with special reference to geographical inequality and health care' In: A. Williams (*ed.*) Health and

Economics. MacMillan, London.

Mooney, G., J. Hall, C. Donaldson, and K. Gerard (1991), 'Utilization as a measure of equity: Weighing heat' Journal of Health Economics, Vol. 10, pp. 475-480.

Morgan, J. (1962), 'The anatomy of income distribution' The Review of Economics and Statistics, Vol. 47, pp. 185-189.

Muellbauer, J. (1987), 'Professor Sen on the standard of living' In: G. Hawthorne (*ed.*) The Standard of Living. Cambridge University Press, Cambridge.

Musgrave, R., K. Case, and H. Leonard (1974), 'The distribution of fiscal burdens and benefits' Public Finance Quarterly, Vol. 2, pp. 259-311.

Muth, R. (1966), 'Household production and consumer demand functions' Econometrica, Vol. 34.

Muurinen, J.-M. (1982a), *An Economic Model of Health Behaviour with Empirical Applications to Health Survey Data.* Unpublished DPhil Thesis, Department of Economics and Related Studies, University of York.

Muurinen, J.-M. (1982b), 'Demand for health: a generalized Grossman model' Journal of Health Economics, Vol. 1, pp. 5-28.

Muurinen, J.-M., and J. Le Grand (1985), 'The economic analysis of inequality in health' Social Science and Medicine, Vol. 20, pp. 1029-1035.

NCHS (1967), 'US National Health Survey health interview responses compared with medical records' Vital and Health Statistics, Vol. 2, Public Health Service, Washington.

Nelson, J.A. (1993), 'Household equivalence scales: theory versus policy?' Journal of Labour Economics, Vol. 11, pp. 471-493.

Nersesian, W.S. (1988), 'Infant mortality in socially vulnerable populations' Annual Review of Public Health, Vol. 89, pp. 361-377.

Newman, J.F. (1975), 'Health status and utilization of physician services' In: R. Andersen, J. Kravits and O. Anderson (*eds.*) Equity in Health Services: Empirical Analysis in Social Policy Ballinger, Cambridge, Mass.

Nolan, B. (1993) 'Ireland' In: E. Van Doorslaer, A. Wagstaff and F. Rutten, *eds.*, Equity in the Finance and Delivery of Health Care: An International Perspective. Oxford University Press, Oxford.

Nolan, B., and T. Callan (1992), 'Cross-national poverty comparisons using relative poverty lines: an application and some lessons' In: D. Slottje (*ed.*), Research on Economic Inequality, Vol. 3. Jai Press, Greenwich, Connecticut.

Nozick, R. (1974), Anarchy, State and Utopia. Basic Books, New York.

O'Donnell, O., and C. Propper (1991a), 'Equity and the distribution of U.K. National Health Service resources' Journal of Health Economics, Vol. 10, pp. 1-19.

O'Donnell, O., and C. Propper (1991b), 'Reply: Equity and the distribution of U.K. National Health Service resources' Journal of Health Economics, Vol. 10, pp. 247-249.

O'Donnell, O., C. Propper, and R. Upward (1993), 'United Kingdom' In: Van Doorslaer, E., A. Wagstaff and F. Rutten, *eds.*, Equity in the Finance and Delivery of Health Care: An International Perspective. Oxford University Press, Oxford.

O'Higgins, M. (1987), 'Egalitarians, equalities and welfare evaluation' Journal of Social Policy, Vol. 16, pp. 1-18.

O'Higgins, M., and S. Jenkins (1989), 'Poverty in Europe: estimates for 1975, 1980 and 1985' Analysing Poverty in the European Community, Eurostat, Luxembourg.

O'Higgins, M., G. Schmaus, and G. Stephenson (1989), 'Income distribution and redistribution: A microdata analysis for seven countries' Review of Income and Wealth, Vol. 35, pp. 107-131.

OECD (1982), The OECD List of Social Indicators. OECD, Paris.

OECD (1984), Reviews of National Policies for Education: Portugal. OECD, Paris.

OECD (1985), Measuring Health Care 1960-1983. Expenditures, Costs and Performance. OECD, Paris.

OECD (1987), Financing and Delivering Health Care: A Comparative Analysis of OECD Countries, OECD, Paris.

OECD (1988), OECD Economic Surveys. Portugal. OECD, Paris.

OECD (1991), OECD Economic Surveys. Portugal. OECD, Paris.

OECD (1993), OECD Health Systems. Facts and Trends 1960-91, Vol.1 and Vol.2, OECD, Paris.

Ordem dos Médicos (1994), 'Ciclo de debates sobre política de saúde' O Médico.

Paneth, N., S. Wallenstein, J.L. Kiely, and M. Susser (1982), 'Social class indicators and mortality in low birth weight infants' American Journal of Epidemiology, Vol. 116, pp. 364-375.

Pazner, E.A., and D. Schmeidler (1978), 'Egalitarian equivalent allocations: a new concept of economic equity' Quarterly Journal of Economics, Vol. 92, pp. 671-687.

Pechman, J. (1985), Who Paid the Taxes, 1966-85. Brookings Institution, Washington DC.

Pechman, J. and B. Okner (1974), Who Bears the Tax Burden?. Brookings Institution,

Washington DC.

Pereira, J. (1990a), 'Equity objectives in Portuguese health policy' Social Science and Medicine, Vol. 31, pp. 91-94.

Pereira, J. (1990b), 'La interpretacion economica de equidad en la salud y atenciones sanitarias' In: J. Montserrat e C. Murillo (*eds.*) Salud y Equidad, Ministerio de Sanidad y Consumo, Madrid.

Pereira, J. (1992), 'Horizontal equity in the delivery of health care in Portugal' Revista Portuguesa de Saúde Pública, Vol. 10, pp. 35-46.

Pereira, J., *ed.* (1994) As Despesas com a Saúde em Portugal. Apifarma, Lisboa (in print).

Pereira, J., A.C. Campos, M.F. Cortês, and C. Costa (1987), 'Equidade geográfica no sistema de saúde português' In: A. C. Campos and J. Pereira (*eds.*) Sociedade, Saúde e Economia. ENSP, Lisboa.

Pereira, J., and C.G. Pinto (1992), 'Equidade no financiamento da saúde em Portugal' Economia, Vol. 16, pp. 105-135.

Pereira, J., and C.G. Pinto (1993), 'Portugal', In: E. Van Doorslaer, A. Wagstaff and F. Rutten, *eds.*, Equity in the Finance and Delivery of Health Care: An International Perspective. Oxford University Press, Oxford.

Pereirinha, J. (1988), *Inequalities, Household Income Distribution and Development in Portugal.* Unpublished PhD Thesis, Institute of Social Studies, The Hague.

Pfähler, W. (1983), 'Measuring redistributional effects of tax progressivity by Lorenz curves' Jahrbücher für Nationalökonomie und Statistik, Vol. 198, pp. 237-249.

Pfähler, W. (1987), 'Redistributive effects of tax progressivity: Evaluating a general class of aggregate measures' Public Finance/Finances Publiques, Vol. 42, pp. 1-31.

Phelps, C.E. (1973), The Demand for Health Insurance: A Theoretical and Empirical Investigation. Rand Corporation, Santa Monica.

Pinto, C.G. (1988), 'A incidência da despesa pública em cuidados de saúde na Área Metropolitana de Lisboa em 1980/81 e 1985' Estudos de Economia, Vol. 8, pp. 389-409.

Pinto, C.G., and J.C.G. Santos (1993), 'Incentivos fiscais e equidade no financiamento da saúde em Portugal' Paper presented to the Third Meeting of the Portuguese Health Economics Association, APES, Lisboa.

Plotnick, R. (1981), 'A measure of horizontal inequity' Review of Economics and Statistics, Vol. 63, pp. 283-88.

Pollak, R., and T. Wales (1979), 'Welfare comparisons and equivalence scales' American Economic Review, (Papers and Proceedings), Vol. 69, pp. 216-221.

Portugal (1992), 'Grandes Opções do Plano para 1993' Lei nº 30-B/92 de 28 de Dezembro. Diário da República, I Série - A, nº 298.

Portugal (1987), Programa do XI Governo Constitucional, Presidência do Conselho de Ministros, Lisboa.

Portugal (1988), Reforma do Sistema Retributivo da Função Pública, Presidência do Conselho de Ministros, Lisboa.

Poterba, J. (1989), 'Lifetime incidence and the distributional burden of excise taxes' American Economic Review (Papers and Proceedings), Vol. 79, pp. 325-330.

Power, C., O. Manor, and J. Tox (1991), Health and Class: The Early Years. Chapman & Hall, London.

Propper, C., and R. Upward (1992), 'Need, equity and the NHS: Distribution of health care expenditure 1974-87' Fiscal Studies, Vol. 13, pp. 1-21.

Puffer, F. (1986), 'Access to primary care: A comparison of the U.S. and U.K.' Journal of Social Policy, Vol. 15, pp. 293-313.

Pyatt, G. (1976), 'On the interpretation and disaggregation of Gini coefficients' Economic Journal, Vol. 86, pp. 243-255.

Pyatt, G. (1980), 'Poverty and welfare measures based on the Lorenz curve' Development Research Centre, World Bank, Washington, D.C. (mimeo).

Pyatt, G. (1985), 'An axiomatic approach to the Gini coefficient and measurement of welfare' In: R.L. Basmann and G.F. Rhodes, (*eds.*) Advances in Econometrics: Volume 4 (Economic Inequality: Measurement and Policy). Jai Press, Greenwich, Connecticut.

Rainwater, L. (1992), 'Inequalities in the well-being of children in ten nations' In: D. Slottje (*ed.*), Research on Economic Inequality, Vol. 3. Jai Press, Greenwich, Connecticut.

Rawls, J. (1971), A Theory of Justice. Harvard University Press, Cambridge, Mass.

Reynolds, M., and E. Smolensky (1977), Public Expenditure, Taxes, and the Distribution of Income. Academic Press, New York.

Rodrigues, C.F. (1993), 'Measurement and decomposition of inequality (1980/81 - 1989/90)' CISEP/ISEG, Documento de Trabalho nº 1/93.

Rodríguez, M., S. Calonge, and J. Reñé (1992), 'Analisis de la (in)equidad en la financiación y la provision de los servicios en el sistema sanitario español' In: B. Cloquell Rodrigo, M.A. Garcia Alonso, and A. Yañez Alonso (*eds.*) El Sistema Nacional de Salud de los 90. Generalitat Valenciana, Valencia.

Roemer, J. (1986), 'Equality of resources implies equality of welfare' Quarterly Journal of

Economics, Vol. 101, pp. 751-784.

Rosen, S. (1987), 'Human Capital' In: J. Eatwell, M. Milgate and P. Newman (*eds.*) The New Palgrave: A Dictionary of Economics. MacMillan, London.

Runciman, W.G. (1966), Relative Deprivation and Social Justice. Routledge and Kegan Paul, London.

Santos, B.S., and P. Hespanha (1987), 'O Estado a Sociedade e as políticas sociais: o caso das políticas de saúde' Revista Crítica de Ciências Sociais, nº. 23.

Santos, J. (1984), 'Escalas de equivalência' Estudos de Economia, Vol. 5, pp. 43-65.

Schokkaert, E., and L. Van Ootegen (1990), 'Sen's concept of the living standard applied to the Belgian unemployed' Recherches Economiques de Louvain, Vol. 56, pp. 429-450.

Schieber, G., and J.P. Poullier (1991) 'International health care spending: issues and trends' Health Affairs, Vol. 10, pp. 106-116.

Schiepers, J. (1992), 'On the choice of equivalence scales' Statistical Journal of the United Nations, Vol. 9, pp. 111-124.

Sen, A. (1973), On Economic Inequality. Clarendon Press, Oxford.

Sen, A. (1974), 'Informational bases of alternative welfare approaches: Aggregation and income distribution' Journal of Public Economics, Vol. 3, pp. 387-403.

Sen, A. (1978), 'Ethical measurement of inequality: some difficulties' In: W. Krelle and A.F. Shorrocks (*eds.*) Personal Income Distribution. North Holland, Amsterdam.

Sen, A. (1980), 'Equality of what?' In: Choice, Welfare and Measurement, Basil Blackwell, Oxford, 1982, pp. 353-369.

Sen, A. (1982), Choice, Welfare and Measurement, Basil Blackwell, Oxford, 1982.

Sen, A. (1983), 'Poor relatively speaking' Oxford Economic Papers, Vol. 35, pp. 153-169.

Sen, A. (1985), Commodities and Capabilities. North Holland, Amsterdam.

Sen, A. (1987a), 'Justice' In: J. Eatwell, M. Milgate and P. Newman (*eds.*) The New Palgrave: A Dictionary of Economics. MacMillan, London.

Sen, A. (1987b), 'The Standard of Living' In: G. Hawthorne (*ed.*) The Standard of Living. Cambridge University Press, Cambridge.

Sen, A. (1990), 'Welfare, freedom and social choice: a reply' Recherches Economiques de Louvain, Vol. 56, pp. 451-485.

Sen, A. (1991), 'Welfare preference and freedom' Journal of Econometrics, Vol. 50, pp. 15-29.

Sen, A. (1992), Inequality Reexamined. Clarendon Press, Oxford.

Shalit, H. (1985), 'Calculating the Gini index of inequality for individual data' Oxford Bulletin of Economics and Statistics, Vol. 47, pp. 185-189.

Shorrocks, A.F. (1980), 'The class of additively decomposable inequality measures' Econometrica, Vol. 48, pp. 613-625.

Silber, J. (1989), 'Factor components, population subgroups and the computation of the Gini index of inequality' Review of Economics and Statistics, Vol. 71, pp. 107-115.

Sloan, F.A. (1979), 'Toward measures of equity in the delivery of health care services' In: W.B. Rogers and D. Barnard (*eds.*) Nourishing the Humanistic in Medicine. University of Pittsburgh Press, Pittsburgh.

Sloan, F., and J.D. Bentkover (1979), Access to Ambulatory Care and the U.S. Economy. Lexington Books, Lexington, Mass.

Slottje, D. (1991), 'Measuring the quality of life across countries' Review of Economics and Statistics, Vol. 73, pp. 684-693.

Smeeding, T., and G. Schmaus (1990), 'The LIS database: technical and methodological aspects' In: T. Smeeding, M. O'Higgins and L. Rainwater (*eds.*) Poverty, Inequality and Income Distribution in Comparative Perspective. Harvester Wheatsheaf, London.

Smith, A. (1987), 'Qualms about QALY's' The Lancet, Vol. 1, p. 1134.

Stockwell, E.G., D.A. Swanson, and J.W. Wicks (1988), 'Economic status differences in infant mortality by cause of death' Public Health Report, Vol. 103, pp. 135-142.

Strotz, R.H. (1955/6), 'Myopia and inconsistency in dynamic utility maximization' Review of Economic Studies, Vol. 23, pp. 165-180.

Sugden, R. (1993), 'Welfare, resources and capabilities: A review of *Inequality Reexamined* by Amartya Sen' Journal of Economic Literature, Vol. 31, pp. 1947-1962.

Suits, D. (1977), 'Measurement of tax progressivity' American Economic Review, Vol. 67, pp. 747-752.

Tanzi, V., and L. de Wulf (1976), 'A distribuição da carga fiscal por grupos de rendimento em Portugal' In: I Conferência Internacional sobre Economia Portuguesa, Fundação Calouste Gulbenkian, Lisboa.

Teekens, R. (1990), 'Inequality and poverty. Portugal compared with Greece, Ireland and Spain' Estudos de Economia, Vol. 10, pp. 111-142.

Theil, H. (1967), Economics and Information Theory. North Holland, Amsterdam.

Tobin, J. (1970), 'On limiting the domain of inequality' Journal of Law and Economics, Vol. 13, pp. 263-277.

Tooley, W.H. (1966), 'Newborn intensive care' In: Proceedings of AMA National Conference on Infant Mortality, Washington.

UNDP (1990), Human Development Report. Oxford University Press, Oxford.

Urbano, J., M. Bentes, and J.C. Vertees (1993), 'Portugal: National Commitment and the implementation of DRG's In: J. Kimberly and G. de Ronvonville (*eds.*) The Migration of Managerial Innovation, San Francisco.

Van Doorslaer, E., R. Janssen, A. Wagstaff, J. van Emmerik, and F. Rutten (1991), 'Equity in the finance of health care: Effects of the Dutch Health Insurence Reform' In: G. López Casanovas (*ed.*) Incentives in Health Systems. Springer-Verlag, Berlin.

Van Doorslaer, E., A. Wagstaff, and others (1992), 'Equity in the delivery of health care: some international comparisons' Journal of Health Economics, Vol. 11, pp. 389-411.

Van Doorslaer, E., A. Wagstaff and F. Rutten, *eds.* (1993), Equity in the Finance and Delivery of Health Care: An International Perspective. Oxford University Press, Oxford.

Van Doorslaer, E., A. Wagstaff, and R. Janssen (1993), 'The Netherlands' In: Van Doorslaer, E., A. Wagstaff and F. Rutten, *eds.*, Equity in the Finance and Delivery of Health Care: An International Perspective. Oxford University Press, Oxford.

Varian, H.R. (1974), 'Equity, envy and efficiency' Journal of Economic Theory, Vol. 9, pp. 63-91.

Wagstaff, R.A. (1985), *Unemployment and Health: An Economic Analysis.* Unpublished DPhil Thesis, Department of Economics and Related Studies, University of York

Wagstaff, A. (1986a), 'The demand for health: theory and applications' Journal of Epidemiology and Community Health, Vol. 40, pp. 1-11.

Wagstaff, A. (1986b), 'The demand for health: some new empirical evidence' Journal of Health Economics, Vol. 5, pp. 195-233.

Wagstaff, A. (1991), 'QALY's and the equity-efficiency trade-off' Journal of Health Economics, Vol. 10, pp. 21-41.

Wagstaff, A., and E. Van Doorslaer (1994), 'Measuring inequalities in health in the presence of multiple-category morbidity indicators' Health Economics, Vol. 3, pp. 281-291.

Wagstaff, A., E. Van Doorslaer, and P. Paci (1989), 'Equity in the finance and delivery of health care: some tentative cross-country comparisons' Oxford Review of Economic Policy, Vol. 5, pp. 89-112.

Wagstaff, A., E. Van Doorslaer, and P. Paci (1991a), 'On the measurement of inequalities in health' Social Science and Medicine, Vol. 33, pp. 545-547.

Wagstaff, A., E. Van Doorslaer, and P. Paci (1991b), 'On the measurement of horizontal inequity in the delivery of health care' Journal of Health Economics, Vol. 10, pp. 169-205.

Wagstaff, A., E. Van Doorslaer, and P. Paci (1991c) 'Horizontal equity in the delivery of health care' Journal of Health Economics, Vol. 10, pp. 251-256.

Wagstaff, A., E. Van Doorslaer, and others (1992), 'Equity in the finance of health care: some international comparisons' Journal of Health Economics, Vol. 11, pp. 361-387.

Wagstaff, A., E. Van Doorslaer, and others (1994), 'The redistributive effect of health care finance: some preliminary cross-country comparisons' Paper presented at the Seminar on Income Distribution, Taxes and Benefits, LSE, December.

West, P. (1986), 'What is wrong with QALY's?' Paper presented to a meeting of the Health Economists Study Group, University of York, January.

West, P. (1988), 'Inequalities? Social class differentials in health in British youth' Social Science and Medicine, Vol. 27, pp. 291-296.

Whalley, J. (1984), 'Regression or progression: the taxing question of incidence analysis' Canadian Journal of Economics, Vol. 17, pp. 654-682.

Whitehead, M. (1987), The Health Divide: Inequalities in Health in the 1980's. Health Education Council, London.

WHO (1970), Technical Report Series N° 440, WHO, Geneva.

WHO (1981), Health Services in Europe. Volume 2: Country Reviews and Statistics, WHO Regional Office for Europe, Copenhagen.

WHO (1985a), World Health Statistics Annual-1985, World Health Organization, Geneva.

WHO (1985b) Targets for Health for All. WHO Regional Office for Europe, Copenhagen.

WHO (1987), Meeting of Principal Investigators on Inequities in Health. Lisbon, 28-30, September 1987, Escola Nacional de Saúde Pública.

Wilkinson, R.G., *ed.* (1986), Class and Health: Research and Longitudinal Data. Tavistock Publications, London.

Williams, A. (1984), 'Human capital and inequalities in health' Seminar presented to Department of Economics, Lund University.

Williams, A. (1988a), 'Ethics and efficiency in the provision of health care' In: J.M. Bell and S. Mendus (*eds.*) Philosophy and Medical Welfare. Cambridge University Press, Cambridge.

Williams, A. (1988b), '... Makes a man healthy, wealthy and wise! (*or* from folklore to system science)' In: G. Duru *et al* (*eds.*) System Science in Health Care, Vol. 2. Masson, Paris, pp. 57-60.

Williams, B. (1987), 'The standard of living: interests and capabilities' In: G. Hawthorne (*ed.*) The Standard of Living. Cambridge University Press, Cambridge.

Yitzhaki, S. (1979), 'Relative deprivation and the Gini coefficient' Quarterly Journal of Economics, Vol. 93, pp. 321-324.

Yitzhaki, S. (1983), 'On an extension of the Gini index' International Economic Review, Vol. 24, pp. 617-628.

Zweifel, P. (1989), 'Consumer incentives in health care' Paper Presented to the First European Conference on Health Economics. Barcelona, September.