

THE UNIVERSITY *of York*

CENTRE FOR HEALTH ECONOMICS
YORK HEALTH ECONOMICS CONSORTIUM
NHS CENTRE FOR REVIEWS & DISSEMINATION

Equality of *What* in Health Policy? Conflicts Between the Contenders

A.J. Culyer

DISCUSSION PAPER 142

**EQUALITY OF *WHAT* IN HEALTH POLICY? CONFLICTS BETWEEN THE
CONTENDERS**

A J Culyer

**Department of Economics and Related Studies,
University of York, Heslington,
York YO1 5DD**

November 1995

The Author

Tony Culyer is Professor and Head of the Department of Economics and Related Studies, University of York.

Acknowledgments

This paper was given at the Third European Conference on Health Economics in Stockholm, 1995. The author would like to thank the discussants, Jaoa Pereira and Terkel Christianson, for their comments and also other participants for their helpful comments. This paper is part of an on-going collaboration between Culyer and Adam Wagstaff (University of Sussex). Wagstaff's comments are also gratefully acknowledged.

Further copies

Further copies of this document are available (at a price of £4.00 to cover the cost of publication, postage and packing) from:

The Publications Office
Centre for Health Economics
University of York
York YO1 5DD

Please make cheques payable to the University of York. Details of other papers can be obtained from the same address, or telephone York (01904) 433648 or 433666.

ABSTRACT

Most people agree that equity has something to do with equality - but equality of what? This paper discusses five relevant "respects" according to which horizontal or vertical equity for individuals or groups of people might be achieved: marginal willingness to pay, per capita expenditure, need as ill-health, need as capacity to benefit, and need as the resources required to exhaust capacity to benefit. They are shown, in general, to conflict with one another (though not with cost-effectiveness) and also to conflict with a sixth, and in my view preferable, relevant respect whose object is to promote greater equality in the distribution of health in the community. This sixth criterion does not necessarily conflict with cost-effectiveness either. A seventh principle, equality of access, is also shown to be inadequate.

The sources of these equity conflicts, in a health care system that maximises health gain given a stock of resources, are interactions between differences in people's initial health endowments, their wealth, their preferences, their health production functions, and the budget constraint, a change in any one of which *ceteris paribus* may change the distribution of both health and health gain.

The general conclusion is that *cheapness* (rather than equality) in initial accessibility of health care is a necessary condition for equity but that *inequality* in resource distribution is also generally required to achieve equity in health.

INTRODUCTION

This paper is in no way a comprehensive discussion of distributive justice in health and medical care (or any of the other determinants of health). Its purpose is much more limited and is, in essence, to demonstrate the inherent conflicts that exist between some common equity concepts and, hence the necessity of choice between them. It draws on the analysis developed recently by Adam Wagstaff and me (see Culyer 1993a, Culyer and Wagstaff 1993) but emphasises the consequences of adopting different concepts rather than advocating any one of them. I shall take it for granted that a concern about equitable as well as efficient allocations of resources is a legitimate topic for economic analysis (Culyer 1993b).

There are some ground rules which need to be made clear from the beginning. First, I shall assume, in the extra-welfarist tradition (Culyer 1989), that the efficiency maximand of the medical care system is health. Second, in that same tradition, I shall assume that, whatever else is sought in the way of equitable distributions, an equitable distribution of *health* is also an objective. I shall *not* examine possible conflicts between these other objectives of an extra-welfarist social welfare function because I want to focus on conflicts that exist within the domain of health itself. However, I am going to assume that at least one commonly asserted conflict does *not* exist, viz, that between equity and efficiency. In this model, efficiency obtains at any point on a health frontier - the locus of health maxima defined by a given stock of resources available for medical care, the available technologies for transforming these resources into health outputs, and the initial endowments (of health) of individuals. This concept of efficiency can be seen to be a close relation of the Pareto criterion - efficiency exists when the health of no individual can be increased without reducing that of another. The equity problem is to select the preferred point on that frontier and I shall assume that we are free to move along the frontier without there be any secondary feedback effects that might cause it to shift.

I shall be completely conventional in my treatment of the Aristotelian principles of horizontal and vertical equity. I shall take horizontal equity to require the like treatment

(in terms of resources) of people who are alike in equity-relevant respects and vertical equity to require the proportionately unequal treatment of people who are unlike in equity-relevant respects. What I shall be exploring is the different outcomes that arise as one redefines these various possible "equity-relevant" respects. The first equity-relevant respect to be considered will be willingness to pay (the same willingness means you should get the same; willingness to pay more means that you should get more in the same proportion as your greater willingness). The second respect will be need in the sense of presenting or current sick state or, as I shall term it, endowment. The third respect will be straight equality of resource distribution on the ground that each has, in principle, some equal entitlement (there is no vertical version of this respect). The fourth respect will be need, in a sense I shall explain shortly. The fifth and final respect will be capacity to benefit, which some call need, but which I shall call capacity to benefit.

My conclusions will not be surprising to many of you. The first is that different equity rules, even apparently quite strongly egalitarian ones, can produce bizarrely unequal distributions of health - distributions that are more unequal than the endowment distributions which redistributive policy was intended to alter for the better. Hence the need to choose, which I take to involve a deeper enquiry into the more fundamental reasons for advocating the distributional principles to be discussed. I shall offer you my own preference but my aim is less to advocate my own view than to force you to recognise that there is much fence coming-off to be done. There are a number of other implications that I shall mention as we go along. The most important - and reassuring - to my mind is that it is *cheapness* of access to medical care that is of prime equity significance rather than equality of access and that, in general, both inequality of access and inequality in the distribution of medical care resources, seem to be required if one wants to achieve equity in the distribution of health. I shall note that the informational requirements of any policy for equity in health are fairly demanding, though I think that one can get quite far both by using the methods I propose as a taxonomising device, in order to get people such as members of health authorities (in the UK) to think more clearly about one of their duties, and through the use of quite crude empirical information supplemented by bold judgments (whose challenge can be expected to lead, of course, to

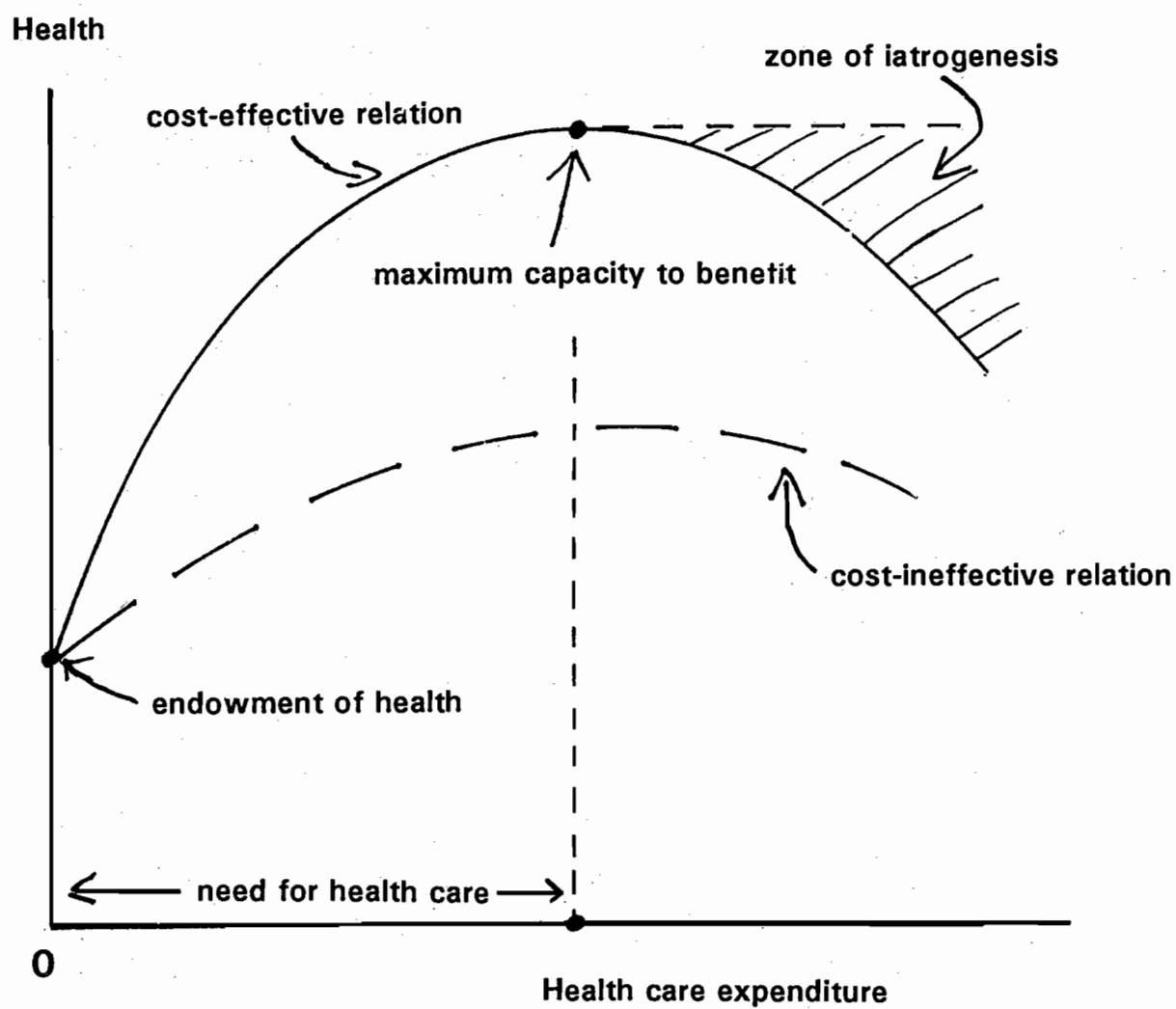
refinement and better future information bases). Finally, a reminder: I shall focus on medical care and health alone. I do not consider that this is adequate and would judge that the most effective methods of achieving real equity in health probably lie outside the medical sphere as usually interpreted, through manipulating other determinants like genetics, child rearing, and the general equalities and inequalities in the societies in which people live. Moreover, health is not the only good nor the only good whose equitable distribution ought to concern us. All these considerations are, however, off limits as far as this paper is concerned.

THE BASIC APPARATUS

Let us begin with the basic geometrical apparatus I propose to use. Figure 1 shows the simple production function we shall be using. The vertical axis is the health measure, which I shall take to be the present value of expected future QALYs for an individual. It is easiest to think of "health" in terms of a fully cardinal measure (viz. one on a ratio scale) and, provided that your preferred measure meets this heuristic desideratum, and is ethically relevant as a distributional concern, you may choose any measure you like. The horizontal axis measures medical care resources, here taken to be expenditure. The higher of the two output or production functions assumes full cost-effectiveness and I shall assume that the outcomes can be predicted with certainty. The lower function indicates a cost-ineffective locus arising perhaps out of X-inefficiency or monopolistic pricing policies. I am going to assume full cost-effectiveness, so the assumptions underlying the (higher) function are the usual ones made in the analysis of the competitive firm or industry. The functions embody strictly diminishing marginal products of health care in terms of health outcome. There is no strategic play in the model, and no discrimination or cream skimming. All producers are assumed to be health or health gain maximisers. If the medical care system is financed out of pocket or via insurance, there is no adverse selection or moral hazard.

The origin of the production function defines the endowment point for an individual, his or her presenting state (the inequities which may have brought this about are not a part of

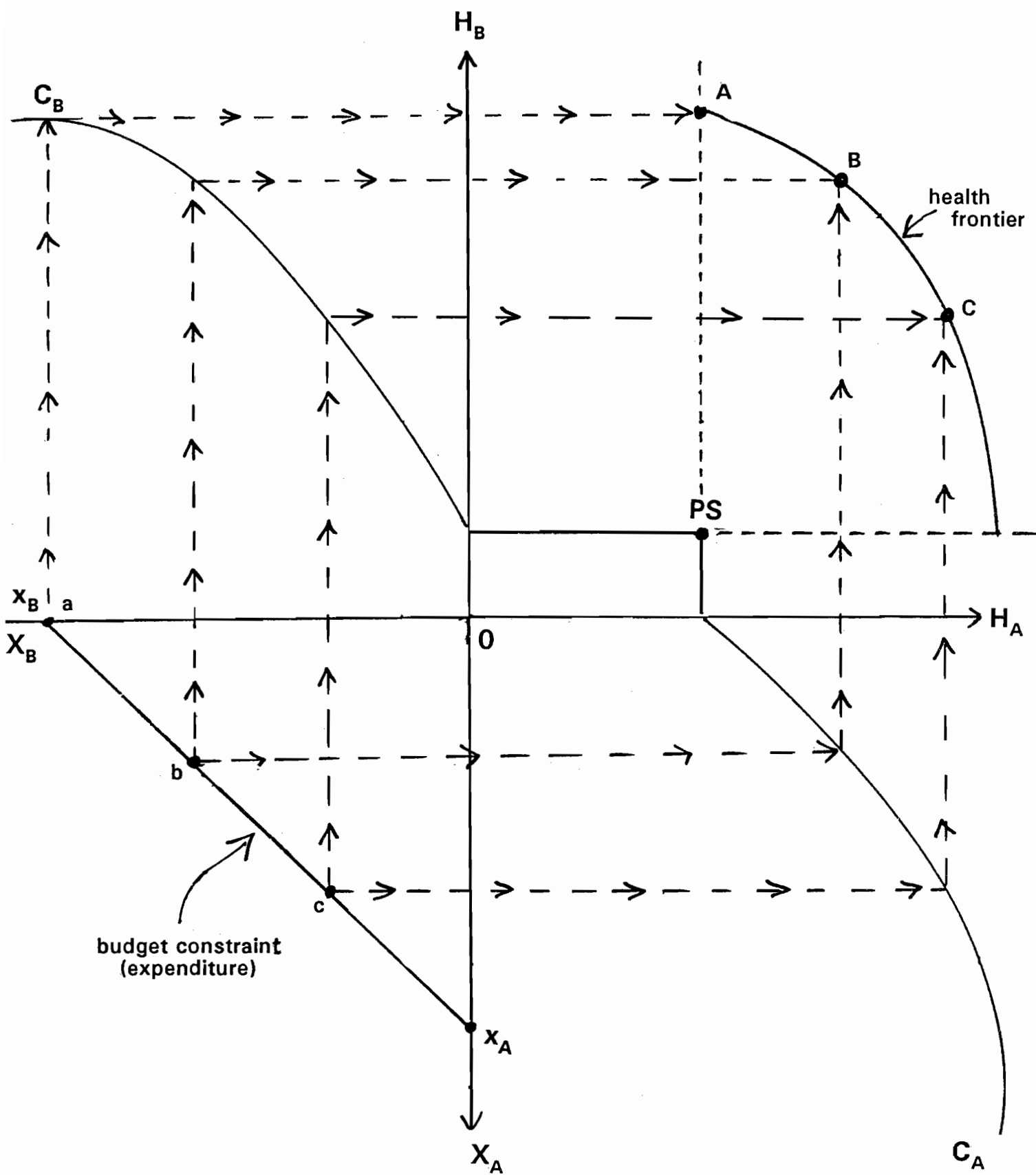
Figure 1 Health as a function of resources (expenditure)



the model). The point at which the marginal productivity of resources falls to zero is of course the maximum of the function which defines maximum capacity to benefit from medical care. In the Culyer/Wagstaff terminology all resources having a positive marginal product for this individual are *needed* by this individual in the sense that they are a necessary condition if the additional health is to be obtained, so the distance marked on the horizontal axis indicates this person's need for medical care. I shall mention the other notions of need later. I term the region in which there is a negative marginal product of medical care to be the zone of iatrogenesis and my assumptions imply that no producer will operate on this section of any individual's function.

Figure 2 is a two-person model of a society in which A and B each have a production function with the characteristics already described. The top left quadrant shows B's function and the bottom right quadrant shows A's. The bottom left quadrant shows the budget constraint; it is either the budget available to the medical care system or the sum of the expenditure chosen by individuals in a market. If the entire budget were expended on A, she would have x_A ; were it expended on B, she would have x_B . The endowments of A and B, their production functions, and the budget constraint together determine the strictly convex health frontier shown in the top right quadrant, whose axes are the health of A and B. The frontier is derived by tracing the maximum health outputs for each individual at a variety of distributions of the budget between them. For example, a division of the budget indicated by b implies point B on the frontier and c implies C . The frontier represents the maximum attainable health for each, given any division of the budget along the linear constraint in the bottom left quadrant. Along the frontier, the health of one can be increased only by reducing that of the other. You will see that I am not making any assumptions about the ways in which individual QALYs are to be added up. In some efficiency analyses, this is done by assuming some appropriate weighting system (for example, that a QALY is of equal social value whoever has it). But my idea of efficiency is more limited than this. The system is efficient when there is no way in which resources can be redeployed to increase the health of one, either in different production technologies or with same technologies but with different shares going to each individual, without diminishing the health gain of the other. In short, all points on

Figure 2 Constructing health frontier



the frontier are (Pareto-) efficient given the health maximand. The dotted axes in the top right quadrant have their origin at the endowments of the two individuals and define the scope for health gain, given extant technologies, the available resources, and cost-effective production techniques.

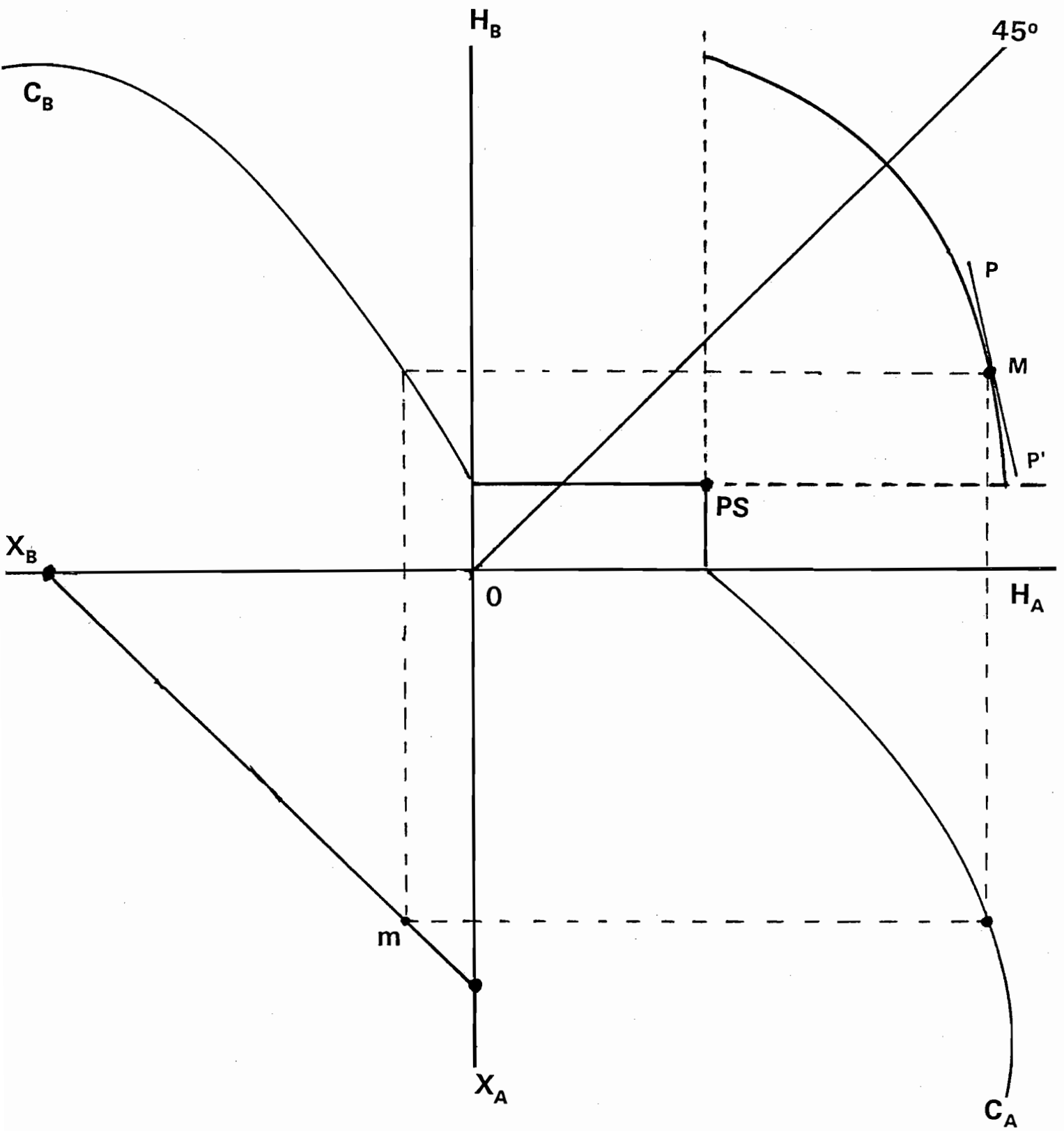
I now propose to use this simple apparatus to bring out the points made in the introduction. In what follows I shall consistently assume that individual A is both richer and at least initially healthier than B. The greater healthiness of A is plain from the location of the endowment point in the top right quadrant. I shall also assume that B, the less healthy poorer person, has a higher capacity to benefit from health care. Note also that A's *need* is larger than B's, despite B's greater capacity to benefit, a possibility that is not widely recognised.

ALLOCATION BY WILLINGNESS TO PAY

Although it is not perhaps usual to consider willingness to pay as an equity criterion, there are some streams of more or less libertarian literature which regard as equitable the outcomes of fair market procedures (even if purchasing power is unequal). Moreover, the "market solution" is a useful benchmark with which to compare the outcome of following other possible rules. So let us suppose that medical care is purchased in the market, with or without insurance, and that the rich will buy more of it.

This therefore implies the selection of some point such as *m* on the budget constraint in Figure 3, to the right of its midpoint, with A buying more than B. This corresponds to point *M* on the frontier. The 45° line in the top right quadrant shows equal distributions of health between A and B. From our assumptions the endowment point (*PS*) is not on this line, showing A to enjoy better health. *M* must lie to the right of the point at which the 45° line intersects the frontier. The point you see in Figure 3 shows the same proportionate distribution of health as existed at the endowment point, though it might lie either to the right or the left of this. The tangent at *M* shows the rate at which the health of A is traded against that of B. As shown, a small reduction of resources transferred

Figure 3 Distribution by willingness to pay



from A to B would yield an increase in the total number of QALYs were each to have a unitary weight. However, in this model, such a move does not represent an increase in efficiency, for M is as efficient as any other point on the frontier. Rather, it would represent a distributional judgment that one efficient point is more equitable than another. The implied distributional weight in Figure 3 suggests that a marginal additional QALY for A is worth something like 100 QALYs for B. And that might be judged to be equitable on procedural grounds provided that the procedures through which wealth was initially distributed and the procedures through which medical care was purchased were also in an appropriate sense "fair", and, of course, provided that the endowments could also be regarded as distributionally "fair".

The "budget" in the bottom left quadrant is, of course, a function of the preferences and wealth of the two individuals, and the market opportunities facing them. It is whatever total expenditure arises out of their decisions. I am now going to treat the expenditure outcome of these circumstances in the market situation as the budget that might be available to a system in which political decisions determine it, in order to compare other possible distributional principles.

So much, then, for the market solution.

DISTRIBUTION BY ENDOWMENTS

The first more conventional equity principle I want to explore is commonly termed distribution according to need, where "need" is taken to be the presenting states of health of individuals. I do not much care for this definition, since it is not at all clear that medical care is actually needed at all in the sense that its receipt is a necessary condition for improving health. After all, the productivity of medical care might be zero or negative for an individual. So I term this principle "distribution by rather than distribution by need" since that describes more accurately what it really is.

Figure 4 Distribution by health endowment endowments"

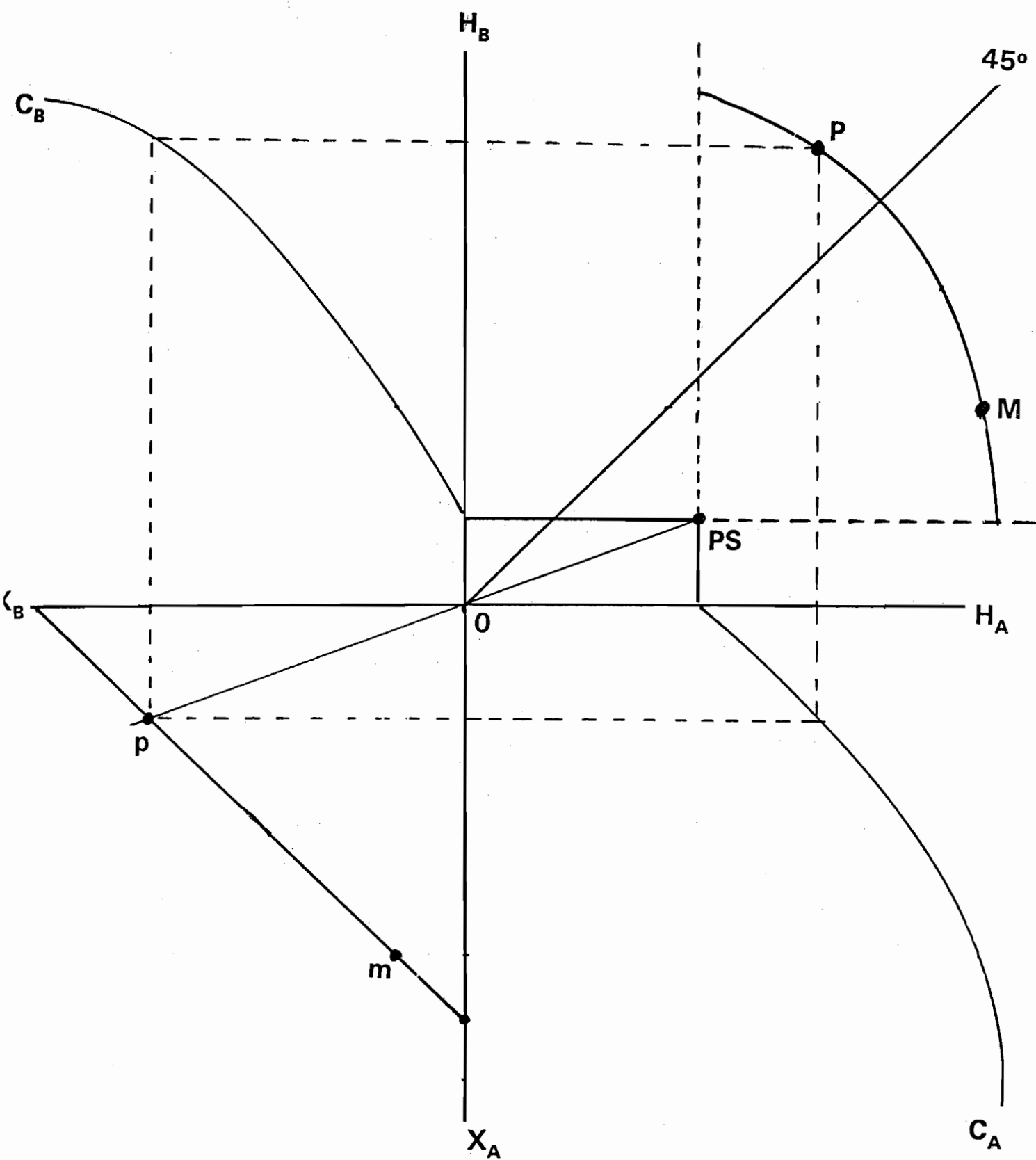
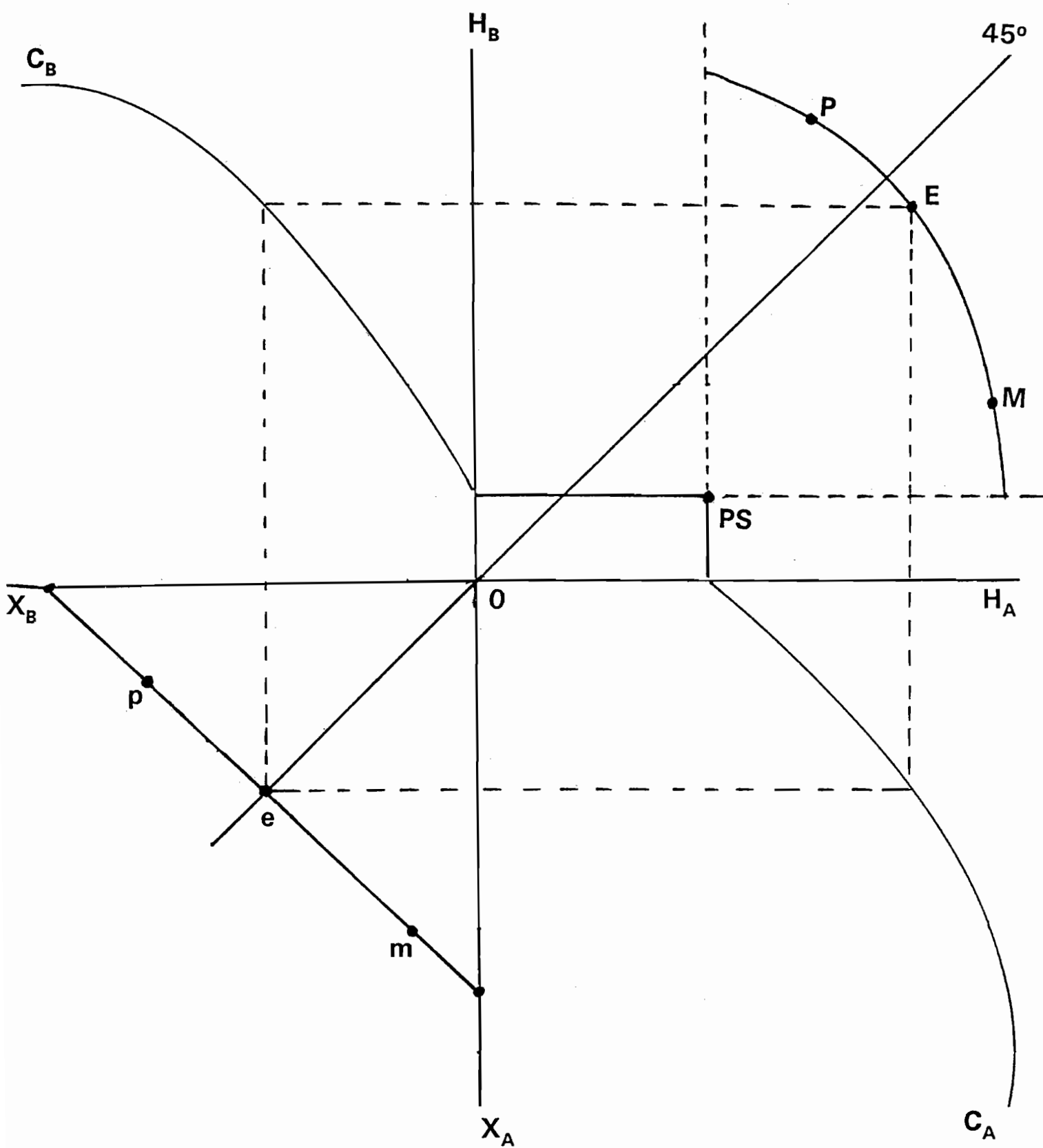


Figure 4 shows the consequences for the same two individuals of allocating resources in proportion to endowments or presenting states of ill-health as at PS . The line through the endowment point and the origin produced into the bottom left quadrant locates point p , at which each receives resources in *inverse* proportion to their endowment of health. This yields point P on the frontier. It is of necessity to the left of point M and, as shown, is also to the left of the 45° line. This illustrates a case of inequality reversal: in endowment, A was healthier than B; after receipt of medical care B is healthier than A. Now a marginal QALY for B trades at about two for A.

EQUAL DISTRIBUTION OF RESOURCES

Figure 5 shows the consequences of distributing resources along strictly egalitarian lines. e is the midpoint of the budget constraint in the bottom left quadrant and it implies point E on the frontier. E must lie between P and M but there is no reason for it to lie on the 45° line, nor at a point on the frontier at which its slope is equal to -1, so it will in general still imply that the marginal social value of a QALY is not the same for each individual; what the marginal social value is, is not an explicit policy choice, but is merely churned out by the interaction of the production functions and the endowment points, given the decision to allocate resources in strict equality. This equity rule would plainly be absurd if taken too literally, for it implies giving the same medical care to those who are as completely fit as they can be as to those who may actually need it (in any sense of need) or those who merely want it. But it is less absurd if the analysis is taken - as it is my guess that it is usually intended to be - at the level of representative individuals, or individuals who are really standing for reasonable well-defined groups within the community (like "the poor", "the rich", "rural dwellers" and so on). At a certain level of aggregation, the rule seems less obviously absurd. Moreover, it may be interpreted as one version of a very common equity rule: equality of access, in a regime where you may utilise only what is available and what is available is equally available to all client groups. Wagstaff and I have commented elsewhere on the limitations of some common approaches to access-equity (Culyer and Wagstaff 1993) and I propose no further

Figure 5 Distribution with equal expenditures



discussion here. My main intent is to bring at least one version of this type of principle into the analysis for comparison though I do not doubt that aficionados of this approach will (rightly) consider this treatment to be something of a travesty.

DISTRIBUTION BY RELATIVE NEED

Figure 6 shows the effects of distributing resources according to relative need (in the sense developed by Culyer and Wagstaff). Marginal need for each individual falls to zero when the individual production functions reach their maxima, which locates the point marked "total need" in the bottom left quadrant of Figure 6. A line through the origin and this point intercepts the budget line at point n , whose implied division of resources in turn implies the point N on the health frontier. The location of the total needs point is independent of endowments but, since the wealthier individual is here assumed to have the greater need, point N must lie to the right of the distribution of health implied by an equal sharing of health care resources (E) and, given that this lies to the right of the 45° line, must correspond to a distribution of health that is more unequal. Of course, although the relative needs of the two individuals are independent of endowments, the health frontier is not. Had, for example, the healthier individual been less healthy than here assumed, the frontier would be shifted parallel and to the left, so that both E and *a fortiori* N would lie further away from the 45° line. Indeed, with equal endowments of health, the upshot would be unequal, with N more unequal than E .

DISTRIBUTION ACCORDING TO CAPACITY TO BENEFIT

Capacity to benefit is used by some, it will be recalled, as a concept of need. Figure 7 shows the consequences of allocating resources in proportion to capacity to benefit. Total capacity to benefit is indicated by the point "total capacity to benefit" in the top right quadrant. Since B has the higher capacity to benefit of the two, the line through the endowment point PS and this point has a slope steeper than +1 and makes the angle α with the 45° line. This angle is used to construct the allocation of resources in the bottom

Figure 6 Distribution by relative need

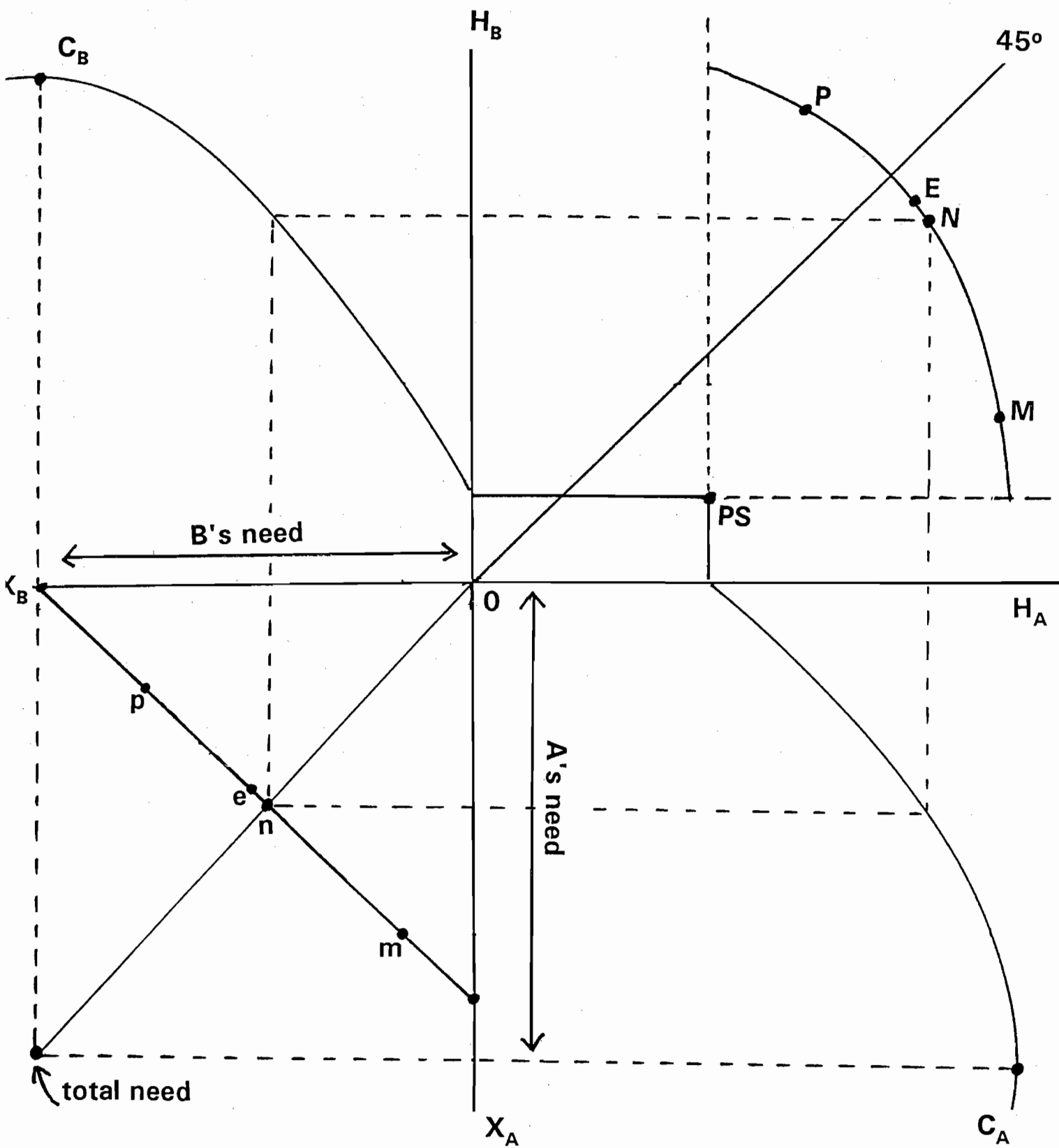
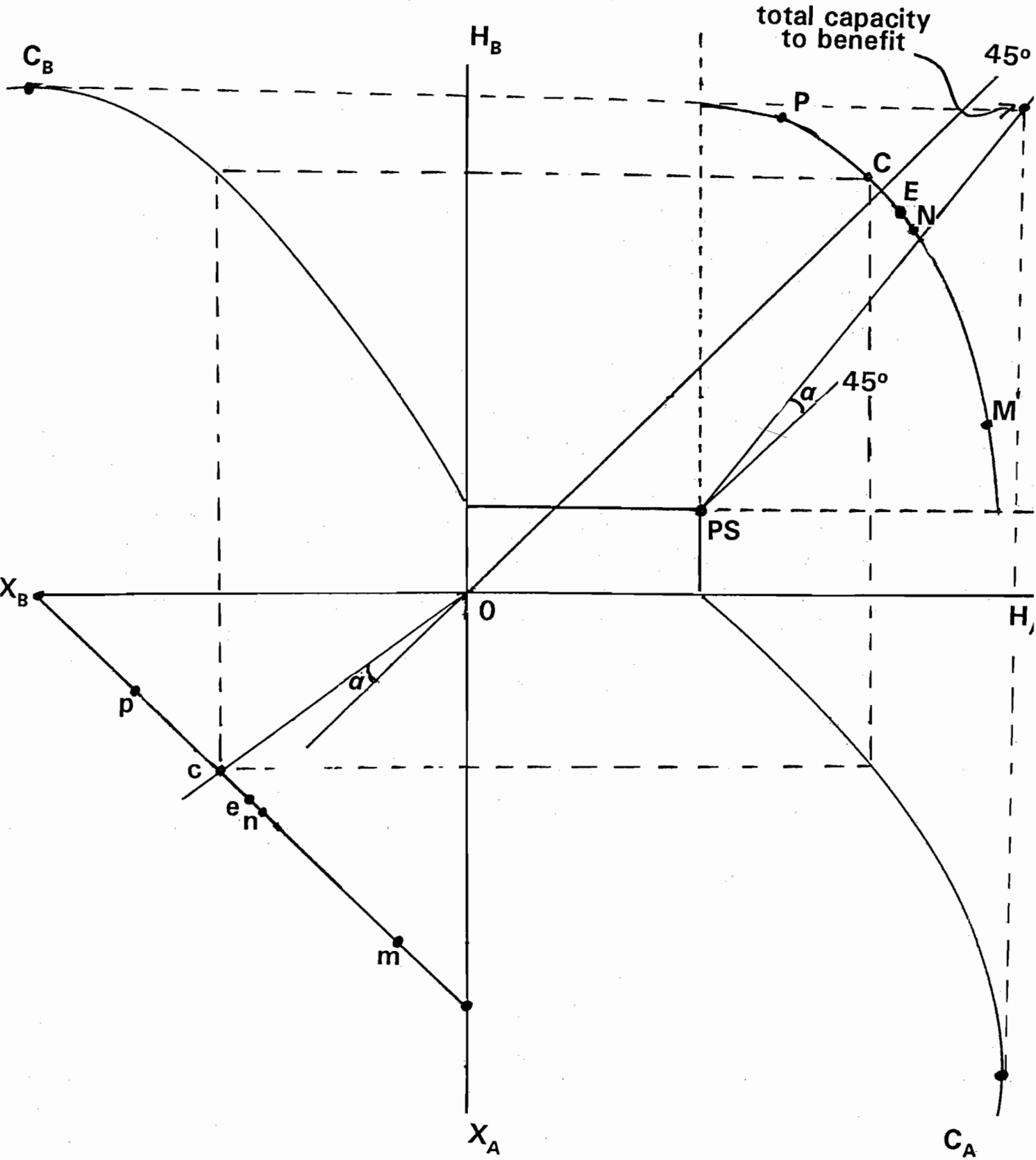


Figure 7 Distribution by relative capacity to benefit



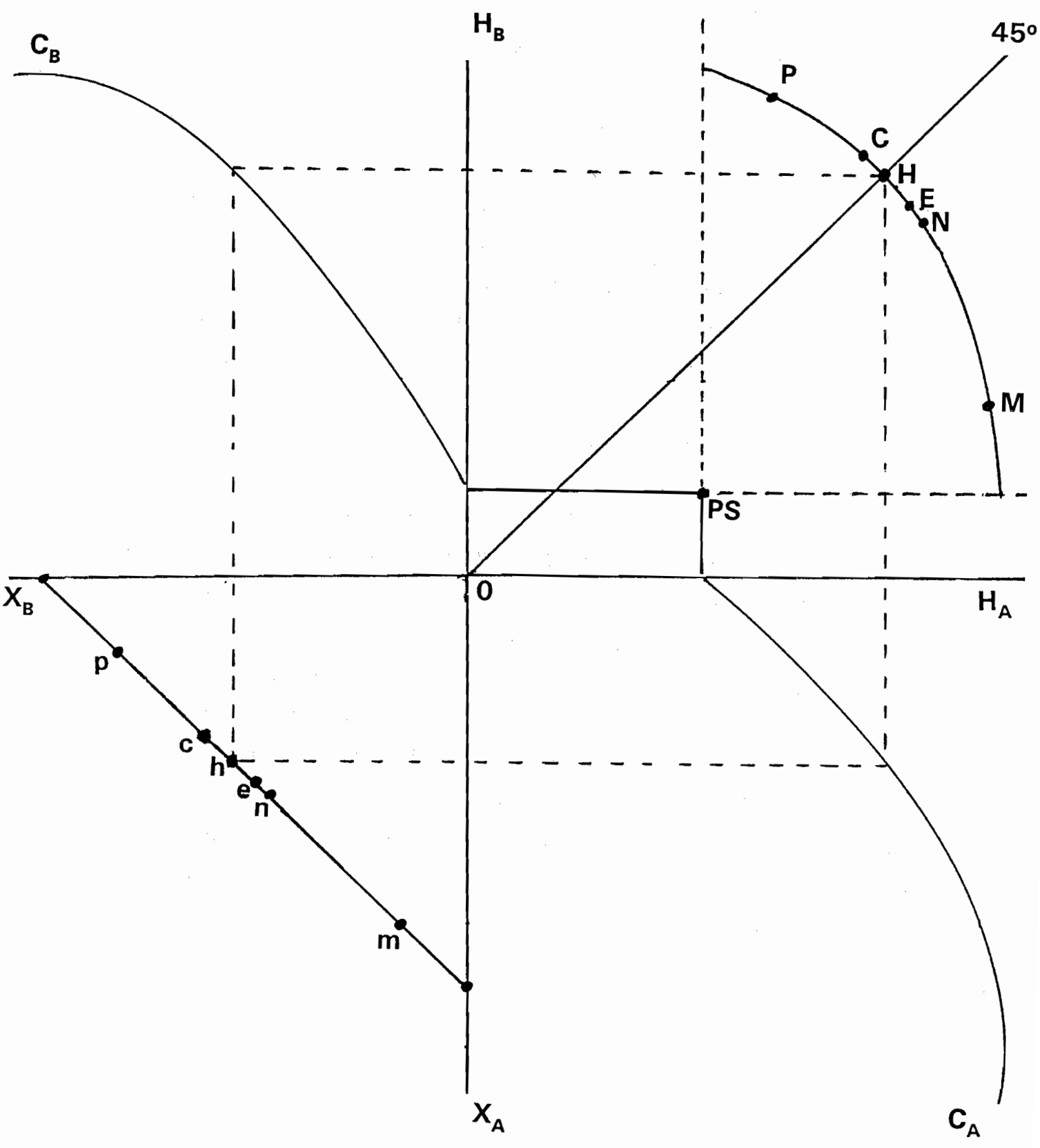
left quadrant, yielding the point c at the intersection with the budget line. This in turn implies point C on the health frontier, viz. a relative favouring of B and, in this case a reversal of the endowment inequality.

A variant of this principle would be to select that point on the health frontier at which its slope is equal to -1, indicating an equal social weight being attached to each QALY at the margin and thus maximising the total QALY count. Such a point would not necessarily coincide with C , since it is dependent on marginal products rather than the maximum relative capacities to benefit. Its location would identify, once tracked back to the bottom left quadrant, the equitable distribution of resources between individuals on the budget line. In Figure 7, as drawn, this point on the frontier would lie to the right of C , implying a smaller inequality reversal. But it could lie to its right.

DISTRIBUTION FOR EQUAL HEALTH

The final distribution to compare is that resulting from a policy aiming at equalising health. In this case, as with maximising the sum of QALYs, we start in the top right quadrant of Figure 8 but now identify the point at which the 45° line intersects the frontier - H . Working backwards, the distribution of resources required to achieve this outcome is found to be h on the budget line in the bottom left quadrant. The allocation is seen to favour B over A, on account of B's larger capacity to benefit, rather than existing relatively poor health. To illustrate, were the origin of A's production function to shift towards 0, the health frontier would shift to the left and a new H would be located on the 45° line to the south west of the original H . This would tend to shift the distribution of resources further in B's favour, even though B was now less relatively sick than before at the endowment point.

Figure 8 Distribution for equal health



SUMMARY AND CONCLUSIONS

Figure 8 also identifies all the points discussed, with the exception of the sum of QALYs maximum, and indicates the potential spread of distributions of health, together with inequality reversals, that might arise from following the various principles whose implications have been discussed. The relative positions of these points depend in the model, of course, on relative endowments, wealth distribution, the budget constraint, the slopes of the health production functions, and the maximum capacities to benefit.

Two considerations immediately leap to one's attention:

- (a) the treatment of *efficiency* as being concerned with identifying the health frontier, which I have termed "health maximisation", and the treatment of *equity* as being concerned with the selection of a point on it and
- (b) the need to consider whether there are more "ultimate" warrants for selecting points on the budget constraint, that is the equitable distribution of *health care resources*, which exclude consideration of the extant and subsequential distributions of *health* in the community.

My own view is that, although the distribution of health may not be the *only* factor to be taken into account, it has a strong claim to be a major one, and mainly on the rather Aristotelian grounds that health is normally a necessary condition for a person to be able to flourish as a human being and, if such opportunities for flourishing are to be made equitable, so then should the distribution of health. This would normally, I suggest, involve seeking as great an equality of health as is feasible (complete equality may not lie in the feasible set of the top right quadrant). I say "normally" on account of my desire to allow the possibility that other factors than the distribution of health might properly be considered in deciding the equitable distribution of resources, and also to allow for the probability that I would not be indifferent between alternative paths through which such an equality were attained - for example, I would not regard it as equitable to seek equality

in health by damaging the health of the healthy. It is perhaps worth noting that this procedure does not involve the selection of *a priori* desirable social weights for attaching to QALYs; the marginal trade off of QALYs for one against QALYs for another is a *consequence* of the equitable distribution of health, not a determinant of it.

The second general conclusion is to note the possibility of quite bizarre inequalities in health arising from the pursuit of apparently egalitarian resource distribution policies. In other words, if you have any kind of equitable concern about the distribution of *health* across individuals in the community, then you need to pay due attention to the kind of *resource* distribution policy that is to be adopted and the circumstances (endowments, relative marginal products, etc.) which will determine the health distributional outcomes of each policy. The third general conclusion is that you need both to be more *precise* about the equity principles that are to govern the allocation of health care resources and to *choose* between equity principles - especially ones that can (in principle) have such widely differing outcomes.

A fourth conclusion related to access. Equality of access is frequently claimed to be an essential ingredient of equity in health care. The analysis here seems to deny this (see also Culyer *et al.* 1992). In so far as the ultimate concern (or one of such concerns) relates to the distribution of health, and the distribution of medical and health care resources is really no more than one possible means to this end, then it seems to me that there are two important conditions that access has to meet. One is that *initial access* (for assessment - diagnosis and prognosis) should be *cheap*, so that capacities to benefit can be assessed (for capacity to benefit is plainly an important determinant of the equitable outcome, even though it is not in my view the ethically dominant determining factor). The other is that access to active health care procedures, whether preventive, palliative or curative, should be allocated selectively in pursuit of both efficiency (cost-ineffective procedures are not warranted) and equity (interpersonal redistributions that do not promote equality are not desirable). The likely resource distribution that is warranted on both efficiency and equity grounds will be characterised by mechanisms to ensure *easy*

(such as cheap insurance and GP gatekeepers with nominal or zero user charges) and by *inequality* in receipt of care to achieve an equitable distribution of health outcomes.

Finally, although I have proffered one or two value judgments of my own in this paper, its analysis and conclusions do not depend upon them - if rejected, for example, the need for choice of distributional principle is in no way diminished nor, unless the conjectured ethical significance of the distribution of health is also rejected, can questions concerning the distribution of resources be considered independently of their consequences for the distribution of health.

REFERENCES

Culyer, A J (1989) The normative economics of health care finance and provision, *Oxford Review of Economic Policy*, 5, 34-58.

Culyer, A J (1993a) Conflicts between equity concepts and efficiency in health: a diagrammatic approach, in A M El-Agraa (ed.) *Public and International Economics*, New York, St Martin's Press, 42-58.

Culyer, A J (1993b) Health, health expenditures and equity, in E van Doorslaer and A Wagstaff (eds.) *Equity in the Finance of and Delivery of Health Care: an International Perspective*, Oxford, Oxford University Press, 299-319.

Culyer, A J, van Doorslaer, E and Wagstaff, A (1992) Access, utilisation and equity: a further comment, *Journal of Health Economics*, 11, 207-210.

Culyer, A J and Wagstaff, A (1993) Equity and equality in health and health care, *Journal of Health Economics*, 12, 431-457.