National Institute for Regional and Spatial Analysis

NIRSA

Working Paper Series
3-Jan02

Cross-Regional Equity in Health Care Funding*

By
Derek Bond and Denis Conniffe

New University of Ulster, Coleraine

Economics Department & NIRSA, NUI Maynooth

*A shorter and less technical paper based on this work and titled "Allocating Funding Across Health Boards – Is Equity Easy?" was presented by the authors to the Dublin Economic Workshop's 24th Annual Economic Policy Conference in Kenmare, October 12th-14th 2001.

National University of Ireland, Maynooth
Maynooth
Co. Kildare
Ireland

nirsa@may.ie
www.may.ie/nirsa

NIRSA is a research institute dedicated to the inter-disciplinary and comparative study of the impact of global processes on regional and spatial development. It involves several third-level institutions in Ireland: NUI Maynooth, Galway-Mayo Institute of Technology, Institute of Technology, Sligo, Mary Immaculate College - University of Limerick, Waterford Institute of Technology.

DISCLAIMER

While NIRSA is happy to facilitate the dissemination of working papers produced by associates of the Institute, the responsibility for content rests solely with the authors.

ABSTRACT

In Ireland, as in many other countries, much health care provision is State funded and State provided. Equity, in the sense of equality of 'treatment' for individuals with the same health needs irrespective of their geographical locations, or incomes, has been, and is, very much stressed as an appropriate policy objective. However, health care delivery in Ireland (and in some other countries including the UK) operates through regionally based Health Boards responsible for provision in their own regions. This implies mechanisms for funding allocations to Boards that allow for the corresponding regional needs, but the choice of mechanism is not at all a simple or non-contentious topic. It is important though, and so a very appropriate topic for research by the National Institute for Regional and Spatial Analysis. This Working Paper is at least a first step in a relevant investigation.

Following much research, formulae have been developed in the UK for the allocation of health care funding across regions with the objective of attaining regional equity in health care. The formulae employ criteria such as regional populations, age and gender distributions, and regional morbidity and socio-economic measures. The UK work is a natural starting point for corresponding Irish formulae. So one theme of this paper is a critical review of the UK's formula based system, looking at the theory and methodology underlying the derivation of formulae and assessing how the system has actually worked in Northern Ireland. However, we think it is also important to critically examine the apparent primacy granted to the objective of equity attainment. As regards the first theme, we conclude some assumptions are implausible and some issues unresolved and we discuss the implications should a formula based regional allocation strategy be desired for Ireland. As regards the second theme, we find that healthcare equity – in spite of its primacy in many minds – is an elusive entity, very hard to measure, let alone achieve. Making it a primary objective in health care provision carries a price as regards attainment of other objectives and this, along with its elusiveness, deserves consideration by health economists and policy makers.

I INTRODUCTION

We have two main objectives in preparing this paper. The first is to review the UK's formula based system for the equitable allocation of health care funding across regions paying particular attention to the Northern Ireland experience. 'Equitable' here means equal 'treatment' for individuals with the same health needs irrespective of their geographical locations (or incomes). Such a review is worthwhile for several reasons, including the tendency for administrative theories and devices introduced in the UK to often appear in the Republic after some time lag. This may be especially likely for health funds allocation since Ireland seems to accord equity attainment, or the aspiration to it, the same primacy that the UK does, or perhaps more accurately, used to do. But even if we do follow the UK to allocation formulae, we need to take a critical look at the assumptions underlying approaches and to assess how they have worked in practice. There may be lessons to be learned from the UK experience and we may need to do better.

The second objective is to look closely at what is realistically achievable in terms of attainable equity and what the implications are for the emphasis that should be placed on other objectives of health policy such as effectiveness, efficiency and operational flexibility. In Ireland, as in other countries, lists of such objectives have often appeared without much consideration of their joint attainability. For example, the Irish Department of Health and Children's (1997) *Statement of Strategy* listed "highest standards of effectiveness, efficiency, equity, quality and value for money in the health delivery system". But objectives, or at least the strategies to achieve them, can conflict[1] and, if one objective is accorded primacy, the attainment of others is affected.

As regards the plausibility of a formula based allocation, a recent paper on capitation funding in the public sector by Smith, Rice and Carr-Hill (2001) said that four possible mechanisms for distributing funds across geographical regions are:

(i) according to size of bids from regions
(ii) on the basis of political patronage
(iii) according to historical precedent
(iv) on the basis of an independent measure of needs

Put as simply as that and assuming that attainment of equity has primacy, it would be hard for anyone to dispute that mechanism (iv) is preferable. It might seem a natural step then to say that allocation of health resources across regions should be based on populations, age/gender distributions, and regional morbidity and socio-economic measures. That is indeed what the formulae purport to do. However, as we will see, there are problems, conceptual as well as practical, just beneath the surface.

Although the notion of equity was implicit in the principles of the UK's National Health Service from its origin in 1948, actual annual allocations for its first 30 years were largely based on those of the previous year with (depending on the state of national finances) some increments for growth. Since then however, there have been ambitious attempts to actually attain equity by using statistical or econometric formulae that relate financial allocations to measures of need. The methodology behind the currently employed formulae dates to the mid-1990's and applications in Northern Ireland are later still. In formulating and deriving these capitation formulae, the equity objective was not only taken as pre-eminent, but other

[1] Although not directly regionally related, anti-smoking campaigns can serve as an example. These have improved health but increased inequality of health, because over the years there has been a greater response from higher income groups (and males). Smoking has become ever more concentrated in the lower income groups. Indeed, in a broader context than health alone, welfare inequality has been further increased because of the regressive nature of increased tobacco taxation, one of the standard components of campaigns.

objectives did not even receive explicit consideration[2]. However, recent UK debates about health care policy and even some innovations in the NHS may signal some change of direction. We will return to the methodology and application of UK formulae in sections III and IV. But first, because of the centrality of the idea of equity attainment to the historical developments, we need to discuss the various interpretations of equity and its perceived primacy.

[2] Actual implementation of allocation formulae did allow for some other factors. For example, the funding changes were phased over years to retain some stability and operational flexibility.

II EQUITY

While there are several possible definitions of 'equity' in relation to health services, when it comes to measurement it is usually equality of expenditure (or a closely related measure) for equal need that is understood. That is, two individuals with the same health needs, but perhaps from different groups (defined by region, or income, or age, or whatever) are being treated equitably when there are equal expenditures on their health care. We will largely be focussing on this equal utilisation of health services for equal need (often called horizontal equity), but it needs to be said the term 'equity' is often understood as not only implying this, but also implying that higher income groups should contribute proportionately more to the funding of the health service irrespective of their utilisation of it (often called vertical equity).

It can be validly argued that unequal expenditures for equal need not imply inequity. Some people may *choose* to receive less treatment than is on offer and while they are perhaps unwise they cannot be subject to compulsion and are certainly not being treated inequitably. So 'access' to equal health care rather than equality of care may be a preferable definition and is, in fact, often used. The obvious problem, though, is that data records, particularly at aggregate level, often cannot distinguish between a choice not to utilise health services and a lack of need for them. Again, it is true that equality of the quantity and quality of health services delivered to people of equal need is what is important and equal expenditures do not guarantee these because efficiency could vary (by region, say). Although a certain amount can be done about this by replacing actual expenditure by an aggregate of standardised components (for example, cost hospital bed nights not by the actual regional costs, but by the average national cost and try to do the same for other components of care), there are great problems to quantifying efficiency and quality.

If achievable, it would be equality of health and not of health care than would be the desirable objective. But the quality of health an individual experiences through a lifetime depends on far more than on available health care. It is true that the rich enjoy better health than the poor (although ill-health can result from consequences of lifestyles associated with prosperity as well as poverty) and it may well be a valid welfare objective for the State to try to eliminate this inequity. But to do so would require extra resources devoted to many other areas that can affect health, such as nutrition, housing, education, etc. Health care expenditures on their own will not create equity in health and may indeed have relatively little impact, at least if past findings are still relevant. McKeown (1976) claimed the main determinants of health in the UK were behavioural, environmental and nutritional and the contribution of clinical medicine was relatively minor – a view that has been frequently re-expressed, for example by Benzeval, Judge and Whitehead (1995). Indeed, for the UK, the Acheson Report (Department of Health, 1998) concluded that health inequalities had not improved since the 1970's, but had even worsened in some respects. These considerations have some implications for both the choice of total health service expenditure and its regional distribution. Contrary to the views of some health policy analysts, it is clear than even on just an equity criterion, the health service budget should not become disproportionate to the budgets for other (health influencing) services. Correspondingly, regional equity of health would demand much more than regional equity of the health service, so that even if the latter were achievable, which we will show is far from straightforward, it could only do so much.

However, these considerations have not weakened the priority accorded in the UK to seeking equity in health care expenditure and the current UK capitation formulae are certainly directed to this end, but the view is not limited to policymakers, but widely held by health policy analysts, nor is it restricted to the UK alone. Gillon (1986) maintained that the view of equity in healthcare that commanded greatest support among health professionals and the public at large is that of distribution according to need, coupled with payment according to ability to pay. Mooney (1986) claimed that throughout OECD countries the public attached

greater importance to equity than to efficiency in health care and the claim has been frequently repeated. Wagstaff and Van Doorslaer et al (1992) believed there was broad agreement among policy makers in at least eight of nine Western European countries that health care ought to be financed according to ability to pay, but delivered according to need. A quotation from Wagstaff, Van Doorslaer and Paci, (1991) probably sums it up - "If there are two persons in equal need of medical care, it would be considered undesirable if the richer of the two were to receive treatment". Health services seem to be seen as somehow different to all other commodities[3]. Majority views in the US do not give the same primacy to equity. Health services are more likely to be thought and written about in the general context of consumer demand for commodities with a consequently greater trust in the role of markets to promote efficiency and less emphasis on equity. Enthoven (1988) could perhaps typify the majority viewpoint. However, there are American health economists, for example, Rice (1998), who disagree and take a 'European' line.

Views in Ireland fit into the European mould. Nolan (1993) saw the Irish perception of healthcare as "distinctive in terms of public attitudes as regards equity", with access to it "generally regarded as a basic right" and that equity concerns related not just to publicly funded health services, but to "the overall use of health services, whether publicly or privately financed or delivered". Indeed, the importance of healthcare equity relative to equity in consumption of other (health influencing) commodities seems to be more highly rated in Ireland than elsewhere. This was not a unanimous view in the past. Tussing (1982), who studied health and poverty in Ireland, said little about equity of health service distribution. He stressed housing, the living and working environment, and health problems arising from hazards peculiar to poverty. He also favoured health care cost control including consideration of pricing and payment systems. But a succession of Government papers - *Health the Wider Dimensions* (1986), *Commission on Health Funding* (1989) and *Shaping a Healtheir Future* (1994) have emphasised equity in healthcare expenditure. Indeed, in contributions to a recent book (Leahy and Wiley, 1998) officials of the Department of Health and Children (O'Dwyer, 1998, and Barrington, 1998) not only see their Department's objective as achieving healthcare equity "across the whole country and between groups with different needs", but of reducing inequality generally - "The biggest challenge for the health services . . . is . . . the reduction of inequality…", the logic apparently being that health care would lead to health which would enable everyone to attain economic and social success[4]. In line with this, Barrington (1998) claimed the Department had made a major contribution to the Irish economic boom of the 1990's. With these views, it is not surprising that these authors had no worries about choice between objectives (equity v efficiency, say) in health care provision, or total expenditure on health care v that on other commodities. The Irish health service was "perhaps one of the most cost-effective in any OECD country" and the very large increases in Irish health expenditure in the 1990's were treated as if exogenous.

To summarise: in Ireland, the UK and Europe, equity is accorded primacy among objectives by policymakers, health administrators and many analysts. But the measure of equity is via expenditure on health care and the connection between that and health outcomes is quite tenuous. In the following sections we will show that in the UK, which has taken the formulaic determination of equitable expenditure allocation further than other developed countries, the resultant distribution of even this expenditure measure is of very dubious equity.

---

[3] Some health policy analysts argue that the 'exceptional' characteristics of health services, make it extremely difficult to say any level of health expenditure is too high. Others (for example: Smith, Frankel and Ebrahim, 2000) deny such conventional economic views as that 'free' supply of a desirable commodity will increase consumption without limit unless strictly rationed.

[4] In fairness, this is an extreme position. More reasonably, the recent publication by Layte and Nolan (2001), while it restates the equity in healthcare aspiration, recognises the limitations of healthcare alone to achieve equity in health outcomes, let alone in other spheres.

III THE UK[5] ALLOCATION FORMULAE

The financial resources required by a regional health authority will depend on the population of the region. So a first stab at equitable allocation would just allocate a fixed total in line with population proportions over regions. But people of different age and gender have different health care needs, so this allocation should be modified to reflect different demographic and gender patterns that may exist between regions. Again, even given age and gender, people can differ in health care needs due to various morbidity and socio-economic factors. For example, with all other things equal, unemployment is probably associated with more ill health. So there should be further adjustment for these factors. All this is very reasonable. The difficulty lies in determining the magnitudes of the adjustments that should be made.

The essential problem can be understood by considering a hypothetical (grossly over-simplified) situation. Suppose we have only two age categories – old and young – and nothing else affects health care needs. Suppose we have only two regions, both with 1 million population – one with 75% old and 25% young and the other with 25% old and 75% young. Suppose that in a previous time period central authority had allocated £1 billion to each region, because they had the same populations, but no allowance was made for age distribution. Now central authority wishes to again allocate an unchanged total of £2 billion, but wants to adjust equitably for age distribution. The regional health authorities are able to give us the actual health care expenditures incurred by old and young in each region in the most recent time period prior to this allocation exercise. They were:

	Old	Young	Everyone
Region 1 Expenditure Population	£800 million .75 million	£200 million .25 million	£1 billion 1 million
Region 2 Expenditure Population	£500 million .25 million	£500 million .75 million	£1 billion 1 million
Whole State Expend. Population	£1300 million 1 million	£700 million 1 million	£2 billion 2 million

The old incurred more health expenditure than the young in a per person ratio of 1.3 to .7 = 1.86. So Region 1 got 'too little' and region 2 got 'too much'. Assigning according to the national ratio gives:

	Everyone
Region 1 Expenditure Population	£1.15 billion 1 million
Region 2 Expenditure Population	£.85 billion 1 million
Whole State Expend. Population	£2 billion 2 million

[5] Actually, Scotland has its own allocation formulae, based on a rather different methodology to the rest of the UK.

Real life situations are much more complicated with more regions, many age categories, plus gender and other factors, as well as various classifications of health care expenditure (acute hospital care, GP services, elderly care, etc.) being allocated separately, rather than as a single health care allocation. However, the essence of the procedure is the same.

The problem with the procedure is that the national, or overall, expenditures on old and young are being taken as providing 'true' measures of age related need. This clearly assumes that even though regions differed originally in relative expenditures on old and young, they did so in a 'random' fashion so that the average, or overall, is a fair estimate. To a considerable degree we are assuming the original allocation was 'fair' to begin with. There are other possibilities. Suppose we believe that some of the old can be relatively easily fobbed off from seeking some of their health entitlements and that this is more likely to happen in a more financially pressured region. Then we would feel use of region 1 in estimating relative need would be biased against the old and region 2 would give a fairer picture. The expenditure on each old person was £2000 and on each young person £667 (as compared with the overall of £1300 to £799) or 3 to 1. Allocating the £2000 on this basis would lead to the allocation:

	Everyone
Region 1 Expenditure Population	£1.25 billion 1 million
Region 2 Expenditure Population	£.75 billion 1 million
Whole State Expend. Population	£2 billion 2 million

This gives a still larger proportion of total funding to region 1. This example, is *not* merely hypothetical, although a vast simplification of its real life counterpart. In Scotland, there has been opposition to using national average expenditures as estimators of need. The case, as argued by Sutton and Lock (2000), is that there is a large literature suggesting that high need groups under-use health care (they cite, for example, Tudor-Hart, 1971) and so regions with substantial proportions of high need groups will, if quite constrained in resources, 'underspend' in relation to these needs. So Scottish estimates of need are based on averages from the regions judged to be less financially pressured. Some subjective judgement may be involved here, but is probably surmountable, for example, by the idea of assessing pressure by hospital waiting lists suggested by Blundell and Windmeijer (2000).

Of course, someone could argue that regional health care professionals and administrators might react to tight financial constraints by concentrating care on the most needy, the old in our example, and (relatively) neglecting the young. Then national figures would over-estimate the needs of the old relative to the young. We are not trying to disagree with Sutton and Lock, but to emphasise that there is a basic circularity in trying to measure the needs of a category of people by the observed expenditure on them. The circularity can be broken if strong assumptions are made. Taking national averages as proportional to need and readjusting the regional allocations to them assumes the original regional differences were unrelated to needs ('random'), so that the original allocation was unbiased, if imprecise. The 'Scottish' approach is different and 'better' if original regional differences were related to needs in the manner they describe, that is if their assumptions are true.

With the real life complications added in, tabular outlines of the needs estimation process become exceedingly cumbersome and a regression equation framework is usually adopted. For a particular category of health care (say acute hospital care) we could write a model[6]

$$y = a + b_1 x_1 + b_2 x_2 + \quad . \quad . \quad . \quad + b_n x_n + e, \qquad (1)$$

where y is expenditure, the x's are needs (dummy, perhaps) variables, the b's are the adjustment coefficients to be determined and e is the 'random' deviation form 'correct' expenditure. But although regression may have computational and presentational virtues in handling complexity, the circularity is *not* evaded by the more sophisticated representation[7]. However we will subsequently be considering the possibility that adding extra ('supply') variables to equation (1) can break the circularity. First it is necessary to explain further complications associated with the estimation of needs coefficients. A model can involve a lot of needs variables and, as there are a limited number of Health Board regions, extra observations are required to permit estimation. This is achieved by relating expenditure to needs variables at electoral ward level thus generating a very large number of observations. The statistical task is gigantic. Taking acute hospital care in England as an example, hospital records corresponding to a chosen base period are comprehensively examined to determine the electoral wards patients originated from, the treatments they received, the bed nights involved, etc. So a dependent variable can be constructed for each electoral ward, as will be explained shortly. From small area census data and other sources, demographic, morbidity and socio-economic variables can be ascertained to give the explanatory, or needs, variable values for each ward. Clearly, in performing such an exercise, certain standards of relevance, reliability and timeliness of data, need to be achieved. The exercise is only feasible for certain 'base' years, so that allocation formulae remain unchanged (other than annual changes to the regional need variable measures) until the next base year.

Actually, the estimation of coefficients in the current UK approach is done in two stages. First, age and gender adjustments coefficients are derived from national totals by the already described procedure that assumes national expenditure averages are proportional to needs. These are used for an initial adjustment to allocations based on 'raw' populations. The second stage should allow for the morbidity and socio-economic variables, but since age and gender effects are already presumed provided for, the dependent variable should be purged of these effects. So national average treatment and bed night costs are ascertained for each age band and gender combination. These 'standard' costs are then applied appropriately to the hospital patients to produce the ward level dependent variable values. So the dependent variable, usually called 'utilisation', is an expenditure type measure, but presumed free of age and gender effects. Skipping, for the present, discussion of some important issues of methodology, the morbidity and socio-economic coefficients are then derived by regression of electoral ward 'utilisation' on the corresponding variable values.

Although these need coefficients are estimated at ward level, they are of course employed for allocation at Health Board regional level by multiplying coefficients by regional needs measures. This introduces complications in itself. A formula estimated at one level of aggregation may not apply at a different one. The 'ecological fallacy' occurs when a quite different relationship may seem to exist at the more aggregated level, but even the more familiar (though possibly related) difficulty known in economics as the 'aggregation' problem

[6] We use a linear model purely for ease of exposition. Most work did not use linear models and this will be returned to.

[7] The familiar econometric condition for validity of regression is that the x's are independent of e. In this case that means regional deviations must not depend on needs.

can cause difficulties unless choice of functional form of relationships is very restricted [8]. We will see later that a particularly important issue is complicated by the ward level estimation.

Probably the most contentious issue in the current methodology relate to the use made of 'supply' variables and the easiest way to appreciate the current situation is to look at how the NHS allocation formulae evolved. As mentioned in the Introduction, allocation formulae were not employed until the mid 1970's and initially they were fairly simple adjusting for age, gender and standardised mortality rate under 75 years of age (SMR75) as a proxy for all other influences on health. But since adjustments used national averages the circularity difficulty was inherent in the exercises. Coopers and Lybrand (1988), commissioned by the NHS, conducted a ward level study and looked at various demographic, social deprivation and morbidity variables. But of more fundamental importance, they employed supply variables to try to break the circularity problem. For example, for each electoral ward they devised measures of GP availability in the ward and ease of access (via a distance gravity model) to hospital beds from the ward. Then, rather than estimating (1), they estimated

$$y = a + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n + c_1 s_1 + c_2 s_2 + \ldots + c_k s_k + e\,, \qquad (2)$$

where the s's are the supply variables. The needs coefficients (the b's) are now being estimated controlling for the supply variables ,that is as if the supply variables were everywhere equal. They then interpreted the b's as having been adjusted for previous resource imbalances between regions and appropriate for use in regional allocation formulae[9]. Clearly, a lot is assumed here about how well the supply variables perform as measures of resources, but there are deeper issues.

The recommendations of the Coopers and Lybrand report were not fully implemented, although some changes were made to formulae. Criticisms of their report and continued dissatisfaction with the actual operation of the allocation formulae led to the NHS commissioning another study, this time by staff of the University of York. The model described in the report by Carr-Hill et al (1994) became the basis for the current NHS allocation formulae. There is much to be said about the methodology of that report, but keeping to the supply and circularity issue for the present, the key point is that the needs coefficients are estimated from a regression of the dependent variable (utilisation) on need variables *alone*, that is, an equation of the form (1). Their argument, at least as we understand it, is that if a region got greater than average resources in the past (represented by supply variables), this was probably for the good reason that its needs were greater, so they did not want to control for supply in estimating current needs coefficients. So they are returning to the original assumption that the national, or overall, distribution of resources to needs is OK and all that is required is some 'fine-tuning' in reducing the random variability between regions. Remember that the first stage age/gender adjustment implicitly involved the same assumption. Presumably, it could be valid, although, as already mentioned, Sutton and Lock (2000) would not accept it for Scotland. Indeed, another Scottish health service analyst,

---

[8] If a linear relationship $y_i = a + bx_i + e_i$ holds for ward i and we know the same relationship holds for all other wards in a region, adding up for the region will give $\Sigma y = a + b\Sigma x + \Sigma e$, that is, the same relationship holds at regional level between the regional totals. But if the ward relationship is $y_i = a + bx_i + cx_i^2 + e_i$, the regional relationship is not $\Sigma y = a + b\Sigma x + c(\Sigma x)^2 + \Sigma e$, but $\Sigma y = a + b\Sigma x + c\Sigma x^2 + \Sigma e$ and these could be quite different if the x's vary much between regions. The practical point is that ward level data and hence Σx^2 will not be known until the formulae are due for re-estimation. At best, the aggregate relationship is only approximate.

[9] Their dependent variable was not actually a utilisation/expenditure measure, but one based on deaths and hospital discharge rates that they felt was an adequate proxy.

Leyland (2001) who (with many others) discussed the York methodology at a Royal Statistical Society meeting on the topic, also disagreed with the procedure and underlying assumption. Bevan (2001), Chisholm (2001) and Sanderson (2001), also had strong reservations about this, with the latter feeling that if various policies aimed at equity or efficiency have any point, existing patterns of utilisation must be regarded as inappropriate. Subjective judgements on how well the NHS allocation system has been working could matter a lot to what assumptions are found plausible.

But whatever about the appropriate assumptions about past *regional* resource allocation, the York procedure must be wrong for a methodological reason. The regression analyses were at ward level not regional level. Within a Health Board region, wards will differ as regards supply variables (as is obvious from their very nature - distance to hospital beds, access to GPs etc.) depending on their location, rural or urban nature and the decisions of the Board as regards siting of facilities. These differences would occur even if all Boards had received exactly the same resources. Unless the needs coefficients are estimated with supply variables in the regression, they will be 'biased' because they contain supply effects, whose magnitude will depend on the patterns of need and supply at ward level. But these supply patterns are hardly likely to relate to historical cross-regional allocations, so it could not be argued that it is appropriate to retain these biases. This point was made by one of us (Conniffe, 2001) at the RSS discussion and elicited the reply from Smith et al (2001) that "this issue is not yet completely resolved" and is "an interesting area for future research".

Supply variables do play a role in the York approach, but only in helping choose which set of needs variables are included in the final regression. There are a very large number of candidate needs variables (type of housing, tenure of same, amenities in same, density of occupation, employment status, social class, car ownership, ethnic origin, welfare recipience, etc) and so selection is required. Simplifying, just for the moment, to make the procedure intelligible, the York team put all needs variables (actually almost all, some are reserved for a reason to be described later) and supply variables in a huge initial regression and proceed to eliminate non statistically significant needs variables (but not supply variables) using the usual, rather mechanical, approach of stepwise regression. Economists usually dislike this approach as being too arbitrary in the final choice of remaining variables, although it has to be said that other social scientists often employ it. However, the York team appeal to their key assumption - that historical needs led to the resources now being measured (or proxied) by the supply variables - to justify the procedure and to make very strong deductions about the needs variables that are excluded or included. They argue that if a needs variable is found non-significant and is excluded, this is because it was taken account of in past allocation of resources and so its effect operates through the supply variables. So it should not play a role in current allocation. On the other hand, a needs variable that is significant, in spite of the inclusion of supply variables, represents something not allowed for in previous resource allocations and so should play a part in the current one. They label included and excluded variables with the emotive titles of "legitimate needs drivers" and "illegitimate needs drivers". Having finally identified a small group of "legitimate needs drivers" they drop out the supply variables and derive the formulae coefficients as already described.

The first point to be made is that, once again, the distinction between regional level and ward level distributions of resources is being missed. Even if past regional resource allocations had been made fully equitably given perceived needs, there is no reason why the postulated process should operate at ward level within regions. For purposes of argument we could imagine a sufficiently large number of regions to estimate a regression equation at that level. Then the argument might have some validity, but only if their assumption about past allocations having been accurate responses to past needs is true. We osay 'some validity' , because even then there are other reasons why needs variables might or might not be statistically significant. A group of needs could be individually important, but statistically highly correlated (it is well known that this is true of deprivation measures) so that when

jointly inserted in an equation they do not register as significant. This is not because the supply variables are picking up their effects, but because they act as proxies for each other. We believe the York approach is actually a highly arbitrary one as regards the final choice of formula variables.

It is worth noting that age and gender related needs are not subject to this selection procedure or to any possibility of classification as illegitimate needs drivers. The two-stage allocation process ensures that age and gender are always adjusted for, in the manner already described. Since age and gender are such obvious variables, it is understandable that anyone would be uneasy about omitting them from an allocation formula. On the other hand their invariable inclusion seems inconsistent with the key notion of the York approach – that past allocations to regions will have (equitably) taken account of perceived past needs. Of course, age distributions will change somewhat over time, but that could also be said of any other variable.

As mentioned, we have simplified our account of the York variable selection procedure, to bring out the essential features and assumptions. features and assumptions. Actually their method is technically much more complex, because they argue that supply is *endogenous* and that consequently multi-equation systems estimation must be used, which we consider quite unnecessary and serving only to obfuscate what is going on. They say (reasonably) that expenditure (utilisation) is affected by supply factors as well as needs and that supply variables are also related to needs. Supply may (they would argue) have been an appropriate response to past needs, but current (meaning the estimation period) needs are no doubt related to past needs also. We do not object to describing supply as *endogenous,* but the framework is quite compatible with standard regression or ordinary least squares (OLS). But they do not realise this and employ two stage least squares (2SLS), which would be appropriate when there is not only a causal relationship from explanatory variables to the dependent variable, but also causation in the reverse direction - from dependent back to the supply variables.[10]

In 2SLS the hypothesised backwards causation from the dependent variable is got round by employing what are termed instrumental variables that can be used to concoct replacements for explanatory variables contaminated by the backwards causation. They must be known to be related to these variables, and also known *not* to be otherwise related to the dependent variable. It can be very hard to find such instrumental variables (IV's) and with invalid IV's, 2SLS can be seriously misleading. There is a large econometric literature emphasising these points. Here the IV's must be related to the supply variables, and *not* otherwise related to utilisation. The York approach appeals to the concept of legitimate and illegitimate needs variables. The definition of the latter – not significant in equations because taken into account in supply – is seen as implying they are suitable as IV's. But these are actually identified during the statistical analysis. So the York analysis commence with a 2SLS analysis including all supply variables and almost all needs variables except for some that they know (although how is not explained) are illegitimate needs variables. These are necessary because 2SLS simply will not work without some IV's. Then the non-significant needs variables in

[10] The implied model is the multi- equation system

$$y = a + b_1 x_1 + b_2 x_2 + \ldots + c_1 s_1 + c_2 s_2 + \ldots + e$$
$$s_i = c_i + d_{i1} x_1 + d_{i2} x_2 + \ldots + u_i ,$$

which is validly estimable by OLS because because the system is recursive. While reverse causation of utilisation on supply (e affecting the u_I) would invalidate this, it seems impossible that there can be any in the point in time data the York team employ. Proximity to a hospital (a supply variable) could increase utilisation, but an electoral ward with high utilisation cannot move towards a hospital. Of course, *over time* new hospitals may get built nearer a ward, but the data refer to a fixed chosen time period. Indeed, if data were available over time, their own assumptions about supply determination would imply that insertion of past values of needs variables in all equations, followed by OLS estimation, would be the appropriate analysis.

the estimated equation are deleted and the equation re-estimated, but with the set of IV's augmented by the deleted variables. The procedure continues, exactly analogously to stepwise regression, except that at each stage 2SLS and not OLS is employed. All our previous remarks made in an OLS context apply to the 2SLS analyses also.

Sheldon and Carr-Hill (1992), two of the authors of Carr-Hill et al (1994) had strongly criticised Coopers and Lybrand (1988) on various grounds, including some which seem remarkably applicable to the York work and to which we will return. But it is interesting that they were not then disposed to interpret supply differences between regions as following from equitable distribution according to need and indeed their discussion of earlier work by Derbyshire (1987) and others includes mention of regions "unfairly allocated too few resources" and "circularity" in allocation. Of course, opinions can change. Something they did stress, with which we would fully agree, is that there is a mystique associated with complex statistical techniques. They remarked this could be and frequently was used as "a statistical fig-leaf" and that there ought to be a responsibility on analysts to make assumptions and "political judgements" explicit. But Carr-Hill et al (1994) is no improvement on Coopers and Lybrand (1988) in that respect.

There are many other issues that could be raised with regard to the York methodology some of which we will examine in the next section when looking at what the approach produced in Northern Ireland for the acute hospital care sector. However, there is one other general issue arising in modelling than we want to address here. The functional form chosen for a model can matter greatly and ought to be theoretically plausible and fit the data well. Smith et al (2001) choose a multiplicative (double-log) model, as did Carr-Hill et al (1994) and HHCRU (1997). This assumes the dependent variable zero if any needs variable is zero, which is plausible for the output of a firm, where nothing can be produced without some labour, some capital, some raw materials etc. We doubt if this is plausible for the consumption of health services. In our opinion, a functional form ought to take into account both behavioural plausibility (perhaps partly deduced from the views of health service professionals about how health services function[11]) and data analysis to test and refine. The York team and their subsequent collaborators in Northern Ireland do not seem to have engaged in any prior assessment of behavioural plausibility. Smith et al (2001) replied to criticism that the multiplicative model fitted well and was not rejected by specification tests, but the latter seem to us to be far from stringent[12]. The most worrying point about incorrect functional forms relate to their use at the final allocation stage, but they can produce strange effects at the earlier data analysis stage. For example, in regression analysis one of the commonest reasons for a variable to seem to have a counter-intuitive role is because the functional form has been wrongly chosen.

[11] Discussions with health care and other professionals ought to elicit useful information. For example, what needs variables might interact suggesting refinements of a model? Or if an increment in a needs variable is expected to increase utilisation, would another increment of the same magnitude be thought to have an equal, lesser or greater, effect on utilisation? The corresponding functional forms would be semi-log, linear, or exponential, respectively.

[12] In Carr-Hill et al (1994) and HHCRU (1997) these seem to amount to assessments of R^2 which are often inappropriate for functional form assessment and endogeneity tests, which we think irrelevant.
The RESET test, which featured in later Northern Ireland reports and is referenced by Carr-Hill et al (2001), does have a role, but much more could be done. For example, semi-log and double-log forms can be compared by the Box-Cox methodology and semi-log and linear through the Box-Tidwell approach, although there are other techniques too.

IV THE NORTHERN IRELAND EXPERIENCE

The formulae derived for England were not applied to Northern Ireland because it was felt there were sufficient differences in circumstances, in the range of potential explanatory variables and in availability and quality of data to warrant separate estimation. Taking the case of acute hospital services, the York methodology (with minor modifications) was applied to the nearly 500 wards[13] in Northern Ireland, with the econometric modelling (HHCRU, 1997) performed by staff of the Health and Health Care Research Unit (HHCRU), QUB in conjunction with the Centre for Health Economics, University of York. As in the exercise for England, the objective, having gone through the stepwise selection process (from nearly 50 candidate needs variables and five supply variables) was to have a final model with good fit (in the sense of adjusted R^2, statistical significance of coefficients and passing an endogeneity test[14]) which was also parsimonious (which they took to mean 4 or 5) in needs variables. In the chosen HHCRU model these variables were: SMRALL – standardised mortality rate at all ages, ISTOTAL - the proportion of people on income support, FAMCRED – a similar measure of receipt of family credit, ELDER61 – the proportion of people (over 75) living alone and PCLBW – a measure of low birth weight frequency.

The corresponding coefficients were used in a second stage modification of the allocations to the four Northern Ireland Health Boards. (A first stage had already adjusted for age/gender patterns). Health Boards, or at least two of them, did not consider the resulting allocations in the least equitable and this is actually how we were introduced to these allocation formulae. We were each approached by (different) Health Boards seeking help in understanding the econometric model that was producing the allocations and making a case against the reduction in its former share. We found that Board officers not only did not understand the econometric methodology, but disagreed with the underlying assumptions, once they understood what these were. Implementation of formulae was contested for a considerable time and eventually accepted only after revision. In a zero-sum game situation (the budget under this heading was fixed for Northern Ireland as a whole, at least at that time), Boards can be expected to fight their own corners, but we learned a lot about the practical performance of the allocation procedure.

In the previous section we have outlined some major problems with the York methodology. But suppose we choose to accept it all – 2SLS, the concept of legitimate and illegitimate needs drivers, the multiplicative functional form and final re-estimation without supply variables. One of us received (via the DHSS and a Health Board) the original data used by HHCRU, so it was possible to repeat their exercise a considerable number of times just making some quite arbitrary and no less plausible choices about variable deletions at various stages[15]. Many of the resulting models were superior in terms of statistical fit to the HHCRU model, although they usually involved more variables. Although we do not agree with the primacy of parsimony, we will limit ourselves here to those models that had as few variables. We will describe four of these models – three of which fit (in terms of HHCRU's own criteria) as well or better than their model and have the same number of variables, while the other fits only marginally less well and has fewer variables. For practical purposes, it would not matter that several equally tenable models exist provided they lead to the same inter-regional allocations, so we show the allocations. For confidentiality, the Boards are just

[13] Wards with particularly small populations were actually combined with nearest neighbours.

[14] These criteria were assessed at the penultimate stage of analysis – before dropping supply variables.

[15] There are many relatively arbitrary variations possible with stepwise variable selection procedures. These include various strategies of variable omission, reassessment and possible readmission; choice of tests used and probability levels chosen for exit and entry. There is *no* accepted optimum procedure and every textbook stresses the arbitrary element. With 2SLS further variations follow from the initial choice of IV's.

labelled A, B, C, D and the allocations are given in percentage terms. The HHCRU model is given first for ease of comparison.

MODEL	**R^2 (adj.)**	***Allocation***			
		A	B	C	D
SMRALL ISTOTAL FAMCRED ELDER61 PCLBW	.52	41.51	24.17	18.00	16.31
SMRALL ISLT65 ISGT65 ELDER61 FAMCRED	.53	41.11	24.15	18.22	16.52
SMRALL ISLT65 ISGT65 PCLBW FAMCRED	.52	41.17	24.10	18.12	16.61
SMRALL ISLT65 ISGT65 AMENIT21 FAMCRED	.53	41.54	23.85	18.08	16.52
SMRALL ISLT65 ISGT65 D FAMCRED	.51	40.96	24.10	18.19	16.75

The variable AMENIT21 measures lack of some basic household amenities, while ISLT65 and ISGT65 are more refined measures of categories of income support.

Now the allocations given by the 4 new models may not look very different from the HHCRU model, which is unsurprising given that we are making only relatively minor variations on the York/HHCRU approach. A major change in line with the criticisms of the previous section would no doubt produce much greater deviations from the HHCRU allocation. But even so, every .1% is a very large sum and for at least one Board the differences in allocations shown would, they claimed, have made the difference between comfortable operation and very harsh service reductions. It is possible that expecting regression equations with substantial residual variation to allocate with great precision is expecting too much. This theme has appeared in the literature already. The R^2 values just shown for the acute care programme in Northern Ireland, are by no means the lowest found for any UK models. In the discussion on Smith et al (2001), Derbyshire (2001) remarked on an equation with an R^2 of .13, while Goldstein (2001) also expressed concern about the magnitude of unexplained variation. Wright (2001), though commenting on education (where funding is also formula allocated), worried about consequent fluctuations in expenditures, saying "although fairness is said to be the most important thing, stability is what makes life bearable". Barrow (2001) made a related point about the importance of stability. Chisholm (2001), for the same reason of large unexplained variation, as well as basic doubts about methodology, felt that the pursuit of equity, while laudable in principle, seriously conflicted with other objectives. In replying, Smith et al maintained their R^2 values were quite good by the standards of equations estimated from large cross-sectional data sets in the social sciences.[16] This is true, but an equation that is good enough to confirm a social scientist's hypothesis about the effect of some factor, in spite of great variability in the data, is not necessarily good enough to allocate budgets between competing regions.

We do believe that the York/HHCRU methodology can be improved upon, probably most fundamentally by addressing some of the issues described in the previous section, but also through relaxing the degree of parsimony desired by York/HHCRU. The number of needs variables obviously needs reduction from the initially large number, but with nearly 500

[16] Sheldon and Carr-Hill (1992) had been critical of the (not too dissimilar) R^2 values achieved by Coopers and Lybrand and, indeed, of their use of R^2 as a criterion at all, again showing how opinions can change.

observations available, there is no justification for preferring a 5 needs variable model to, say, a better fitting model with a few more needs variables. We have verified that extra variables added to various 5 variable models are highly statistically significant and improve fit and performance in relation to specification tests. The larger models may cure other problems. For example, in the Northern Ireland analyses (unlike English models) dummy variables for regions[17] were included to control for some supply related regional factors (the nature of which was unexplained). However, unlike the supply variables, these were retained in the final OLS regression. Dummy variables like these could easily pick up the effect of an omitted needs variable that differed between regions. The HHCRU (1997) report said they "feel secure the coefficients obtained are not a reflection of an 'omitted' need variable - precisely because so many variables have been included in the modelling process". But the coefficients of these regional dummy variables do not depend on the number of variables included in the general modelling process, only on the number still in after the stepwise procedure, which was just five, because of the parsimony argument. We found the coefficients did reflect omitted variables. Different combinations of need variables, as in the models described earlier, led to very different coefficients of dummy variables. However, as the number of needs variables was increased from 5 to 8 the magnitudes of coefficients decreased considerably.

But although we think that various issues need resolution and that modelling could be better, it need not follow that the eventual equations and formulae based on available data would be good enough. It is possible that some relevant variables are not being identified, let alone measured, and that the quality of measurement on some others is inadequate. Reliable formulae will not result from poor data. Creating the HHCRU database was a huge task and we do not want to belittle the effort, but there are some worrying features that probably deserve comment, certainly in case the approach is to be attempted for the Republic. That supply affects demand (utilisation) is a key idea, but the supply variables may not be good measures of true supply. They ought to be highly related to the dependent variable (in an equation without needs variables) and if the supply variables are proxies for resources previously allocated in response to past needs they ought to be well related to current needs. Neither York (1994) nor HHCRU (1997) provided information on these points, but in conducting our own analyses on the HHCRU data, we did not find any relationships very impressive. For example, regression of GPALL (the supply variable measuring availability of GP services to a ward) on *all* (almost 50) needs variables had an R^2 of .22, nor was this the lowest. There are several reasons why a relationship may be poor, including the possibility that the assumption that resources were allocated according to need is wrong, but clearly the definition and measurement of supply variables could be a critical matter.

We also noticed frequent occurrence of 'odd' relationships. For example, the explanatory variable CAR31 (proportion of households without a car), which would be expected to indicate less well off areas, had good explanatory power in equations, but it had the wrong sign, indicating increasing, not decreasing, utilisation. The same occurred with other variables including the unemployment measure. With highly correlated data, coefficients can sometimes seem counter-intuitive, but even so, the findings are suspicious. One possible explanation would be incorrect functional form, which would relate back to a theme of the previous section. But another alarming possibility has to do with data. Age, which is never in the equations, could explain the odd results. Young people are less likely to have a car, less likely to be ill and more likely to be unemployed. But, of course, age should not matter since

[17] With English data, Carr-Hill et al (1994) treated such Board level effects as a random variable, with, obviously, a common value for all wards within a Board region. Then a mixed ("multi-level") was estimated. With only four Boards in Northern Ireland Board effects were treated as fixed and represented by dummy variables. Either way, the nature of the Board effects are left unspecified. When Coopers and Lybrand (1988) did this for English data, Sheldon and Carr-Hill (1992) described it as "absurd . . . the very core of the analysis should be the attempt to explain this variation".

the dependent variable (utilisation) is supposed to have deleted age and gender effects through use of national standardised costs. But if that process was defective age effects could persist. Then either age effects have to come in again at the second stage allocation, or more logically, the two stages should be combined with age coefficients estimated simultaneously with other needs coefficients. The key point here is that the validity of the entire analysis depends on the logical plausibility of the standardisation procedure and the quality of the data used in it.

Turning now to the implications of the HHCRU model for equity at ward rather than Board level, there were substantial differences between wards in utilisations. That, in itself, need not mean there was substantial inequity because supply and need may have varied correspondingly between wards. But if so, the equations should not be leaving 50% of variation in the dependent variable unexplained. If we take the HHCRU modelling results (and earlier York work in England) at anything like face value, there is a lot of inequity existing within Board regions. Of course, for all the reasons already discussed, there may be many faults in the equations. But it is also possible there really is substantial inequity between wards. We can certainly say that these analyses have not revealed equity between wards – if by that we mean equal utilisation for equal measured need.

Before concluding this section it is worth considering the extent to which consideration of private medical services provision complicates the allocation and estimation issues and to what extent it has been taken into account in the modelling. In a context of the State meeting full expenditures, a regional Board could either provide medical services or purchase them from a private provider (or from another region). There have been moves in the UK in recent years to separate the purchase and provision of health services at regional level. This should not alter the principles of allocation of funding to regions since needs would be unaltered, although if there were efficiency differences between types of care providers there could be short term pressures. The estimation could be somewhat more complicated, however, not only because of the need to get proper data representation from public and private providers, but also because there could well be more supply variables if, say, access to a public hospitals is considered different in kind from access to public hospitals. Private payment for services introduces still more complications. Equity in regional allocation will presumably require adjustments for the proportions of regional populations opting for private payment, but re-examination of what is meant by equity would be implied also. Would regional equity in allocation now mean equal 'utilisation' for equal need for those availing of the public service or should it mean equality of utilisation between public and private?[18] And what would the implications be for the modelling procedure?

There may be deeper issues too. When a State chooses to fully fund and publicly provide all health care (which more or less used to be the situation in the UK with the NHS), deliberately excluding private sector involvement, it needs administrative allocation mechanisms to function at all. It is behaving like a 'command economy' as regards the health care sector and there are no automatic or market forces influencing allocation. Research into formulae developed in that context. How might matters have evolved in an environment with a much greater private sector role? Naturally, these topics have received little attention in the York or HHCRU modelling studies, with the latter just remarking that that private bed availability and utilisation in Northern Ireland was very small. However, some reservations about the implications for modelling have appeared, for example, Barrow (2001), Chisholm (2001) and Glennerster (2001). If allocation models are to be developed for the Republic, the topics will require far more consideration.

[18] The Irish authors quoted in section II would seem to take the latter interpretation.

V SUMMARY AND CONCLUDING REMARKS

We think that the preceding sections will have made clear that formula based regional allocation of health care finances as currently practised in the UK, is not the panacea it has sometimes been claimed to be. Although the objective is focussed on equity, this is interpreted as equity in expenditure, which, as we discussed in section II, is a far from ideal measure. Even then, as we showed in section III, the theory and methodology behind the derivation of formulae are replete with strong assumptions that lack widespread support and in some respects are, we think, quite wrong. In section IV we turned to the practical implementation of estimation and allocation as exemplified by application in Northern Ireland, where it was certainly not seen, by at least some Health Boards, as an improvement on previous approaches. We pointed out various worrying features such as inadequate stability of allocation over equivalent models and the substantial unexplained variation. Overall, there is little support for the argument that formulae lead to equitable allocation between regions, which is one, but only one, necessary condition for equitable allocation between individuals.

At this point it is only fair to say that there are those who believe the allocation formulae do deliver equity, or are at least better in that regard than the sort of alternatives described in section I. The Government commissioning agencies and those distributing funding to Health Boards (the DHSS in Northern Ireland) believe in them. But they would have little hope of having them implemented if they did not display confidence in them. The, often eminent, researchers who developed theory and methodology believe in them also, although perhaps sometimes less dogmatically that do the agencies that commissioned them. Although we earlier cited many critical comments made at the RSS discussion of Smith et al (2001), not all contributors were critical. Also, as described in section II, equity was an objective of the NHS long before the current formula methodology and indeed a key objective of European health policies whatever the allocation mechanisms employed in the various countries. There are papers in the health and economics literature that assert or assume that equity has been more or less achieved. It would take much too long to explain why we disagree with all these papers, so a couple of examples must suffice.

Propper (2001), in an important review of UK health care policy, saw regions approaching equity. But her measure of this was the extent to which actual regional health expenditures were converging on the formulae determinations (as previously mentioned, scope was allowed for phasing change). So she was defining equity as compliance with the formulae, that is, implicitly assuming they were well founded. She also reviewed the international literature on equity attainment and while quite sceptical about the quality of a lot of it, quoted the multi-country findings of Van Doorslaer et al (2000 and earlier) in support of equity attainment elsewhere. She was critical of the neglect of a quality dimension and doubted that equity in expenditure implied equity in health service deliveries let alone health outcomes. Even so, she was probably crediting Van Doorslaer et al with more evidential value than they deserve.

These multi-country studies used data form household surveys in the various countries to perform huge regression analyses of health expenditures[19] on income and other socio-economic variables, and on needs variables like age, gender and self assessed health status. They concluded that in all countries lower income groups were more intensive users of the State funded health care systems, but that taking their greater needs into account, equity seemed achieved in most respects. While the first conclusion is probably unassailable, the second is very doubtful, if unsurprising given the methodology. The problem is that assessing need through regression of expenditure on a collection of needs variables runs straight into the

[19] Actually utilisations of health services priced by national average costs for the services.

circularity problem discussed in section III. Without 'supply' variables, or any other device to try to break circularity, the implicit assumption to begin with is that the availability of health services is equitable between households apart from purely random variation.

Of course, it may be possible to overcome some of the issues we have reviewed in this paper. The notion of commencing from regional populations must be sound and the next and key step is to find unbiased ways of adjusting for differences in the (average) health service requirements of individuals in the different regions. The major problems seem to arise through trying to use existing health expenditures to deduce the adjustments. Given suitable data, it is certainly possible to form estimates of a person's likelihood of various health problems and, presumably, consequent treatment needs[20], from knowledge of their personal circumstances and characteristics. Actuaries perform this kind of estimation for many situations where relative risks require assessment. The potential value of such studies based on data on individuals is often mentioned by authors, including indeed, Smith et al (2001). Clinical, epidemiological and observational studies (perhaps even from other countries), as well as hospital record data, might all have a role. Presumably both York (1994) and HHCRU (1997) felt that data gaps precluded this approach and, of course, they believed their theoretical formulation permitted them to proceed using utilisation and supply variables.

So if the Republic should aspire to a formula based regional allocation strategy, some hard thinking about the objectives and their likely attainment would be called for. We must not slavishly imitate the UK approach, but we can learn from their experience and especially from debates such as that following Smith et al (2001). Econometric theory and equations can disguise, even if unintentionally, assumptions and value judgements with which many might be uneasy. We need to recognise that equity is an elusive entity, hard to measure, let alone achieve. If we decide to pursue equity, via formulae, we should try to produce a database that will enable us avoid the circularity and other problems that have dogged the UK approach.

The existence of substantial regional inequity and geographic[21] inequity within regions, must have implications for that viewpoint on equity that is heavily focussed on income. The sort of comment
"If there are two persons in equal need of medical care, it would be considered undesirable if the richer of the two were to receive treatment" (Wagstaff, Van Doorslaer and Paci, 1991) not only assumes a supreme primacy for equity, but even in that context, must hang on a supposition that regional or geographical inequity is either non-existent or else that region is a perfect proxy for income. If in a less favoured region the rich choose to pay private providers to obtain the same measures of health care as are publicly provided in a well favoured region, they are improving equity, partly by raising their own level of care and also by saving resources for the others in their less favoured region. For example, the less favoured region may have long waiting times for public treatment, leading those who can afford it to opt for private provision rather than queue, which they might not do if in a favoured region. It is possible some authors, holding the viewpoint typified by the quotation, may believe there is total 'crowding out' of public by private provision - that is, anyone getting private care is depriving someone else of public care. But if so, they ought to try to explain why such a scenario could be credible.

Perhaps more importantly, the effects on equity of policies to promote efficiency can be similarly debated. Propper (2001) reviewed the literature assessing some recent innovations

[20] Although some authors (e.g. Bevan, 2001) have argued that discretionary choices by physicians can be a considerable source of variation.

[21] Note that the York/HHCRU analyses found there was large variation between wards within regions, unexplained by any variables. That is, wards with the same needs variables could differ greatly in utilisations. What about variation between individuals with the same needs? It is almost always the case that the more disaggregated the level of data the greater the variation.

in the UK health service. At regional level these included permitting Boards to purchase health service provision from other regions or from private providers and conversely to allow a region's public providers to bid to service other regions. Within regions, some GP practices were grouped into GP Fundholders – who not only provided GP services, but purchased hospital treatment etc. as they deemed fit, so that they had considerable discretion over the spending of the total funds allocated to them and could retain funds in their practices.[22] Propper reported that competition and market type forces do seem to improve efficiency in the sense of reducing costs and (in the case of GPF's) reducing hospital referrals. Now, for example, had there been equity in geographic areas at the start, before formation of GPF's it could be argued that the reduced hospital utilisations on GPF patients compared to normal GP practice patients was increasing inequity (at least as measured by expenditure). But if areas were inequitable to begin with, it could be argued that the incentive for doctors to form GPF's and their subsequent reductions in hospital referrals would be greatest in resource favoured areas. This would reduce, not increase, inequity. Of course, quality of service is not being taken into account here, but as remarked in previous sections, there are measurement problems to doing so.

There is also the theme raised at the close of the previous section. Allocation formulae certainly fit into a 'command economy' approach to health care. If health care was treated as just another market commodity to be provided by the private sector there would be no modelling or formula formulation problem. However, that is an unrealistic, and probably undesirable, extreme. But the possible future context of a substantial public/private mix of provision and funding differs from that in which the UK formulae were developed and that requires consideration. These themes need much more development, but we as econometricians, not health policy analysts, have probably taken them as far as we should. If development of a formula based regional allocation approach to health care funding is visualised for Ireland, the proposal will deserve scrutiny by a multidisciplinary team, able to see the allocation issue within the context of plausible scenarios for the future health care sector.

[22] These GPF's could be seen as modest adaptations of health organisations more familiar in the US.

REFERENCES

Barrow, M. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 244.

Barrington, R. (1998) "The future political, legislative and social framework of the health services", in (eds. Leahy, A. L. and Wiley, M. M.) *The Irish Health Care System in the 21st Century*, Dublin: Oak Tree Press.

Benzeval, N., Judge, K. and Whitehead, M. (1995) "The Role of the NHS", in (eds. Benzeval, N., Judge, K. and Whitehead, M.) *Tackling Inequalities in Health: an Agenda for Action*, London: King's Fund.

Bevan, G. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 243-244.

Blundell, R. and Windmeijer, F. (2000) "Identifying demand for health resources using waiting times information, *Health Economics*, 9, 465-474.

Carr-Hill, R. A., Hardman, G., Martin, S., Peacock, S., Sheldon, T. A. and Smith, P. (1994) *A Formula for Distributing NHS Revenues based on Small Area Use of Hospital Beds*, York: University of York.

Carr-Hill, R. A., Dixon, P., Browne, S., Stevenson, M., Jamison, J. and O'Reilly, D. (2000) Sheldon, T. A. and Smith, P. (1994) *Development of Needs Indicators for the Elderly Programme of Care in Northern Ireland*, Belfast: HSCRU QUB.

Chisolm, M. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 248.

Commission on Health Funding (1987) *Report*, Dublin: Stationery Office

Conniffe, D. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 248-249.

Coopers and Lybrand (1988) *Analysis for the Review of RAWP*, London: Coopers and Lybrand

Department of Health (1986) *Health the Wider Dimensions*, Dublin: Stationery Office.

Department of Health (1994) *Shaping a Healthier Future – A Strategy for Effective Health Care in the 1990s*, Dublin: Stationery Office.

Department of Health (1998) *Independent Inquiry into Inequalities in Health Report* (Acheson), London: The Stationery Office.

Department of Health and Children (1997) *Statement of Strategy*, Dublin: Stationery Office.

Derbyshire, M. (1987) "Statistical rationale for grant related expenditure assessment (GREA) concerning personal social services" with discussion and reply, *Journal of the Royal Statistical Society* **A**, 150, 309-333.

Derbyshire, M. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 241-242.

Enthoven, A. (1988) *Theory and Practice of Managed Competition in Health Care Finance*, Amsterdam: North Holland.

Fahey, T. (1998) "Population ageing, the elderly and health care", in (eds. Leahy, A. L. and Wiley, M. M.) *The Irish Health Care System in the 21st Century*, Dublin: Oak Tree Press.

Gillon, R. (1986) *Philosophical Medical Ethics*, Chichester: Wiley.

Glennerster, H. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 249.

Goldstein, H. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 242-243.

HHCRU - O'Reilly, D., Carr-Hill, R. A., Reid, J., Browne, S., Jamison, J., Stevenson, M.and Merriman, B. (1997) *Report of Study to devise a Formula to Assist in Allocating Resources for Acute Hospital Services within Northern Ireland*, Belfast, HHCRU, QUB.

Leyland, A. H. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 245.

Manning, W. G. (1987) "Health insurance and the demand for medical care; evidence from a randomised experiment", *American Economic Review*, 77, 251-277.

McKeown, T. (1976) *The Role of Medicine*, London: Nuffield Provincial Hospital Trust.
Mooney, G. (1986) *Economics, Medicine and Health Care*, Brighton: Wheatsheaf.
Nolan, B. (1993) *Charging for Public Health Services in Ireland: Why and How?* Policy Research Series No. 19, Dublin: ESRI.
O'Dwyer , J. (1998) "Reflections on future structures in the health services", in (eds. Leahy, A. L. and Wiley, M. M.) *The Irish Health Care System in the 21st Century*, Dublin: Oak Tree Press.
Propper, C. (2000) "The demand for private healthcare in the UK: a review of the issues", *Journal of Health Economics Studies*, 19, 855-876.
Propper, C. (2001) "Expenditure on healthcare in the UK: a review of the issues", *Fiscal Studies*, 22, 151-183.
Rice, T. (1998) *The Economics of Health Reconsidered*, Chicago: Health Administrative Press.
Sanderson, C. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 252.
Sheldon, T. A. and Carr-Hill, R. (1992) "Resource allocation by regression in the National Health Service: a critique of the resource allocation working party's review", *Journal of the Royal Statistical Society* **A**, 155, 403-420.
Smith, G. D., Frankel, S., and Ebrahim, S. (2000) "Rationing for health equity: is it necessary?" *Health Economics*, 9, 575-579.
Smith, P. C., Rice, N. and Carr-Hill, R. (2001) "Capitation funding in the public sector" with discussion and reply, *Journal of the Royal Statistical Society* **A**, 164, 217-257.
Sutton, M. and Lock, P. (2000) "Regional differences in health care delivery", *Health Economics*, 9, 547-559.
Tudor-Hart, J. (1971) "The inverse-care law", *Lancet I*, 405-412.
Tussing, A. D. (1982) "Health, health policy and poverty", in (eds. Joyce, L. and McCashin, A.) *Poverty and Social Policy*, Dublin: IPA.
Van Doorslaer, E., Wagstaff, A. et al (1992) "Equity in the delivery of healthcare: some international comparisons", *Journal of Health Economics*, 11, 389-411.
Van Doorslaer, E., Wagstaff, A. et al (2000) "Equity in the delivery of healthcare in Europe and the US", *Journal of Health Economics*, 19, 553-583.
Wagstaff, A., Van Doorslaer, E. and Paci, P. (1991) "On the measurement of horizontal inequity in the delivery of health care", 10, 169-205.
Wagstaff, A., Van Doorslaer, E. et al (1992) "Equity in the finance of health care: some international comparisons", 11, 361-387.
Wright, N. (2001) "Contribution to discussion of Smith, et al ", *Journal of the Royal Statistical Society* **A**, 164, 246.