Timid, unduly respectful, naive and conventional – a critique of the quantitative programme in Social Policy – David Byrne

dave.byrne@durham.ac.uk

Stream: Theoretical and methodological innovation

Social Policy must have a quantitative programme. It is impossible to grasp any of: the nature of the societies in which social policies operate, the character of social interventions, or the effects of those interventions, without recourse to measurement. Likewise, whilst quantitative work is not the only way in which we can explore causality, it remains a crucial part of any causal programme. This applies both in terms of overall critical description of the way in which social policy interacts with the social world and with particular force in the evaluation of specific policy and practice interventions. Such exploration of cause requires modelling. So we have to measure and we have to model. The argument in this paper is that quantitative social policy does both these things badly. And it does so in large part because is has relied on the usually, misconceived quantitative programmes of other disciplines and fields and has accorded to those programmes a wholly undue measure of disciplinary subordination and respect. Moreover, it has, with some exceptions, largely ignored recent quantitative methodological innovations in relation to measurement and the exploration of cause.

The paper will argue the following:

1. The quantitative programme of contemporary Economics is almost wholly useless in relation to social policy issues. This derives from Economics' fundamental reliance on posited equilibria as a foundation for modelling and on the development of linear models as a way of exploring cause in economic relations. Neither element is commensurate with the reality of causal relations in a complex world.
2. The rhetoric of the supremacy of experimental methods which is the foundation of the unequivocal assertion by the Campbell Collaboration (the Cochrane Collaboration is a little more tentative) of a hierarchy of evidence, is not only misplaced in relation to much of clinical practice, but wholly wrong in relation to the generation of quantitative evidence in policy and practice interventions in a complex world.
3. There is an alternative quantitative programme available which draws on understanding measurement as primarily a process of classification, recognizes the value of set theoretic methods as opposed to reductionist variable analyses, and understands that what is required is a quantitative framework for the description of complex systems and the exploration of change in the trajectories of those systems through time.

This paper is about causality – about how quantitative methods are deployed in trying to establish causal relationships in relation to social policy issues. It agrees absolutely with Dannermark et al's prescription:

> 'The explanation of social phenomena by revealing the causal mechanisms that produce them is the fundamental task of social research.' (2002 1)

So it will proceed to examine how this task is undertaken and consider how we might, as applied social researchers, do it differently and do it better.

I want to begin by examining what Social Policy academics actually do in terms of quantitative methods in relation to social causation. As a way into this I have copied the approach of Payne et al. (2004) and looked to journal content. However, whilst they reviewed a range of sources over four years, I have confined my attention to the *Journal of Social Policy* but have examined all issues since 2000, that is Volumes 29 to 37 in toto and the two issues of volume 38 to hand when I wrote this piece. Table 1 enumerates my results descriptively and serves both as a summary of my classification and enumeration and as an illustration of what here will be called "simple tabulation".

Table One – Pattern of articles published in JSP 2000-2009

TOTAL ARTICLES	238
% of total articles quantitative	37
% of quantitative articles with simple tables	84
% of quantitative articles with complex tables	12
% of quantitative articles with modelling	28
% of quantitative articles using econometric approaches	2
% of quantitative articles using numerical typology approaches	2
% of quantitative articles using longitudinal data	7
% of quantitative articles using factor analyses	3

The above is very rough and ready but perfectly adequate for purposes of illustration. I include as quantitative any article where numbers are presented in a formal way as part of the account and / or argument. Simple tables are tables of frequencies i.e. counts or percentages as in Table One. Complex tables involve relating one variable to another in a contingency table and no article did this for more than the two variable cross tabulation. Modelling involved the development of a causal model using some variation of regression techniques. Econometric usage was any development of modelling in which coefficients having been estimated were used to project results. Numerical typologies were classifications using clustering techniques. Longitudinal examples involved the use of both longitudinal data and the quantitative exploration

of trajectories of cases over time. Factor analyses were the identification using factor analytic techniques of 'underlying causal factors' in relation to measured indices describing cases. There is substantial and deliberate double counting in the above table. In other words each occurrence of an approach was counted and many articles incorporated multiple quantitative approaches. In particular almost all articles incorporating other techniques also included simple illustrative frequency / percentage count tables.

It is interesting to compare this examination of the UK's leading Social Policy journal with Payne et al.'s results for Sociology journals. They found only 21% of articles in the main stream Sociology journals had any quantitative content although this figure was 55% for *Work, Employment and Society.* They also did not examine, as they admit, important Sociology or Sociology related journals in health and related fields and it may well be that the proportion of quantitative articles would be higher in such journals. The higher percentage of quantitative pieces in *WES* reflected the degree to which articles in that journal drew on major employment related surveys, although there are also examples of more sophisticated methods, including causal modelling, to be found there.

Let me dispose first of those articles in *JSP* which contained reportage from no quantitative method more advanced than simple tables. This set included more than half of all the articles with any quantitative element. Here typically one or more uni-variate i.e. frequency, tables were used to illustrate the characteristics of a set of cases. The cases could be individuals, households, institutions, or even nation states. The quantitative information was used to indicate the overall character of the group of interest and then typically the article proceeded to a qualitative account deploying historical material based on documents and / or the products of interviews with a rather limited use of ethnographic observation. This approach is absolutely fine and I have no criticism of it. It is essentially the use of quantitative information to enhance description and hence understanding of the objects of interest. My criticisms are addressed at the approaches which attempt to go beyond this into some kind of account of causality.

It is perhaps important to point out that the distinction between description and causal account is by no means absolute. Many accounts conflate the two by using quantitative description as part of a textually developed argument describing how particular conditions come to be. These conditions can be the states of any level from individual through household, neighbourhood, locality, region, nation, block (e.g. EU) up to the global system as a whole. That hierarchy is constructed in terms of units beyond the individual which have a spatial component but we could equally talk of classes, schools, education systems or of GP practices and secondary units, local health economies, and national health care systems. There is absolutely nothing wrong with this. It is a perfectly valid mode of representation of the dynamics of social reality. However, it is interesting to note that for reasons which we can equate with the privileging of 'scientism' in social accounts, such integration of quantitative description into qualitative accounts of causality are both formally and often in practice regarded as inferior modes in contrast with formal modelling using mathematical tools. We can see this explicitly in the hierarchies of evidence asserted by the Cochrane Collaboration and although the Campbell Collaboration is somewhat more tentative, it still tends towards the same conception of what counts as evidence.

The focus of my criticism is directed at quantitative approaches which are fundamentally founded around the logic of explanation of 'The General Linear Model'. Abbott has distinguished between what he calls the 'representational' and the 'entailment' deployment of general linear modelling thus:

> 'The phrase 'general linear reality' denotes a way of thinking about how society works. This mentality arises through treating linear models as representations of the actual social world. This representational usage can be opposed to the more cautious use of linear models in which the analyst believes that some substantive causal process logically entails patterns of relations between variables, patterns which can then be tested by that model to discover whether the actual state of affairs is consistent with the substantive mechanism proposed [the entailment model].' (2001 38)

I have to say that I am not convinced by this distinction although it has some heuristic value. In other words it seems to me that even the entailment model inherently subscribes to the position which Abbott outlines absolutely specifically:

> 'To use such a model to actually represent social reality one must map the processes of social life onto the algebra of linear transformations. This connection makes assumptions about social life: *not* (original emphasis) the statistical assumptions required to estimate the equations, but philosophical assumptions about how the social world works. Such representational use assumes that the social world consists of fixed entities (the units of analysis) that have attributes (the variables). These attributes interact, in causal or actual time, to create outcomes, themselves measurable as attributes of the fixed entities. The variable attributes have only one causal meaning (one pattern of effects) in a given study, although of course different studies make similar attributes mean different things. An attribute's causal meaning cannot depend on the entity's location in the attribute space (its context), since the linear transformation is the same throughout that space. For similar reasons, the past path of an entity through the attribute space (its history) can have no influence on its future path, nor can the causal importance of an attribute change from one entity to the next. All must obey the same transformation.' (2001 39-40)

It is really very important to appreciate the significance of this denunciation of linear modelling. We can best do this by taking a complexity fix on Abbott's formulation. Another word for entity would be case – the actual 'thing' which possesses properties. Ragin and Becker (1992) asked 'what is a case?' Ragin in that collection proposed that we need to pay very serious attention to processes of 'casing' – we need to recognize that just as traditional consideration of the quantitative approach in social science urges careful attention to operationalization – to the processes by which we construct variables through our measurement acts – so we have to recognize that our processes of definition construct cases. This is not to say that cases are simply reified constructs, although in some instances they may be. Here Cilliers' account of the simultaneous reality and construction of our specification of complex systems is directly applicable:

> Boundaries [of complex systems] are simultaneously a function of the activity of the system itself, and a product of the strategy of description involved. In other words, we frame the system by describing it in a certain way (for a certain purpose) but we are constrained in where the frame can be drawn. The boundary of the system is therefore neither a function of our description, nor is it a purely natural thing. (2001 141)

In order to not over-complicate this discussion let us accept the reality of the cases we are dealing with which is perhaps easier in policy context since our cases tend for have a pre-existing, if fuzzy, form. For example the cases might be institutional entities such as Local Strategic Partnerships, Schools, Local Government Units. They might be individuals or households – and households have a particularly dynamic character. However, over time, even institutional entities change. Administrative reorganization alters boundaries. The components of LSPs change regularly in terms of personnel. LSPs can certainly be regarded as fuzzy cases since they integrally, and indeed intentionally, intersect with other institutional entities. The entity has a continued existence but that existence is itself plastic. In other words we can regard entities quite properly as themselves being complex systems. Let us be clear about the implications of that system. The Generalized Linear Model does not grasp the complexity of social entities. Rather in line with the Newtonian origins of linear modelling, it sees entities as simple objects without inherent causal powers. Instead the causal powers reside in forces external to the entities which can be described as 'variables'. Abbott puts it like this in a passage which bears a deal of repetition:

> 'The people who called themselves sociologists believed that society looked the way it did because social forces and properties did things to other social forces and properties. …. Sociologists called these forces and properties "variables". Hypothesizing which of these variables affected which others was called "causal analysis". The relation between variables (what these sociologists called the "model") was taken as forcible, determining. In this view, narratives of human actions might provide "mechanisms" that justified proposing a model, but what made social science *science* (original emphasis) was the discovering of these "causal relationships."' (2001 97)

We can go further. The process of partialling – that is of decomposing overall causal effects so that we can assign specific effects to particular single causal variables – reinforces the whole conception that it is variables which act and social entities which are acted upon. Moreover, in line with the traditions of scientific measurement which began in the seventeenth century and which are very strongly represented in the social sciences and particularly in Economics, variables are understood to be always continuous i.e. possessing the full properties of numbers in conventional mathematical form. This means that significant change is always incremental and that causality is assumed to operated proportionately. We can always express outcomes Y in terms of an equation where $Y = f(X)$. Note f can be a power or even an exponential function, but that function is linear. Pearson, one of the founders of statistical methods, understood measurement in tetrachoric terms. That is he believed that anything which is expressed in categorical terms – in other words where the distinction we draw is one of kinds rather than of degree – actually is merely the manifestation of an underlying continuous distribution. HOWEVER, if categorical distinction is actually real in itself then we might consider, and indeed in real life as opposed to statistical

modelling usually do consider, that what matters is the kind of thing something is. It has been said, completely accurately, that:

> ' … policy researchers, especially those concerned with social as opposed to economic policy, are often more interested in different kinds of cases and their different fates than they are in the extent of the net causal effect of a variable across a large encompassing population of observations. After all, a common goal of social policy is to make decisive interventions, not to move average levels or rates up or down by some miniscule fraction.' (Rihoux and Ragin 2004 18)

Let us consider what this means in relation to the application of the methods of conventional contemporary Economics to social policy issues. I could find only one explicitly Econometric article in the JSP over the period I examined and this was actually a rather interesting, if perhaps over-elaborate use of the construction of coefficients to examine the impact of religion in relation to deprivation in Northern Ireland (Boorah 2000). In fact econometric methods are not much employed in the social policy literature as represented in JSP although they seem to have a kind of symbolic significance in relation to the discussion of social policy issues. One domain where they are deployed is in relation to health policy. An excellent discussion of econometric methods in health contexts is provided in Tappenden et al (2004).

Before going through this useful piece one thing has to be made clear. So far as I can determine – and this could provoke useful and interesting arguments – the practical i.e. applied use of health economics does not depend on the endorsement of some central principles of conventional economics as a discipline, despite frequent assertions in the literature describing the field that it is dependent on those principles. In other words the practical uses are largely about 'micro-economic evaluation at the treatment level' and the fundamental principle of neo-classical economics, that of the achievement of equilibrium state, is not relevant to this practice. It is precisely the reliance on the notion of equilibrium – the notion that markets always achieve an equilibric outcome – which renders neo-classical economics not isomorphic with reality: that is to say which renders it about as much use as a chocolate fireguard for practical purposes. The position is fully argued out in Ormerod (1994) and I will not elaborate it here. In brutal summary the world consists of far from equilibrium systems and they are characterized by non-linear transformation whereas neo-classical economics sees equilibria as natural and relies, predominantly, on the General Linear Model.

Tappenden et al. do recognize this and their approach is one which raises some interesting issues for our understanding of what 'Economics' is in practice in relation to policy and practice issues. If 'Economics' is, what economists do, then there may be some value in some of the techniques suggested, with reservations. If 'Economics' is what neo-classical economic theory says it is, then the chocolate fireguard is well in place. Tappenden et al. were reviewing: 'Methods for expected value of information analysis in complex health economic models' and the most relevant part of their discussion is an extended review of the methodological framework for undertaking this kind of analysis and in particular of issues surrounding Metamodelling. Essentially metamodelling is a response to the problem of the representation of complex systems in even the most advanced computing systems. This is often

presented in terms of the computing cost issue and this is quite real. Running full simulations of complex systems, even if those simulations are considered to be accurate representations – re-presentations – can take a great deal of time. More fundamentally for me there is a serious question as to whether simulations based on continuous measurements of 'variables' can ever be accurate representations of complex systems where the appropriate level of measurement is a categorical description of state or perhaps we might allow ordered categorical descriptions of states at most? That on one side for the moment – but we will return to this crucial issue – the problem is of course the non-linearity of systems. Let us begin by letting Tappenden et al. define meta-models for us:

> 'For this reason [the cost and complexity of direct simulations], one approach to modelling the original problem which has become more popular over the past quarter of a century is the use of metamodels as replacement for the original simulation models. Metamodels are effectively models of models, or mathematical approximations to the input and output functions of a model, and can be seen as simplifications of the original model.' (2004 19)

They go on to discuss the following approaches:

- Linear regression
- Neural networks
- Response surface methodology using polynomial regression
- Multivariate adaptive regression splines
- Gaussian processes involving non-linear regression.

The key issue is the problem of non-linearity. All the methods discussed other than simple linear regression involve efforts to cope with non-linearity in outcome determination – in other words with non-proportionate changes which might be described in terms of catastrophic transformations but are more easily understood as changes not of degree, but of kind. Neural network approaches (see Garson 1998 for a full description) should not be understood as a variant of mathematical modelling in any meaningful sense. The others are modifications and developments of the linear approach in an effort to cope in various ways with non-linearity although mathematicians might dispute that description.

Tappenden et all devote a good deal of attention to 'importance analysis' (2004 31 and following). They do so because the use of any of the kind of modelling techniques described above becomes difficult when there are large numbers of parameters to be incorporated in the model. So it is necessary to have an approach which can 'partial out' the contributions of the parameters and determine which actually matter in determining outcomes. The importance of 'importance analysis' is a very clear indication of the significance attached to variables as real causes in approaches of this kind. There are very different ways of understanding causality and we will address these in a moment. First, having reviewed serious efforts to address complex causality through mathematical modelling, let us knock on the head the value of simplistic understanding in relation to what is increasingly being described as the 'gold standard' in applied social research – the randomized controlled trial.

The central principle of the randomized controlled trial is control by randomization – that is obvious from the name but we have to consider rather carefully what it actually means. In a bench experiment the experimenter exerts actual physical control over a sub-set of reality abstracted from general reality. For example in what in my days was O level physics we worked by keeping voltage constant to maintain a constant difference by using a fixed power battery in a circuit, varied resistance by using a rheostat, and plotted out current flowing using an ammeter. We then plotted current against resistance and found a downward sloping straight line so that current was inversely proportionate to resistance given constant difference. This demonstrates a simple causal relationship – current flowing depends on the resistance in the circuit in a direct fashion. However, this is, as Znaniecki noted in the 1930s, very different from the procedures of control involved in any statistical experiment where the control derives from the random allocation of cases to treatment and non-treatment. Certainly the understanding of causation remains simple. We consider the treatment to be the cause of the observed effect. However, even here there is a problem because the outcomes vary across the cases. All electrical circuits in this universe will display the simple relationship between current and resistance described above. In an RCT not all cases will display the same, or indeed even any or in the same direction, effect given a supposedly simple causal intervention. This is the case even when we are dealing with something where we might postulate simple causation e.g. a pharmaceutical intervention in relation to a medical condition. Even here we will see variation even if average out across the cases there is a significant difference between the cases in the intervention group and the cases in the control group. So even here what Pawson and Tilly (1997) describe in their equation thus:

Mechanism & Context => Outcome

holds. Mechanism in interaction (Byrne 2002's modification of Pawson and Tilly) holds. In other words the individual physiology of cases modifies the impact of the drug with differing results in different cases even if there is really a simple causal relationship between the drug intervention and the disease process.

However, when we have multiple and complex causation – i.e. when lots of things operating together generate causes in complex ways and different things operating together might generate the same outcome, then RCTs are wholly useless. Lots of things operating together in context is another and better way of saying that any model that fits the data must allow for interaction terms including interaction terms of a high order. But that still leaves us stuck with the model, the unique and specific model that fits the data. Of course with any real data set lots of models will actually fit the data. We used to insist on parsimony, on Ockham's razor, and go with the simplest model but now people usually use non-quantitative theory to justify the model selected to explain a pattern of outcome data. What we really need are procedures which allow us to see that multiple models actually fit the data – that there is more than one way to come to an outcome, more than one way to skin a cat.

Lots of things have been said so far, not for the first time it has to be admitted! Let us spell them out in bullet point form so as to make them clear:

- Significant changes are categorical – changes of kind not of degree. This implies that the level of measurement which should be normal in social

research should be one which specifies the kind of thing something is – i.e. nominal – specifying the kind of thing something is.
- Social Causation is usually complex.
- Social Causation can be and often is multiple.

So what kind of quantitative methods should we be employing? I am going to propose two approaches which can be deployed in sequence. The first is 'sorting things out' (See Bowker and Starr 1999) i.e. classification using numerical typology approaches. Cluster analysis procedures, readily available in most statistical packages including SPSS, are ways of using measurements describing the 'variate traces' of cases (see Byrne 2002 for what that term means) to sort them into kinds based on the maximization of discrimination among the kinds generated. Of course the kinds are fuzzy sets and there are issues in relation to the absence of any kind of inferential basis for clusters generated from sample data. However, fuzziness is real, there are ways around the inferential problem (for example randomly partitioning the sample and reclustering to check robustness), and in any event we often now have information about all the cases of interest and are not dealing with samples at all.

Only two articles in JSP used cluster analysis techniques, primarily in relation to arguments about macro scale classification and debates with Esping Andersen. None used the approached pioneered by Dyer (2006) which is to explore the character of change in systems, i.e. to map out system trajectories, through constructing time ordered clusterings and describing movement of cases among these typologies over time. In fact such is the dominance of variable based approaches that authors regularly carry out latent class analysis using variables as a rather clumsy way of trying to map such trajectories. Time ordered clustering does this simply and elegantly. Clustering describes and we can see, following Dyer (2006) how it does this not just for cases considered statically i.e. described at one point in time, but also for cases considered in a dynamic fashion. We can describe dynamism in terms of sequenced sets of changes of kind.

This is description. How do we deal with cause? Following Byrne and Williamson and Dyer (both in Byrne and Ragin2009) I suggest that we use Qualitative Comparative Analysis (see Ragin 2000). Essentially QCA opens up the possibility of allowing for multiple causes by establishing configurations – sets of related causal factors in a truth table which are associated with an outcome. In other words QCA specifically allows for complex and multiple causation. Moreover it does so in terms of 'set theoretic relationships'

> … cases are viewed as configurations – as combinations of characteristics. Comparison in the qualitative tradition thus involves comparing configurations. This holism contradicts the radically analytic approach of most quantitative work. (Ragin 1987 3)

For a developed discussion of all this see Byrne and Ragin (2009). The point is that there is another way to go which is isomorphic with complex social reality and is applicable at all scales – macro, meso and micro. I will conclude with that assertion.

Abbott, A. 2001 *Time Matters* Chicago: University of Chicago Press

Becker, H. and Ragin, C. 1992 *What is a Case?*Cambridge: Cambridge University Press

Bowker, G.C. and Star, S.C. 1999 *Sorting Things Out* Cambridge Mass: M.I.T. Press

Boorah, VK, 2000 'Is There a Penalty to Being a Catholic in Northern Ireland? **...** *Journal of Social Policy*, 29:2, pp. 281-301

Byrne, D. 2002 *Interpreting Quantitative Data* London: Sage

Byrne, D. and Ragin, C. 2009 *Sage Handbook of Case Based Methods* London: Sage

Cilliers, P. 2001 'Boundaries, Hierarchies and Networks in Complex Systems' *International Journal of Innovation Management* 5 2 135-147

Dannermark, B., Ekström, M., Jakobsen, L., and Karlson, J.C. 2002 *Explaining Society* London: Routledge

Dyer, W. 2006 The Psychiatric and Criminal Careers of Mentally Disordered Offenders Referred to a Custody Diversion Team in the United Kingdom. *International Journal of Forensic Mental Health*. Vol. 5(1) pp.15-28.

Pawson, R. and Tilley, N. 1997 *Realistic Evaluation* London: Sage

Ragin, C. 1987 *The Comparative Method* Berkley: University of California Press

Ragin, C. 2000 *Fuzzy Set Social Science* Chicago: University of Chicago Press

Rihoux, B. and Ragin, C. 2004 'Qualitative Comparative Analysis (QCA): State of the Art and Prospects', paper presented at : APSA 2004 Annual Meeting, Panel 47-9, Chicago

Ormerod, P. 1994 *The Death of Economics* London: Faber

Payne, G. Williams, M. and Chamberlain, S. 2004 'Methodological Pluralism in British Sociology' *Sociology* 38 1.

Tappenden, P., Chilcott, J., Egginton, S., Oakley, J. and McCabe, C. 2004 *Methods for expected value of information analysis in complex health economic models: developments on the health economics of beta interferon and glatiramer acetate for multiple sclerosis*. Technical Report. Gray Publishing, Tunbridge Wells.