I. Introduction

In one sense a discussion about theories and methods in comparative social policy is 'about everything'. As Joan Higgins has put it "acts of comparing are part of our daily lives". For example, parents compare their babies "to see whether they are unusually fat or thin, or small or large" (Higgins, 1981, p.7) and the photographer places a figure in the foreground of a mountain.

In another sense, it may be said to be 'about nothing'. Many of the issues surrounding the theories and methods in comparative work are not exclusive to cross-national studies, or to social policy. There is no distinct social science 'cross-national method' although such research highlights some of the issues in making scientific, as opposed to impressionistic, comparisons. For example, while access to a wide and more varied range of data than is provided by single country material is one of the main advantages of cross national work, the non-equivalence of the data can be a major issue. Cross-national research also underlines the problems involved generally in developing research methodologies and in conducting research which has explanatory power or moves off in new directions.

We may begin by trying to understand some of the issues by reference to the natural
sciences. A ‘comparative method’ is used in the natural sciences where, for example, two different substances are injected into the same constant tissue to compare their effects. In the social sciences similar experiments are attempted when the results of a social intervention (e.g. social work) are evaluated by comparing the condition of two groups, where one has been subject to the intervention and the other, an identical ‘control’ group, has not.

Although there may well be problems in the natural science of how to measure the results of the experiment, it is clear that the social science experiment is less ‘rigorous’. The effect of the intervention cannot be isolated from other factors which may have affected the clients, the controls can rarely be matched exactly, and, indeed it may be difficult to know how many, and which, variants they need to be matched for. Nonetheless, there is a well-developed evaluative tradition in comparative social policy (see Bolderson, 1988). We discuss some examples in section II.

There are also other ways in which comparison in the social sciences is used as a substitute for experimentation in the natural sciences. Scientists use experiments to advance the development of a theory. For example, a physicist may write down a model, in the form of equations, to explain a set of observations, and then conduct experiments to test whether the implications of the model are valid. This methodology has an affinity to the way in which nineteenth and early twentieth century sociologists used the comparative method. They tested theories about the development of societies (e.g. Comte, Spencer); formulated universal social laws (as in Durkheim's work on the causes of suicide); created typologies and models of society (e.g. Hobhouse); explained particular configurations (as in Weber's account of the rise of Western capitalism) (see Fletcher, 1972). The descendants of this tradition in comparative social policy include Wilensky (1975), who was explicitly concerned with
examining the industrialisation theory of welfare development, and Esping-Andersen (1990) who correlated three models of the political economy of capitalist countries with different welfare configurations. Their work is discussed in sections III-V.

Other distinguished contributors to comparative social policy have moved away from comparison as a `controlled experiment'. Historians in particular are prone to argue that it is not possible to standardise or control sufficiently to test models across countries. Furthermore, an emphasis on `testing' imposes competition on theories which can be made to be complementary. Baldwin deals with the battle between socio-economic and political explanations of the welfare state in a single sentence: "[s]ocial issues play an important role here, but only as filtered through the parliamentary membrane" (1990, p.54). Clearly unimpressed by "legions of empirical studies identifying a link between the labor movement and the expansion of the welfare state" (p.42), Baldwin opts instead to examine several national cases over a long span of time. In historical case studies, theory is the servant rather than the master, generating explanations and guiding the search for data.

A feature of this work is its willingness to grapple with the historical and institutional specifics of the countries under study. This is characteristic of the classic studies of the policy process in a comparative perspective. Heclo (1974) used inductive and interpretative methods to understand the interaction of party politics, interest groups, bureaucrats and socio-economic factors in the formation of social policy. Notwithstanding its attention to detail, Heclo's study draws out theoretical implications and illuminates patterns of development in public policy-making. Historical case studies often also aim for commonalities in their explanations, while acknowledging idiosyncratic features. For example, Baldwin ranges over time and country to develop the thesis that welfare is the product of the interests which
different social groups have in arrangements which share risks to their advantage.

One methodological issue presented by the use of case studies is that it is hard to identify a single method or system in the gathering and organisation of data. The most distinguished contributions in the field seem to suggest that years of immersion are needed to write a good case study. In section VI we turn to examine the methodology adopted by researchers who have not had the luxury of immersion, but who also wish to avoid imposing a particular theoretical approach. We give some examples of designs of cross-national research projects which do not use the standardising methodology which is characteristic of the classic comparative method.

The material which follows is, by necessity, highly selective. Many distinguished contributions to comparative social policy are not mentioned, and some do not 'fit' in any of the methodological groupings which we have identified. Furthermore, our selection of examples is biased towards comparative studies in social security. This is partly because social security is our own area of specialism, but it also reflects the early development of standardising methodologies for analysing social security. Comparative studies of other areas of social policy, such as health care systems, have developed apace in recent years (see, e.g. Wall, 1996), and this has highlighted new issues in comparison and contributed to the development of new methodological approaches.
II. Evaluative studies

We begin by reviewing some studies which can strictly be said to be comparing social policies, as opposed to comparing welfare systems or comparing whole social configurations. This means that they focus on a narrowly-defined set of interventions, such as cash benefits for various contingencies. A grounding in social policy is reflected in the care taken to describe the allocation system (conditions of eligibility and entitlement), and in an evaluative underpinning concerned with the effectiveness of the government in achieving an allocation of resources which meets needs. This means that researchers are usually interested in, for example, the adequacy of provision, the coverage of the population, and the equity or otherwise of treatment of different groups in the population.

For example, in their comparative study of child support in fifteen countries, Bradshaw et al set themselves two objectives: "to quantify the value of the total [child support] package and assess its contribution to meeting the needs of a variety of family types at different income levels" (Bradshaw et al, 1993, p.256). In addressing the first objective, the researchers emphasised the importance of developing a comprehensive and consistent measure of the child support package, including cash benefits, tax allowances, housing benefits and the provision of health care, schooling and pre-school facilities. As we discuss further in section III, many researchers have used aggregate statistical data on government social expenditure to measure social policy activity or 'welfare effort'. Statistical services, such as Eurostat, even provide breakdowns of welfare expenditure by function which would seem to allow researchers to study welfare effort in particular social policy areas. However, Bradshaw et al rejected the possibility of using aggregate data on expenditure on family policy to compare child support across countries. Not only do the available
data have some obvious omissions (e.g. the value of tax allowances is not included), but also the high level of aggregation means that it is not possible to compare outputs across different households or to detect differences between countries in the composition of interventions within each functional area.

Aggregate expenditure data were, therefore, inadequate for Bradshaw et al's evaluative task, which required information on social policy measures as they affected a variety of family types in different economic circumstances. The researchers therefore opted for the `profile' or `model families' method, whereby they calculated entitlements to the child support package for a range of different households. This enabled them to draw comparisons which would shed light on issues of concern in UK social policy, such as the extent of `targeting' towards low income families, the treatment of lone parents, and the implicit equivalence scales in the assistance offered to families of different compositions.

To some extent, the data on the child support package for different family profiles enabled policies to be evaluated against certain axiomatic concerns or values (effect on incentives, uniformity of treatment, responsiveness to different needs). However, Castles and Mitchell (1992; see also Mitchell, 1991), have argued strongly and influentially that evaluation should focus on the effectiveness of social policy in achieving the ends of reducing poverty and redistributing income. Castles and Mitchell also showed that there was no simple linear relationship between welfare expenditure and redistributive impact. By the same token, Bradshaw et al could not assume that those countries with the most generous child support packages would also have the lowest prevalence of child poverty. Instead, they had to turn to another data source for information on child poverty, and then examine the correlation between 'package size' and poverty outcomes.
Both Castles and Mitchell and Bradshaw et al used the Luxembourg Income Study (LIS) to provide comparative data about poverty and redistribution. LIS is a data bank where household income and expenditure surveys from different countries are held and developed to facilitate comparative analysis of the distribution of income (Smeeding et al, 1990). The difficulties of comparing income distributions cross-nationally, and the contribution made by LIS, are discussed by Atkinson et al (1994, 1995). Among the issues to emerge, it is shown that a number of countries have crossing Lorenz curves, which means that they cannot be ranked as having more or less inequality without taking a view about the relative importance of inequality at different points on the income scale. Focusing on the position of the poorest groups leads to the use of inequality-based measures of poverty (e.g. percent of population below half average income). While this focus reduces the complexity of comparison, rankings of poverty rates are also sensitive to the exact line chosen, the equivalence scales used, and other details.

The development of LIS gave a major boost to research in what Uusitalo (1985) has called the `redistribution paradigm'. This approach to the analysis of social policy is based on a conceptualisation of household income formation where the market provides the foundation (`primary' income distribution) and different instruments of state social policy modify the primary distribution en route to generating a final income distribution (Diagram 1).
Diagram 1

| Wages and salaries, income from self-employment, rents and interest = FACTOR / PRIMARY INCOME |
| + cash benefits = GROSS INCOME |
| - direct taxes and social security contributions = DISPOSABLE INCOME |
| - indirect taxes, benefits in kind (health, education, housing services etc) = FINAL INCOME |

(Adapted from Uusitalo, 1985, Figure 1)

LIS provides data on each stage in this welfare-generating process up to row 3, and some data on row 4. This means that it is possible to calculate income-based poverty rates for different countries and to compare the role of the primary income distribution, cash benefits and income taxes in generating (and ameliorating) poverty. Mitchell (1991) used LIS to evaluate the ‘targeting efficiency’ and ‘poverty reduction efficiency’ of ten countries' social security systems. She established that the Australian social security system had a significantly higher level of poverty reduction efficiency than the other nine countries, reflecting the high use of income-testing in the Australian system and the setting of benefits at a level which took families up to the poverty line but not significantly above it. Other findings were more unexpected. It was notable that social insurance-based systems showed high targeting efficiency (most money went to the pre-transfer poor). This was linked to a general pattern whereby countries with low post-transfer poverty rates tended to have high pre-transfer poverty rates. This pattern is not as paradoxical as it sounds. It may reflect ‘second order’ effects of the transfer system on the primary income distribution. For example, where there is a generous state pension, people rely on it and would be in poverty without it (pre-transfer), whereas where the state pension is minimal, there is more private provision, which lowers pre-transfer poverty.
What insights can the analysis of income and expenditure datasets give to comparative social policy? Two limitations are widely acknowledged. First, the datasets do not provide any account or explanation of policy. As O'Higgins et al note: "[LIS] creates a demand for additional historical, institutional and legal knowledge about the countries being studied" (1990, p.159). They speculate that this demand might be met, at least in part, by building into the database descriptions of the different countries' social programmes (p.169). A related suggestion from Atkinson was that LIS needed a complementary policy database, which would allow the income-generating process to be modelled explicitly (1990, p.xxiii). While research along these lines has proceeded, it is evident that significant areas of social policy cannot be `parameterised'. For example, benefit entitlement depends not only on income, family circumstances and age (variables routinely captured in household surveys) but also on administrative judgments about disability, other special needs or unemployment.

This problem is related to another acknowledged limitation of the redistribution paradigm: its very strong cash income orientation. O'Higgins et al note that data on noncash income, especially the value of public services (row 4 in Diagram 1), can significantly alter the cash income picture of inequality and poverty (1990, p.165). However, it is not clear how data on access to services should be interpreted. Where the data have been assembled, they have tended to suggest that services do not have a strong redistributive impact between income classes (Le Grand, 1982). However, Ringen (1987) has argued that the essential goal of service provision is not income redistribution, but the promotion of equal access in relation to needs. Success in achieving the latter aim cannot be verified from household income and expenditure statistics: the arrangement of services to meet needs is closely connected with, and indeed often dependent on, the assessment of needs by expert professionals.
Looking at how the use of LIS datasets has developed in the 1980s and 1990s, we would suggest that the contribution to policy analysis has been less than commentators expected at its foundation. This is partly because of the limited development of complementary policy information, and partly because the `redistribution paradigm' does not capture all the aims of social policy (we return to the multiple aims and objectives of social policy in section IV below). Instead, LIS has proved to be a rich source for developing a research agenda around the idea of a `welfare society', which goes beyond public sector welfare measures to examine the interactions of market, family and state in the generation of welfare (Higgins, 1981; Rein and Rainwater, 1981, 1986; Shalev, 1996).

Comparative evaluative work in social policy has contributed greatly to our understanding of welfare effort, both in showing how welfare effort can be made up of a range of interventions and in drawing attention to the complex relationship between effort and outcomes. In the discussion of regime theory in the next section, we show how these insights have informed `grand theorising' about the welfare state. First, however, we take a step back in the chronological development of comparative analysis. There is an important body of work which assumed that welfare effort was unproblematic to measure and unidimensional in nature, and which sought to explain varying levels of welfare effort across countries by reference to macro social, economic and political factors, and to it we now turn.

III. Common factors explanations and models

The studies discussed in this section are concerned with developing and testing hypotheses about the development of the welfare state. By contrast with the research
discussed in section II, these studies tend to treat the welfare system as a whole rather than examining particular areas of social policy. Furthermore, they are not evaluative in intention, in that they are not attempting to examine how effectively needs are met by particular policy designs. One consequence is that the research is often careless of the distinction between inputs, outputs and outcomes, tending to view all three as aspects of the welfare 'configuration'.

This section discusses how theories of welfare effort have been developed and tested using regression analysis. Section IV turns to regime theory, which is not susceptible to a linear regression approach, instead requiring the researcher to identify and analyse the interactions within 'clusters'. A leading exponent in the first area of research was Wilensky; in the second Esping-Andersen is a key figure.

Both Wilensky and Esping-Andersen draw on 'grand' theories of political and economic development to derive hypotheses about the welfare state. In Wilensky's case, key ideas came from theorists of 'the logic of industrialism', 'the end of ideology' and 'convergence' (e.g. Kerr, 1962; Bell, 1960). All these theories stressed increasing similarities between societies as a result of industrialisation, and underplayed the significance of political differences. Part of Wilensky's achievement was to link these theories to a separate empirical tradition which had tested the role of economic development in determining social security benefit levels (see the discussion of studies by Aaron, 1967; Cutright, 1965 and Pryor, 1968, in Wilensky, 1975, pp.16-18).

As noted above, the use of regression analysis is one of the main methodological features of this area of research. Regression analysis imposes a strong methodological structure on the research. First, the researchers must choose a 'dependent variable' which is what they are seeking to explain. The theories under discussion here are
about `welfare effort', but this is an abstract concept which has to be 'operationalised' or measured if the theory is to be tested empirically. Finding `empirical referents' to correspond to theoretical concepts is an important process which can reveal limitations in the theories. Two measures of welfare effort have been widely used: transfer payments alone, and government expenditure on welfare service provision plus transfer payments. Only relatively recently have these differences in measurement been bestowed with a theoretical interpretation of how different factors might explain the development of a `transfer state' and a `service state' (Huber et al, 1993).

The next step is to undertake a theoretically-informed search for `independent variables' which might explain the level of the dependent variable across countries and through time. Theory provides some clues about where to look for explanations, but, again, empirical referents have to be found to correspond to theoretical concepts. A good example is the detailed discussion of how to operationalise `degree of state centralisation' in Huber et al (1993). They note that most measures of this concept "appear to have been dictated by data availability, and the fit between available measures and the concepts proposed in the comparative historical and theoretical literatures is not very good" (p.720). These problems partly reflect the nature of the data requirements imposed by regression analysis. For each independent variable, consistent observations must be obtained for all the countries and time periods to be included in the regression.

Once the dependent and independent variables have been chosen, regression techniques generate three main statistical findings:

(a) coefficients which measure the relationship of the independent variables to the dependent variable;

(b) measures of significance, indicating how important each of the independent
variables are in explaining the dependent variable;

(c) measures of fit, indicating how much of the variance in the dependent variable is explained by the independent variables.

Note that a single coefficient (and measure of significance) is generated for all the observations of each variable included (thus the terminology `common factors explanations'). This means that the analysis requires that the independent variable has the same effect on the dependent variable across countries (in cross-section analysis), through time (in time-series analysis) or both (in the analysis of pooled data). It is possible to include `interaction effects' in regression analysis, whereby variable A is expected to work one way in the presence of variable B and another way in the absence of B, but only a limited number of interactions can be allowed before the analysis runs into statistical problems such as inadequate degrees of freedom (too many explanatory variables relative to the number of observations to be explained).

In principle, regression analyses can adjudicate between different theoretical accounts of the same phenomenon. Wilensky (1975) sought to explain differences in welfare effort across countries and through time using indicators of economic development. Other researchers have stressed the importance of political factors (Castles and McKinley, 1979; Castles, 1982; Alber, 1981; Korpi, 1980). A thriving research programme has developed as these alternative accounts have been tested and new explanations added (e.g. the role of religion (Wilensky et al, 1985; Castles, 1994) and the structure of the state (Huber et al, 1993)). Wilensky himself has developed a further model which combines the influence of industrialisation (convergence) and politics (divergence) and includes the role of democratic corporatism and the power of mass-based Catholic and 'Left' political parties (Wilensky, 1990).
However, it has to be admitted that, on the whole, findings from large-scale regression analysis have failed to resolve theoretical debates. There is no agreed view on the role of social democracy, the effect of policy maturity, economic growth or population aging on the level of welfare effort. Partly this is because of a shortage of good empirical referents for some of the theoretical concepts; this contributes to a situation where "too many theoretical propositions are being represented by too few measures" (Amenta, 1993, p.757). Partly the problem is that country-specific factors ensure that no regression of common factors can achieve a good fit. Finally, there are problems with taking a unidimensional view of welfare effort, instead of looking at the whole configuration of social policy.

What, in the end, has been the contribution of the comparative study of the determinants of welfare effort to our understanding of the welfare state? Underpinning the debate between different researchers over the choice of explanatory variables was a fundamental issue about `convergence'. Wilensky put forward the hypothesis that "economic growth makes countries with contrasting cultural and political traditions more alike in their strategy for constructing the floor below which no-one sinks" (1975, p.27). Economic development was given a central place; political choices and preferences were relegated to the sidelines.

While many researchers who disputed the convergence theory tackled Wilensky on his own ground by refining the regression analysis of welfare effort, it is also evident that some of the issues presented by convergence theory were methodological. One problem concerned the countries included in the dataset. Convergence theory was most strongly supported when the data were drawn from countries across a wide range of levels of economic development. Differences among the developed countries were not explained. Regression analysis was less suited to analysing a developed
country subsample because of the smaller number of cases. Regression analysis also needs a well-defined dependent variable, but differences among developed countries concerned the nature as well as the quantity of welfare effort. The multidimensional analysis of welfare states required different methodological approaches.

IV. Regime theory and the multiple objectives of social policy

Regime theory extends and enhances common factor analysis in two main directions. First, the idea that there are different types of welfare state replaces the emphasis on ranking welfare effort. Second, regime theory looks at whole configurations of policies, emphasising the conjunctions and interactions between social policy and "its reciprocal political and economic institutions. .. [R]egimes are the specific institutional arrangements adopted by societies in the pursuit of work and welfare. A given organisation of state-economy relations is associated with a particular social policy logic." (Esping-Andersen, 1987, pp.6-7). Whereas regression analysis requires that the relationship between variables obeys a common structure across countries, regime theory allows for different patterns of relationships.

To understand how Esping-Andersen's *Three Worlds of Welfare Capitalism* (1990) has come to be one of the most-cited books in comparative social policy, it is useful to distinguish between these two aspects of the work: the multi-dimensional characterisation of welfare effort and the interaction between politics, economic conditions and social policy. The discussion of Esping-Andersen in the social policy literature has centred on the multi-dimensional characterisation of welfare effort. The data for this work are multi-country information on the detail of social policy interventions (e.g. eligibility and entitlement conditions for benefits, etc). Using such
data, countries are grouped according to certain generic similarities in their social policies.

The data for the analysis of interactions between political systems, economic structures and social policies are of a different order. We would argue that the analysis of regimes in this sense has proceeded most fruitfully through case studies. That this is not how regime theory is usually understood in social policy reflects the particular methodology employed in *Three Worlds*, but this methodology is not representative of the `reciprocal' analysis of the political, economic and social policy dimensions of regimes generally. We leave the discussion of `reciprocal' regime theory to the next section, and concentrate here on the multi-dimensional characterisation of welfare states.

Esping-Andersen focuses on two key dimensions:

1. Decommodification: the degree to which social policy makes individuals independent of the market;
2. Stratification: the degree to which the welfare state differentiates between different groups (e.g. according to occupational status). The opposite of stratification is solidarity, whereby the welfare state offers the same benefits and services to the whole population.

These concepts originate in political theory. Decommodification is a social democratic political strategy which, Esping-Andersen hypothesised, would be reflected in the design of social policy interventions. Stratification is a conservative strategy which one might find signs of in both state-corporatist and liberal regimes, although a liberal regime might leave stratification to the market.
As discussed above, an important step in developing an analysis like this is to find empirical referents for the concepts. For stratification Esping-Andersen produced a range of indicators (1990, Table 3.1) and refrained from combining them into a single index. For decommodification, however, Esping-Andersen constructed a composite indicator which combined average and minimum benefit levels (positive for decommodification), contribution periods, waiting days, individual pension finance share (negative) and coverage (positive). While coverage can be measured by comparing the size of the apparently eligible group with the number receiving benefits, this measure was not used where benefits were means-tested, and instead an arbitrary weight was assigned. An ad hoc adjustment was also made to the contribution period variable in the case of means-tested programmes.

The decommodification index combines two things: the generosity of benefits and the conditionality attached to their award. While one aspect of conditionality - that created by requiring prior insurance contributions - can be measured readily, other aspects of conditionality to do with intrusive and rigorous administration are harder to measure. Esping-Andersen sees means-tested benefits as "highly conditional in terms of offering rights" (1990, p.54), but this is assumed, not revealed by the data. The example of the decommodification index shows how difficult is the search for measures and indicators which can be applied across countries in a uniform fashion. Coverage `worked' as an indicator of conditionality for some systems, but the coverage figures for Australia and New Zealand were too high to support what Esping-Andersen wanted to argue about the relationship between means-testing and decommodification, necessitating an ad hoc adjustment to those countries' decommodification scores.

Other comparative researchers in the `modelling' tradition have suggested that it is
possible to construct a multi-dimensional picture of the welfare state without recourse to such elaborate statistical manipulation. Castles and Mitchell (1992) show how four 'worlds' can be located with a simple two-dimensional structure obtained by cross-tabulating welfare expenditure and benefit equality (difference between minimum and average benefits). More recently, Bonoli (1997) has offered a two-dimensional classification with welfare expenditure and contribution financing as the key variables. In both these examples, the old measure of welfare effort has reappeared, albeit as just one dimension within the analysis. One can recognise a familiar basic point in both these classifications: aggregate expenditure is not linearly related to redistributive impact because of the presence of insurance-based benefits. Castles and Mitchell detect the presence of insurance through benefit (in)equality; Bonoli through the share of contribution financing.

Titmuss (1974) made an early contribution to the multi-dimensional analysis of welfare effort by identifying three conceptions of the role of social policy and linking each conception to principles of entitlement to benefits and services. Titmuss argued that earnings-related contributory systems reflected an 'industrial achievement-performance' conception, universalism an institutional conception and means-testing a residual conception of social policy. The purpose of Titmuss's classification is to explain the logic of advocating a particular basis for entitlement (means-testing, etc) by connecting it to a wider view of the role of social policy. While Titmuss noted that some countries (or, more precisely, some national traditions) appeared to give more weight to one principle than another, he did not attempt to explain each conception with a theory about there being different types of welfare state. Indeed, because policies are introduced under particular sets of conditions at different historical moments for a wide variety of target groups, elements of all three bases for entitlement are likely to coexist in a country's social policy.
In Bolderson and Mabbett (1995), we suggested that different rationales for social security payments could be analysed systematically to enable different countries' social security systems to be located within a common framework. We argued that "broad brush" characterisations of systems did not survive confrontation with the complex detail of actual arrangements. For example, adherence to insurance principles varies widely across different contributory systems. We found that "[r]ules governing benefits in insurance systems can be very complex, and inside the web of rules lie important differences between systems in the extent to which strict relationships between contributions and benefits or earnings and benefits are maintained" (1995, p.124).

We presented an analytical framework from which we generated a set of questions which could be used to locate different countries' social security systems along different axes or dimensions. Developing the questions revealed how some widely-used social policy concepts had system-specific meanings which made it difficult to apply them to comparative analysis. For example, 'residualism' is associated with close matching of assistance to the gap between needs and means, and means-testing is often taken as the central indicator of residualism. However, this ignores the 'needs-testing' side of the allocative system. Furthermore, adjustment of assistance to needs and means can be done by a variety of administrative procedures ranging from close enquiries into personal circumstances to use of a few easily-verified indicators. While the former approach to administration harks back to the Poor Law, the latter more resembles a modern taxation system.

There are similar problems with 'universalism' and the closely-related concept of social citizenship. Looking at social security, we find that benefits often described as
universal, such as child benefit, are more accurately categorical (payable to those in the category 'caring for children'). Universality is often defined by what it is not: not means-tested and not contributory. Yet this can encompass a multitude of benefits using categories which entail widely varying levels of selectivity. The ' universality' of child benefits or old age pensions reflects their straightforward eligibility conditions (age, and little else). There are benefits payable to reflect additional needs arising from disability which are neither means-tested nor contributory, but they are not always envisaged as 'universal' because proving eligibility (membership of the category 'disabled') is often an onerous process.

In the end, we would have to say that the comprehensive multi-dimensional classification of social policy interventions in Bolderson and Mabbett (1995) ended up being too complex to be useful. The problem with deconstructing 'broad brush' characterisations is that the results militate against clear and all-encompassing cross-country contrasts. The study was more revealing about the range of objectives, rationales and instruments which characterise social policy than it was about the differences between countries.

V. Comparison and Convergence

In the discussion at the start of section IV, we noted that regime theory was not just about the multi-dimensionality of welfare effort. Regimes can be characterised in terms of the reciprocal relations between political institutions, economic structure and social policy. Esping-Andersen analyses these relations in Part II of The Three Worlds of Welfare Capitalism by "chang[ing] gears, methodologically speaking" (1991, pp.142-3). Instead of developing comparative indicators for a number of countries, Esping-Andersen confines his discussion to three countries, selected as 'exemplars' of the
different regime types. He argues that the three regime types exhibit distinct configurations of macroeconomic policy, wage bargaining institutions, labour force participation, welfare sector employment and transfer payments.

One important implication of regime theory is that different types of welfare state can coexist in the global economy. Each regime type is affected by, and responds to, international competitive pressure in a distinctive way. This idea is captured in the subtitle of Esping-Andersen (1996): `National Adaptations in Global Economies'. The 1996 study is part of a burgeoning comparative welfare state literature which consists of country case studies organised around the theme of globalisation, often more or less explicitly addressing the question: `can the welfare state compete?' (see, e.g., Pfaller et al, 1991; Cochrane and Clarke, 1993).

Wilensky's argument that welfare states would `converge' in the course of economic development and industrialisation provoked a vigorous response from researchers who argued that `politics mattered' (e.g. Stephens, 1979; Korpi, 1983). One can see similar divisions between convergence theorists and those who attach importance to `national adaptations' in the debate over the impact of globalisation. The new convergence theory is that all market economies are subject to competitive constraints on state welfare activity. Extensive welfare states cannot compete with those which make minimal, residual provision in an increasingly integrated global economy. Competitive pressure on the ability of the state to supply welfare is accompanied by intensified demand, arising from the aging of the population. This convergence theory is accompanied by convergent policy advice, with international organisations advocating a common set of solutions to these global pressures, involving improved efficiency in public sector provision, curtailment and `targeting' of cash transfers, and privatisation.
A methodological gulf can be observed between proponents of the view that globalisation will enforce (downward) convergence on welfare states, and those who argue that distinctive national solutions remain possible. The argument for convergence rests on logical deduction rather than observation. The central argument is that high welfare expenditure leads to higher labour costs which result in lower profitability. Since mobile capital will migrate to high profit areas, it follows that capital will migrate away from redistributive welfare states, eventually causing an economic crisis unless there is reform. While it is possible to find illustrations which seem to support this argument (such as the recent economic crisis in Sweden), the theory does not rely on instantiation for its validity. There has not been a uniform process of welfare state retrenchment, but adherants of the globalisation argument can claim that this is because of (ultimately doomed) 'holding operations' rather than reflecting a capacity to adapt. The logical-deductive structure of globalisation theory also means that it is inherently 'universal' (applicable in any setting) rather than 'comparative' in the sense of gaining insights by highlighting differences between countries.

By contrast, the 'national adaptation' arguments are based primarily on 'instances' (country studies) but are relatively weak on theoretical insights. Arguably, this is because they are not really comparative enough (Pierson, 1995, p.202). The theory that responses to international competitive pressure vary by regime type seems to have some validity (Rhodes, 1996), but the difficulties of classifying countries into regimes and the variety of reform paths taken by different countries mean that this theory works only at the most general level.
VI. Uniqueness and Generalisation in Case Studies

In the context of comparative social policy, the characteristic feature of a case study approach is that the research examines the specific institutional, historical and political features of each country covered, instead of imposing a `standardising' framework whereby only pre-selected items of data are accepted for incorporation into the analysis. Furthermore, case studies allow an open view to be taken about causal linkages, with historical analysis often used to examine how events unfolded.

An early and deliberate step away from standardising methodologies and mono-causal frameworks and explanations was taken by Heclo (1974) in his seminal study of the policy processes involved in the development of income maintenance provisions in Sweden and the UK. In "tracking the intractable" (p.10) he rejected "the analysis of general correlations among aggregate variables" in favour of "inductively building up generalisations from detailed if somewhat less tidy accounts" (p.12). The data for his case-study were documentary and conversational, drawing on some original material, but also that from other scholars.

As we noted in the Introduction, the relationship between theory and observation is more interactive in this type of study than in those that use a `standardising' methodology and attempt to test hypotheses. Heclo related his data to existing theories about the influence of political parties, interest groups and bureaucracies, drawing out the important part played by the "middle men at the inter-faces of various groups [who had] transcendable group commitments, in but not always of their host body" (p.308). In developing this theme, Heclo was generalising from observation, i.e. developing theory inductively. However, he also sought to ensure that this theorising was backed by a rationale or interpretation, stressing that "the opposite of being
definitive is not...to be arbitrary...While it is no doubt unrealistic to expect clear-cut proofs, it is nevertheless possible to produce reasons and adduce evidence for thinking that some relationships are more likely than others" (p.16).

Heclo showed, *inter alia*, that the administrative structure of the state can exert a pronounced influence on the development of social policy. The analysis of the role of political decision-making structures, government systems and bureaucracies is the domain of the flourishing `institutionalist' school in comparative social policy. While researchers have tried to incorporate institutional variables in big comparative studies of the type discussed in section III (e.g. Huber et al, 1993), institutionalist arguments are ideally suited to the case study technique because specifics of the organisation of government are elevated to a central place in the analysis, and the argument is often historically contingent. Detailed policy studies can, nonetheless, yield generalisable theoretical insights. For example, Pierson (1994) uses studies of the USA and Britain as the basis for generalisations about the politics of welfare retrenchment, offering up not only a classification of retrenchment strategies but also some basic `ground rules' for the study of retrenchment.

Another permutation on the relationship between theoretical insight and the comparative use of case studies can be found in Baldwin (1990). Baldwin's analysis has a very strong analytical direction, whereby the welfare state is seen as a set of institutions for risk-sharing, and the research question concerns the breadth of this sharing or, in other words, the extent of solidarity. Baldwin identifies "limited, but crucial areas" of social policy and historical episodes and waves of reform to focus the selection of material (1990, p.53). Compared with Heclo (1974), Baldwin's study reads as if it was founded on a more developed set of theoretical preconceptions, although of course one cannot necessarily infer the path travelled by the researcher from the
organisation of the finished product.

If the research does not begin with a strong theoretical direction, and the researcher does not have the luxury of a long period of immersion to allow issues and themes to rise to the surface, then, we would argue, it is important to adopt a research methodology which is systematic yet 'open' in its approach to gathering comparative material. In a study of the governance and delivery of social security in five countries (Bolderson and Mabbett, 1997), we focused on contingencies (incapacity, unemployment, lone parenthood, elderly with an insufficient insurance pension) rather than specified benefits, in order to capture country-unique provisions and institutions within a comparative framework. Respondents were asked for descriptions of the delivery processes, standardised by stages (entry, review, exit), taking in all eventualities of entitlement, disentitlement, return to work, availability of additional or substitute benefits, etc. The research provided an account of the problems of benefit delivery as perceived by respondents in the organisations involved in social security in each country. It became apparent that there were certain commonalities across countries in the nature of the problems faced. It proved to be possible to develop a theoretical account of structures for the devolution of decision-making power and their 'alignment' with financial responsibility. However, this analysis provided only a partial structure for the information gathered in the research (while at the same time raising further questions which were not anticipated, and therefore not answered, in the data-gathering exercise). Other theoretical directions could also be taken; for example, it is possible to compare the evidentiary basis of benefit decision-making (Mabbett and Bolderson, forthcoming).

However, as in all the studies which try to get closer to culture-specific factors, there are problems of applying the methodology. The sources for the data included on site
visits and face-to-face interviews, and these are especially resource intensive in the cross-national context. Moreover, the research was about processes which involve discretionary decisions which are vulnerable to local and idiosyncratic interpretations. Full scale surveys, followed by in-depth interviews are difficult to conduct across several countries (see Bolderson, 1988, describing Mitton, Willmott and Willmott, 1983) and the best has to be made of a research `safari' (Hantrais and Mangen, 1996, p.4), complemented by information from other national sources.

Perhaps the greatest challenge in comparative work is to understand not only the idiosyncrasies of national conditions but also the conceptual frameworks of actors in each country. There are differences between countries in the way in which social policy interventions are interpreted and understood. Despite the internationalisation of education and the efforts of international organisations to promote common analytical frameworks, distinctive national intellectual traditions survive. Therborn (1993) has suggested that nations belong to families, defined by at least four types of connection (lineage, separated siblings, affinity groups and partnerships). Affinity groups, connected by processes of diffusion and policy borrowing, have been analysed in social policy, while the importance of lineage is well-established in the comparative law concept of `legal families'. One issue of current importance is how `contractual bonding' among states in the EU (a `partnership' family connection) will be affected by the diverse cross-cutting family ties of member states.

In a trenchant analysis of the use of the concept of `poverty' in international comparative work, Atkinson (1995) argued that comparing poverty rates by using a `supranational' definition of poverty implicitly imposes a supranational set of policy objectives and priorities. If the purpose of the analysis is to assess the comparative effectiveness of anti-poverty policies, this should be done against national objectives,
which means measuring poverty by using national rather than supranational definitions. Atkinson's discussion illustrates how standardisation of data results in the loss of information about distinctive national social policy institutions and ways of thinking. It may be frustrating to find that categories, institutions or policies such as `economic activity', `heads of households', working hours, average wages and `family benefits' are differently defined and calculated in different countries, but it is also interesting (see Hantrais and Letablier, 1996). National statistics do not differ just for administrative or other practical reasons but because they reflect different ways of thinking about the institutions they describe (Desrasières, 1996). The problem is to find methodologies which can reflect these deeper issues without losing their way in a mire of detail.

Conceptual frameworks and cultural differences affect not only policy-making but also practice in welfare institutions. An example of the fruits of recognising structural and cultural diversity among welfare professionals is the recent work of Hetherington et al (1997). The research examined the social work role in child protection in England and other European countries, following an earlier work which described the legal and administrative functioning of the French child protection system (Cooper et al, 1995).

The method involved groups of practitioners, in each of the countries, who were both the subjects of the research and the researchers. They completed questionnaires about a case-study, video-recorded discussions about it, and commented on each other's videos. The method was developed further using English and French practitioners who were paired and followed each other's cases.

The methodology recognised that "no equivalence could be assumed" between the meanings of words or the roles and functions of practitioners (Cooper et al, 1992, p.34).
However, recognition that ethnocentric assumptions would have to be suspended did not mean that the comparisons ran into sand. The researchers considered that agreements could be forged about the meanings of analytical categories such as `structure', `function' and `culture'. These were broad and presumably culturally `neutral' enough to act as comparators and to incorporate within them the embedded "bits of systems" or "ways of doing things" (p.35) which by themselves could not be compared or transplanted.

In drawing up the case-study for discussion by the practitioners, it was not possible to take a real case from either country without major adjustments for purposes of standardisation which however made it unintelligible. A fictional case had to be invented jointly by the French and English practitioners/researchers "which was neutral with respect to the systems of both countries, but recognisable" (Cooper, 1992, p.5). We draw attention to this, as it corresponds with the attempts made to trace the delivery processes of benefits in circumstances unencumbered by country-specific structures and provisions in the otherwise quite different study of the delivery of social security benefits (see above).

The approach was clearly useful as a learning exercise. It fulfilled the oft cited purpose of comparative work as `holding up a mirror to one's own society'. The practitioners from each country sought not only to understand the others' structures, values and practices, but also reflected on how they differed from their own and on how these might change in the light of new knowledge.

Where the purpose of comparative work is to `learn lessons' from other countries, (often a hoped for result when governments commission comparative work, but notoriously dangerous to attempt using more conventional methods) the interactive
method used in this study clearly signals a big step forward. Beyond this, the methodology also enabled the researchers to develop a hypothesis about the effects on practitioners of the different relationships between social work and the law in the two countries.

VII. Conclusion

Cross-national social policy-orientated studies describe, analyse, and map different countries' welfare configurations, their specific policies, or their responses to 'common' issues. They are used to test or develop theories or hypotheses, to construct models, to conduct evaluations, to show more clearly the contours of one country's arrangements, to promote "peripatetic learning" (Goodman and Peng, 1996) or the 'borrowing' of policies or practices.

This chapter has shown that a wide variety of methods are used in cross-national social policy-oriented studies. This is not surprising, since the projects undertaken differ in their scale and focus and the researchers who conduct them come from diverse academic disciplines, and use different combinations of deductive and inductive thinking. Their selection of methodological tools is likely to depend on their expertise, inclination and funding arrangements as well as on the more scientific requirement of 'fitting' the method to the objective of the research.

Despite the diversity of methods used, we can see that these different comparative studies face common problems. In the evaluative studies discussed in section II, the problem is to ensure that the research is comparing 'like with like'. In the child support example discussed, the researchers anticipated that a comparison of cash
benefits for children would lead to an incomplete and biased comparison, and they developed the concept of a `child support package' to ensure that what they were measuring across countries was equivalent. One of the limitations of early work comparing welfare states on a `macro' level (section III) was that the researchers assumed the cross-national equivalence of `welfare effort' as measured by the statistical branches of international organisations.

To make progress in explaining the development of the welfare state comparatively, it was necessary to acknowledge the national differences which lay behind aggregate measures of welfare state activity. However, recognition of difference would not in itself advance the progress of comparative research. Progress called for the development of equivalising concepts which captured differences and allowed them to be explained systematically. We have discussed how some of these concepts have more theoretical coherence than others (e.g. the limitations of the trichotomy of insurance, universal and means-tested benefits) and some are so overburdened with theoretical significance that it is a struggle to operationalise them (notably `decommodification').

Comparative analyses which use case studies approach the problem from a rather different direction. It is tempting to suggest that case studies do not utilise equivalising concepts, but this would be an oversimplification. The case study approach allows the researcher to tell a country-specific story, and thereby escapes the problem of operationalising concepts in a uniform way across countries (usually by attaching numbers to them). However, on another level, concepts are operationalised in case studies through exemplification, allowing complex theoretical ideas to be developed as the story unfolds. An attractive feature of the best examples of this genre is that they can be read on several levels. Baldwin (1990) can be raided for facts on the
countries he includes, or used to provide ideas for application in the context of other countries.

It may be that there is a lifecycle in the development of concepts and measurements in comparative social policy. ‘Birth’, or inspiration and discovery, require a certain methodological openness, which tends to dictate research techniques such as visits, use of archival sources, and other forms of ‘immersion’ which cannot be applied to more than a few countries at a time. As hypotheses are developed and explicated, it becomes possible to employ standardising methodologies, construct equivalent measurements, and test hypotheses over large numbers of countries. It is only at this stage that theories lay claim to generality or universality. With this widening of view, one might expect some refutations of the small-sample conjectures, but the reality is that ad hoc modifications are made and the theories become more and more complex and less and less accessible to proof or disproof. The research discussed in section III shows some signs of this tendency.

Many comparative researchers would challenge the idea that there is some natural or desirable progression from case studies to universal propositions supported by standardised measurement of key concepts. Amenta argues that “[c]ase studies and close comparisons of countries have been and will remain the main means for understanding the development of public spending policies” (1993, p.760). He notes particularly the weakness of quantitative research relative to historical analysis in untangling issues of causality. Advocates of small studies also argue that attempts to develop general theories for countries with a wide range of socio-economic and political structures are misguided. While nothing is certain, we feel confident in concluding that comparative social policy will continue to be a terrain contested not only by different theoretical accounts, but also by a wide range of methodological
approaches.
Further reading

Each of the main areas of comparative social policy discussed in this chapter has yielded recent additions and extensions. European integration has provided an additional stimulus to comparative work with a policy focus, and recent studies such as Eardley et al (1996) and Millar and Warman (1996) exhibit wider scope and more flexible methodologies than those reported in section II. Janoski and Hicks (1994) report on the application of new techniques to statistical methods of comparison. Van Kersbergen, a student of Esping-Anderson, brings both case study and `standardising' methodologies to bear on a study of Christian democracy and the welfare state (Social Capitalism, 1995). Esping-Anderson's `Three Worlds' also helped to provoke a flurry of analyses of the gender dimension of welfare state comparisons (Lewis, 1992; Sainsbury, 1994). Following in Heclo's footsteps are the 'historical institutionalists', including Skocpol (1995), Pierson (1995, 1996) and Steinmo, who has produced a unique study of the revenue-raising side of comparative social policy (Taxation and Democracy, 1993).

References


between Politics, Instruments and Outcomes. *Governance*, 5(1), 1-26


Cooper, A. (1992) Some Methodological and Epistemological Considerations in Cross-National Comparative Research. Paper given to Doctoral Seminar, Faculty of Educational Studies, University of Paris, Nanterre


Press.


