Douglas McCulloch has been a lecturer in economics at the University of Ulster since its creation in 1984, and at the Ulster Polytechnic since 1975; his PhD was awarded in 1998. In that time, he has also worked in the Northern Ireland Government service, and at the Irish Centre of Pharmacoeconomics in Dublin, as a health economist. Publications include numerous articles, a textbook on health valuation, and the Irish Health Technology Assessment Guidelines. He is Chairman of the Health Economics Association of Ireland.

Human Economics - a response

DOUGLAS W. MCCULLOCH

Frank Rotering's theory of human economics is an unrivalled analysis of the sustainable economy. Bringing together insights from several fields besides economics, it is a bold essay on the problem of developing the economics of sustainability. Douglas McCulloch gives his reactions to the "Rotering Papers", based on subject expertise and more than twenty years experience of teaching economics.

The "Rotering Papers" describe a bold and interesting approach to a general economic theory of sustainability, which is probably the largest single topic available for study by economists. The theory (hereafter "Human Economics") faces considerable challenges, as well as being (as Rotering recognises) a very large enterprise indeed, but one on which a beginning has been made. Taking into account the Feasta web-site discussion as well as the main papers, we look at the structure and approach of human economics, and then discuss the human framework, health measurement, and the economist's tools.

The Structure and Approach of Human Economics

The aim of Human Economics is to "formulate economic concepts and analytical tools that permit the maximization of human well-being, subject to ecological constraints"; this is ambitious. Given this aim, the approach taken - to divide the subject matter into human, ecological, and functional frameworks - does seem to make sense, though the origins of this structure are not clear. It is rather difficult to stand far enough back to get a detached view for a theory of such a large area.

While the breadth of the approach is impressive, is it the best way forward? Rather than working out a theory of the entire sustainable economy, it might be more productive to make progress with the sustainable economics of particular sectors, for example, to devise a strategy to make the use of energy in Ireland sustainable. Would the value of such an analysis be worthless just because it was not part of a larger whole? In particular, the data categories developed might be expected to be objective, and of lasting value, whether or not a strategy based on them was ever implemented. Working on different parts of the general problem may be more productive than tackling the whole system, both in terms of actual environmental improvements achieved, and in terms of eventually developing a whole system. What is common and general may be more likely to emerge from several separate fields than from an attempt to solve the very large problem of achieving system wide sustainability in a single theory. Working on specific sectors would mean more interaction with the world, testing our ideas more often, and impacting more on those who do not share our view of the importance of sustainability.

The Human Framework

As the paper on Human Economics states, the adoption of a single standard for value and cost follows standard economics. No explanation is given for taking this particular approach, which is surprising, because there is widespread disquiet about the focus on the "bottom line" throughout the private sector. In the United Kingdom this focus has now spread to the public sector, in the form of targets and indicators; the brutal application of these measures is severely damaging the UK's health and education sectors. When people focus on a single measure, wider considerations are left off the agenda, and human and environmental damage can result.

Presumably, maximising human health is not possible without a single measure, because costs have to be traded-off against benefits, as we find in conventional cost-benefit analysis and in Human Economics. This assumes that different values (equity, intrinsic values of habitats, aesthetic values) can be reconciled into a single agreed value, and that production functions can be identified, throughout the range of output levels, for each production process. These are very large assumptions, which tend to ignore the political problems of progress towards sustainability.

Also, the expansion of each production activity to the point where health impact is maximised assumes that the production activities and technologies have been decided, and that information about them can be collected and acted upon. There is an infinite number of different ways of combining technologies to produce what we need; even with modern computers, it may not be possible to select the least cost production techniques from this array, and then calculate the optimal levels of output of each according to a health metric. Pro-market economists do have a point, when they say that markets generate information without cost; the information requirements of Human Economics are enormous.

We now turn to the central problem of measuring health.

Measuring Health

Health is described (in the paper "Human Economics") as objective and quantifiable; the theory relies on this assumption to determine the levels of output of each consumption product in the economy. It follows that if human health cannot be objectively quantified, human economics will have to think again about its measure of human welfare. On the basis of the work done in health economics, it appears that human health cannot be objectively measured. Since it seems that this work must be relevant to the development of Human Economics, we consider health outcome measurement, and go on to the idea of objective medicine.

Health Outcome Measurement

When the prospect of limitations on national health budgets became evident in the late 1960s, economists turned their attention to the measurement of health, with the aim of informing the choices between medical interventions. For example, at that time, considerable resources were being applied to providing surgical treatments of heart disease (bypasses and heart transplants), and considerably fewer were being allocated to hip replacements. While cardiac disease was life-threatening (and more so then than now, because of drug development since), many times more people needed hip

replacements, were in debilitating pain, and could have had their quality of life restored by a fairly simple and relatively inexpensive procedure. The main medical priority was saving lives; in resource allocation decisions, quality of life tended to be ignored, and heart surgery won more resources. Hip replacements were adequately funded later; the raising of the issue of quality of life effects by economists[1] probably had some influence on this development.

This is only one example of the prioritisation problems of health services in developed countries, which are aggravated to-day by the complexity of treatments, the variety of conditions to be treated, and greater longevity. As health economics developed, it appeared that prioritisation might be improved by the measurement of a procedure's impact on quality of life and on survival, amalgamated into a single measure, the "quality-adjusted life year" or QALY. Over the past thirty or so years, several QALY measures have been developed; perhaps the most successful is the EQ-5D, the result of the work of the "EuroQol" team[2]. None of these measures is objective; all rely on patient or carer evaluation of health states, and, despite a generation's work, no QALY measure has achieved the status of an accepted objective measure of health which permits the comparison of all interventions in cost per QALY terms.

There is a whole literature on QALYs, which we leave to one side[3]; it contains important arguments about principles, which have not been resolved[4], though not all of them are relevant to the current discussion. We begin with a definition of the QALY, and go on to discuss the reliability and validity of QALY measures.

The Quality-Adjusted Life Year

The assumption is that patient health states before and after all the interventions between which resources are to be allocated can be described. Further, it is assumed that these states can be valued on a cardinal scale, from zero (death) to unity (full health). If H1 and H2, and S1 and S2, respectively, are the average health state values and expected years of survival before and after a particular intervention, then the value in QALYs of the intervention is: (H2 x S2) - (H1 x S1).

For example, suppose two cancer drugs are the same in every respect (including cost), except their impacts on survival and health-related quality of life (hrqol). On the average, Drug A increases hrqol from a health state valued at 0.6, to one valued at 0.8, and increases survival from 2 years to 3 years; therefore, it produces:

(0.8 x 3) - (0.6 x 2) = 1.2 QALYs.

Drug B increases hrqol from 0.6 to 0.7, and increases survival from 2 years to 4 years; it produces:

(0.7 x 4) - (0.6 x 2) = 1.6 QALYs.

If A and B cost the same, the QALY measure indicates that B gives more health per unit of resources, and is the preferred option.

This measure seems to be pretty close to what is required for the Human Economics project; if QALYs are not objective (reliable and valid) then the human health

measure proposed in the Rotering Papers is also unlikely to be objective. We consider the objective nature of the QALY in terms of validity and reliability.

Validity

A measure is said to be scientifically valid (to have "criterion validity") if it can be shown to measure what it purports to measure; this means that we need another measure to "validate" it. There is no "gold standard" measure of health-related quality of life[5]; criterion validity for such a measure cannot be found.

Also, economists have criticised the QALY approach on the grounds that maximising QALYs means ignoring outcomes which the QALY measure does not capture, such as transport costs, carer time and health, or productivity effects. When QALY values are developed and applied, only the chosen aspects of health and the values expressed by the original respondents are taken into account. Similar problems seem likely to arise with the human health measure proposed for Human Economics.

Reliability

A reliable instrument delivers the same or similar results from its administration on different occasions, or by different observers, to the same or similar individuals. We would expect a QALY measure to be reliable, according to this definition, but the subjectivity of health state valuation may give rise to problems.

Suppose a sample numbering 10 per cent of the UK population is chosen, in such a way that it represents all the UK's age, sex, and disease groups. This would be a vast undertaking, but it would give us a sample size adequate for almost any conceivable scientific purpose. Now suppose a set of health states is specified, and a valuation exercise is carried out, to determine the sample's mean valuation of each health state, on a scale from zero to one. Would the values found be reliable? The answer has to be "no", because while these values would be representative of the values of the population at the time of the sample, there is no good reason to suppose that a repetition of the exercise would find the same values. Despite the very large sample we have imagined, the fact that individuals can change their minds implies that we are not observing an unambiguous variable, like the number of cars sold, but recording the outcome, on one or more occasions, of individuals' judgement processes. The normal rules of statistics do not apply to this kind of situation. The publication of a newspaper article, or a harrowing television documentary, could alter the values from the sample; we simply do not know. Unless more comes to be known about health state valuation, it seems unlikely that scientifically reliable QALY values can be obtained.

Conclusion

However large the sample of respondents and whatever the research method, the statements and values established cannot have objective status. At best, one can say that a set of values has been established in a rigorous and (as far as possible) scientific process, with each part of the process explicitly defined and followed. Even given

such conditions, how is the chosen process to be justified? It seems that every possible process could be criticised, because there is no objective basis on which to stand the method adopted.

These problems seem likely to apply to a measure devised as the foundation for Human Economics; they may be even more difficult to resolve, because of the much wider application of the health measure in that theory.

Objective Medicine?

How scientific are physicians? The basis of scientific medicine is the randomised controlled trial (RCT), in which the treatment and the control group are compared under controlled conditions, to be certain that the new drug (or procedure) really does have the benefical effect indicated. However strong the results found in a trial, the physician usually has to make a judgement about whether or not to apply a treatment to a particular patient. Trial group patients have been carefully selected to make sure they are as similar as possible. To get statistically significant results, sample size must be as large as possible, and the effect must also be as large as possible (if it is present), so it is important to refine the selection of the groups. This can mean that the trial groups make up a small sample of the population, and are untypically homogeneous, to avoid the influence of third (or fourth, or fifth…) factors on the effect found. The physician seldom has scientific results which apply directly to the patient in front of him/her, because few patients conform to the trial protocols. There is even a small literature calling into question the basis and the validity of the RCT procedure[6].

Furthermore, the physician may have to allow for the superior effectiveness and resources of the doctors carrying out the trial; for example, they might have had better equipment than is usually available to look after patients who develop side-effects. Should the physician prescribe the drug in the absence of the same facilties?

As an example, it is reported[7] that those patients included in cardiac invasive procedure trials were only a small proportion of those referred for invasive treatment. In the Coronary Artery Surgery Study, for example, only 12.7 per cent conformed to the trial's inclusion criteria, and of these only 37 per cent were randomised, amounting to only 4.7 per cent of all cardiac referrals. The conclusion is inescapable: physicians have to make life or death decisions based on evidence which has to be interpreted, not simply applied. It is their judgement, as much as their science, which saves our lives and improves our health; that judgement is unavoidable, though it is of course informed by scientific evidence.

An additional question arises with respect to the genetic differences which are being found between human beings, for instance, in the Human Genome Project. How standard is the human organism? If it is not standard, can its health be measured universally on the same basis, and can the health effects of production processes be identified and measured?

It appears that, in some individuals, there may be a genetic predisposition to cancer after exposure to toxic chemicals; while such chemicals are toxic if ingested in any quantity, for most people, exposure to the very small quantities usually involved is harmless. However, perhaps one person in one thousand will develop cancer even when exposed to the typical very small quantity. If only one person in a thousand has this predisposition, an RCT is not going to pick it up, because RCTs search for an effect in the population as a whole, and usual sample sizes will not be large enough to

detect it. If an objective measure of health can be developed, it may be difficult or impossible to measure the impacts of production processes on human health, because such effects cannot be determined.

The Economist's Tools

According to the discussion, some of the tools of economics can be used by those adopting the value of sustainability, while others must be rejected. The grounds for selection are unclear; for example, what is meant by "ideologically suspect" (a term used in the discussion)? If some ideas are to be discarded, we have to be careful about what we keep, and what we reject, and we should know the grounds on which either decision is taken.

One example is the idea of opportunity cost, which is central to any kind of economics; it is neutral, in my view, though this is contradicted by the support paper on economics. There, the idea is considered to apply to inputs only, and it is rejected on the grounds that such inputs (measured in money) do not include the human costs of production. My version of opportunity cost is the following. Scarcity means that, from a set of alternatives, we have to choose one, because that is all which the available resources permit us to have. When we choose one, it is, to us, the best alternative. The opportunity cost of this alternative is the one which we value next best. The real cost of our choice is this "next best" alternative. In the real world, costs are values, and values are costs.

This does not sound exactly earth-shattering, but it emphasises the fact of scarcity, and requires the clarification of alternatives. The very idea of choice is the first challenge to those public and private sector individuals and institutions who would much prefer to leave the alternatives undefined, so that the choices they have made are not questioned.

Looking more widely, we may need the tools of conventional economics if we are to challenge the justification of market forces. In the appropriate context, externalities, economies of scale and public goods are important considerations which have to be weighed in the balance against the benevolence of competition which is so often assumed or even propagandised. Similarly, the defects of GDP as a measure of welfare have been recognised for many years. It follows that conventional economics does contain ammunition with which to challenge its abuses; my own view is that if we do rule an idea out, we should be clear about why it is unsound.

There seems to have been quite a reaction to the graphs in the papers; I found them quite impressive, but my teaching experience has made me aware that those who do not have substantial experience of economics often find such analysis difficult. Simply grasping the idea of a function can be difficult for some of my students, and then understanding that the actual meaning of the diagram may not be precise, or may be true only under specific assumptions, is a step too far for many of them.

It is perhaps a worry that some of this sympathetic audience did not see what the variables were, or at least found it difficult to relate the variables to their own hypotheses about the world. It is not obvious, either, that graphs are always superior to words; the key test is surely what best gets the meaning across. If people do not understand, the argument has not been made, and discussion cannot begin.

Conclusion

The objective we share is progress towards a clearer view of how the sustainable economy might work, and the Rotering papers have made a large contribution to that. The main problem in the papers is the measurement of health, on which the theory relies; the problems identified by health economics seem obviously to be relevant to Human Economics. It also seems relevant that, if health could be measured objectively, the pharmaceutical multinationals would have found a way. Most developed countries now have in place guidelines for the acceptance of new compounds which require the companies to demonstrate cost-effectiveness using a measure of health outcome. Any objective measure which enabled the companies to clear this "fourth hurdle" would have been developed and in use by now.

Another serious problem is the cost of information, which is common to Human Economics and (system-wide) health economics; prioritising health provision, system-wide, when health is free of charge at the point of use, is a sub-set of the problem specified in Human Economics. Prioritising on the basis of the cost per QALY requires large assumptions about what is known (or could be known) about costs and outcomes, especially the relationship between each and the level of production, across the whole range of possible outputs.

Considering production further, what is produced depends on consumer wants; in determining the mix of production technologies, we have to know what wants we are meeting. The division between consumption and production may, for a sustainable economy, be an "ideologically unsound" assumption of conventional economics. The taking of holidays (consumption) would not be necessary if people's working lives (production) were adequately healthy and congenial; the production of health services would be much less if many individuals' consumption patterns (diet, drugs, alcohol) were different. There are many possible examples. The division in our thinking between consumption and production may conceal important possibilities for re-structuring our economies towards sustainability.

Finally, while I am not sure that a global approach of Human Economics will be as productive as the development of policy frameworks and strategies for particular sectors, I am convinced of the value of the Rotering papers. They constitute a substantial contribution to the working out and the eventual development of the sustainable economy, whatever form it takes, and however we get there.

NOTES

1. For example by Williams (1985), at a consensus conference of the British Medical Association on the Coronary Artery Bypass Graft procedure

2. There is a substantial literature, which includes applications of the EuroQol to different treatments and conditions, such as stroke. Two good starter articles are: Brooks, R. with the EuroQol Group (1996), "EuroQol: the current state of play", *Health Policy,* Vol. 18, pp. 870-877; and Euroqol Group (1990), "Euroqol - a new facility for the measurement of health-related quality of life", *Health Policy* 16(3), pp. 199-208.

3. See the references.

4. For example, see chapter 3 in *Economic Evaluation in Health Care: Merging Theory with Practice,* by Drummond M and McGuire A (Oxford University Press 2001).

5. There are many sources which support this important argument, which is accepted by most health economists; the ones I have selected are:

   Kind, P. and Rosser, R. (1988), "The quantification of health", *European Journal of Social Psychology,* Vol 18 pp 63-77
   Nord, E (1992) "Methods for the quality-adjustment of life years" *Social Science and Medicine,* Vol 34(5), pp 559-569
   Richardson, J (1994) "Cost-Utility Analysis: What should be measured?" *Social Science and Medicine,* Vol 39(1), pp 7-21
   Whynes, D.K., and Neilsen, A.R. (1993) "Convergent validity of two measures of the quality of life" *Health Economics,* Vol 2, pp 229-235.

6. The references are:

   Bailey, K.R., (1994), "Generalising the results of randomised clinical trials", *Controlled Clinical Trials,* Vol. 15, pp 15-23.
   Davis, C.E. (1994), "Generalising from clinical trials", *Controlled Clinical Trials,* Vol. 15, pp. 11-14.
   Fayers, P.M. and Hand, D.J. (1997), "Generalising from phase III clinical trials: survival, quality of life, and health economics", *The Lancet,* Vol. 350, pp 1025-1027.
   Herman, J. (1995), "The Demise of the Randomized Controlled Trial", *Journal of Clinical Epidemiology,* Vol. 48(7), pp. 985-988.
   Rothwell, P.M. (1994), "Can overall results of clinical trials be applied to all patients?", *The Lancet,* Vol. 345, pp. 1616-1619.
   Rubins, H.B. (1994), "From clinical trials to clinical practice: generalising from participant to patient", *Controlled Clinical Trials,* Vol. 15, pp, 7-10.

7. Gunnell, D. and Smith, L. (1994), *The Invasive Management of Ischaemic Heart Disease,* Health Care Evaluation Unit, University of Bristol.

REFERENCES

Barer ML, Getzen TE, and Stoddart GL *Health, Health Care, and Health Economics* Wiley 1998.
See Uwe Reinhardt in this.

McCulloch D *Valuing Health In Practice* Ashgate 2003
Assumptions are examined, empirical work is discussed, and a new approach to prioritisation is developed.

Nord E *Cost Value Analysis in Health Care* Cambridge University Press 1999
Probably the most important criticism of the idea of a health outcome measure based

on a cardinal scale.

Sloan F *Valuing Health Care* Cambridge University Press 1995 Conventional scientific approach to health outcome measurement.

Williams, A. (1985), "Economics of Coronary Artery Bypass Grafting", *British Medical Journal*, Vol. 291, pp 326-329
The most famous QALY article.