

<add>**Address for correspondence:** Clare Bamba, Wolfson Building, Durham University, Queens Campus, Stockton, TS17 6BH. Email: clare.bamba@durham.ac.uk</add>

<t>'Sifting the Wheat from the Chaff': A Two-dimensional Discriminant Analysis of Welfare State Regime Theory</t>

<au>Clare Brambra[qu1]</au>

<ha>Abstract<ha>

<ind>Welfare state modelling has long been an important strand within comparative social policy. However, since the publication of Esping-Andersen's 'Worlds of Welfare' typology, welfare state classification has become particularly prominent and a multitude of competing typologies and taxonomies have emerged. Each of these is based on different classification criteria, and each is trying to capture what a welfare state actually does. The result is that the literature is in a state of confusion and inertia as it is unclear which of these rival systems is currently the most accurate and should be taken forward, and which are not and should perhaps be left behind. This article extends Bonoli's two-dimensional analysis of welfare state regimes by using multivariate analysis of variance and discriminant analysis to compare and contrast the various classifications on universal criteria. It also examines the usefulness of the two-dimensional approach itself and suggests how it can be enhanced to benefit future attempts at holistic welfare state modelling. The article concludes that there are some welfare state classifications that are more useful than others, especially in terms of reflecting a two-dimensional analysis: it thereby 'sifts the wheat from the chaff' in terms of welfare state regime theory.</ind>

<ha>Keywords</ha>

Bonoli; Esping-Andersen; Multivariate; Regimes; Social expenditure; Welfare state

<ha>Introduction</ha>

Welfare state modelling has long been an important strand within comparative social policy, serving as a means of reducing the complexity of cross-national

welfare state provision (Wilensky and Lebaux 1958; Cutright 1965; Titmuss 1974; Wilensky 1975). However, it is only since the publication of Esping-Andersen's 'Worlds of Welfare' thesis in 1990 that the welfare state classification literature has become particularly prominent (Pierson 1998). Based primarily on the examination of labour market decommodification, Esping-Andersen proposed a threefold welfare state typology whereby Western countries fell into one of three regime ideal-types: Liberal, Conservative or Social Democratic. Esping-Andersen's work has provoked an extensive ongoing debate in the literature, about which principles should be used to classify welfare states (Alber 1995; Korpi and Palme 1998; Castles 1998; Abrahamson 1999; Kautto 2002; Bamba 2005a, 2005b); in which regimes particular countries belong (Ginsburg 1992; Leibfried 1992; Castles and Mitchell 1993; Ferrera 1996); the number of different regime types (Leibfried 1992; Castles and Mitchell 1993; Ferrera 1996; Bonoli 1997; Pitzurello[qu2] 1999; Bamba 2005b); the methodology of regime construction (Kangas 1994; Ragin 1994; Pitruzzello 1999; Bamba 2006); and the nature of gender stratification within different types of welfare state (Lewis 1992; Orloff 1993; Sainsbury 1994[qu3], 1999; Bamba 2004).

As a result of these criticisms a number of rival welfare state typologies have emerged, each based on different classification criteria, and each trying to capture in its own way what a welfare state actually does. Some have extended the remit of the welfare state modelling literature to include aspects of gender stratification and defamilization (Lewis 1992; Siaroff 1994; Lewis and Ostner 1995; Esping-Andersen 1999; Sainsbury 1999; Korpi 2000; Bamba 2004; Pascall and Lewis 2004), while others have examined the role of welfare state services (Bamba 2005a, 2005b). In addition, the methodological critique of Esping-Andersen's work (for an overview, see Bamba 2006) has led to the production of numerous welfare state taxonomies, many of them based on the reworking of Esping-Andersen's data, yet still suggesting alternative regimes and country classifications (Kangas 1994; Ragin 1994; Shalev 1996; Obinger and Wagschal 1998; Pitzurello 1999; Wildeboer Schut *et al.* 2001). This means that within the contemporary comparative social policy literature, even among those typologies and taxonomies that only examine income protection or the labour market aspects of welfare state provision, there are a number of contrasting claims for the existence of three (Esping-Andersen 1990, 1999; Ragin 1994; Shalev 1996; Wildeboer Schut *et al.* 2001), four (Leibfried 1992; Castles and Mitchell 1993, Kangas 1994; Ferrera 1996; Bonoli 1997; Korpi and Palme 1998), and even as many as five different types of welfare state regimes (Obinger and Wagschal 1998; Pitzurello 1999).

The result is that the welfare state modelling literature is in a state of confusion and inertia as it is unclear which of these competing systems of classification is currently the most accurate or useful and which are less so. It is the purpose of this article to establish some clarity in this regard by comparing the classifications and determining which are currently of the most utility: it will thereby 'sift the wheat from the chaff' in terms of welfare state regime theory.

<ha>Typologies and Taxonomies</ha>

Esping-Andersen's original (1990) analysis of decommodification levels in eighteen countries produced an initial threefold typology of welfare states: Liberal (Australia, Canada, Ireland, New Zealand, UK, USA), Conservative (Finland, France, Germany, Italy, Japan, Switzerland), and Social Democratic (Austria, Belgium, Denmark, Finland, Norway, Sweden).¹ In using decommodification, that is 'the extent to which individuals and families can maintain a normal and socially acceptable standard of living regardless of their market performance' (Esping-Andersen 1987: 86), his approach was an attempt to examine what a welfare state does, rather than how much money it is afforded (Esping-Andersen 1990: 2). Conversely, though, many of the criticisms directed at his typology were made on the basis that it had not in fact adequately measured what welfare states do, nor indeed managed to correctly classify many countries. This led to the construction of alternative typologies, and taxonomies, each based on slightly different measures of welfare state labour market protection and each producing slightly different results (see tables 1 and 2). Typologies are theoretically informed categorizations of welfare states while taxonomies are purely empirical. **[T1][T2]**

On the basis of the cross-classification of the same eighteen nations in terms of aggregate expenditure levels and degrees of benefit equality, Castles and Mitchell (1993) argued for the existence of a 'Radical' (Liberal subgroup) fourth regime (UK, Australia and New Zealand). Institutional analysis (Korpi and Palme 1998), cluster analysis (Kangas 1994), BOOLEAN comparative analysis (Ragin 1994), factor analysis (Shalev 1996), and principal component analysis (Wildeboer Schut *et al.* 2001) also supported the existence of a fourth regime type (see table 2). Cluster analysis by Obinger and Wagschal (1998) and Pitzurello (1999) suggested another subgroup, this time within the Conservative regime, and thereby laid claim to a possible fivefold welfare state regime typology (table 2). However, by extending the variety of countries used, Leibfried (1992), Ferrera (1996) and Bonoli (1997) identified another distinctive 'Latin', or

'rudimentary', fourth regime type (Spain, Portugal, Greece and, to a lesser extent, Italy and France – table 1).

Tables 1 and 2 show the different results of each of the welfare state classifications, both in terms of the regimes asserted by each typology and the constituent countries within them. A number of countries are positioned in the same regime type in almost all of the classifications. For example, Canada and the USA are positioned in the least generous (Liberal) regime type in all of the typologies and taxonomies; similarly, Denmark, Norway and Sweden always appear together in the most generous regime type (Social Democratic) regardless of which indicators are used to construct the typology. Furthermore, Greece, Portugal and Spain are considered to be the core countries of the Latin regime, and Germany is the one exemplar of the Conservative ideal-type model. The positioning of the other countries, however, is a more disputed matter. For example, Australia, Ireland, New Zealand and the UK are positioned in either the Liberal or the Radical regime type; Austria, Belgium, Finland and the Netherlands are placed in either the Conservative or the Social Democratic regime type; and Italy and France are placed in either the Conservative or the Latin regime type (see tables 1 and 2). Most contentious is the case of Switzerland, which is placed in three different regime types: Liberal (Castles and Mitchell 1993; Ragin 1994; Shalev 1996; Korpi and Palme 1998; Obinger and Wagschal 1998), Conservative (Esping-Andersen 1990; Ferrera 1996; Pitzurello 1999), and Latin (Bonoli 1997).

Table 1 also shows the variety of factors used to construct each of the different welfare state typologies. All of these typologies are designed to capture the income maintenance aspects of welfare state provision, yet they all do it using different indicators: decommodification (Esping-Andersen 1990); basic income (Leibfried 1992); aggregate expenditure (Castles and Mitchell 1993); poverty rates (Ferrera 1996; Korpi and Palme 1998); or social expenditure (Bonoli 1997; Korpi and Palme 1998). It could be, therefore, that each welfare state typology is measuring a different aspect of welfare state provision and thereby drawing attention to different aspects of what welfare states do.² However, many of the factors used to devise the diverse typologies are very similar, such as decommodification (Esping-Andersen 1990) and replacement rates and benefit coverage (Ferrera 1996), while others such as social expenditure as a percentage of GDP are shared by different typologies (Bonoli 1997; Korpi and Palme 1998). Perhaps, therefore, these diverse typologies and taxonomies are ultimately measuring merely slightly different aspects of the same underlying dimensions of the welfare state.

<ha>The Two Dimensions of Welfare</ha>

Indeed, Bonoli (1997) argues that the welfare state typology literature can be divided into two halves: one of which examines 'how much' (i.e. the quantity of welfare provision), while the other examines 'how' (the Bismarck–Beveridge funding dichotomy). Bonoli argues that welfare state typologies (such as those outlined in table 1) measure one of these two underlying aspects of the welfare state more than the other: they are one-dimensional categorizations. For example, he asserts that Esping-Andersen's (1990) typology measures the 'how much' issue whereas the typology of Ferrera (1996) encapsulates the 'how' aspect more thoroughly. Therefore, regardless of the particular individual factors used by each typologist, there is in fact only one of two underlying dimensions being measured. This leads Bonoli to use just two factors (social expenditure as a percentage of GDP and the proportion of contribution financing and tax financing of social expenditure) to draw up a two-dimensional typology. He argues that, taken together, these two factors represent both the 'how much' – the quantity issue – and the 'how' – the Beveridge tax-funded universalism versus the Bismarck contribution-based social insurance system. As table 1 shows, his typology is very similar to those drawn up using a more varied or numerous array of factors.

Bonoli's work is very important in the context of this article as it suggests that, rather than emphasizing the differences between how the various welfare state typologies have been constructed and thereby resigning the literature to one of indecision and incomparability, it is possible to use his two-dimensional approach to compare all of the different typologies and taxonomies on the same terms. It is therefore possible to determine which, if any, of the competing theories of welfare state regimes are the most useful in terms of accounting for welfare state variation in one-dimensional (how much or how) and/or a two-dimensional manner (how much and how).

This article therefore utilizes a two-dimensional approach to examine the competing typologies and taxonomies that exist within the contemporary comparative welfare state literature. The aims are fourfold: firstly, to determine which of the income maintenance-based classifications are the most useful in terms of accounting for welfare state variance; secondly, which of the typologies measure one dimension more than the other and which dimension has dominance overall in welfare state modelling; thirdly, to explore which of the two underlying dimensions distinguishes most between different welfare state regimes; and finally, to compare the results of the more theoretically derived typologies with

the more empiricist taxonomies. Additionally, the value of Bonoli's two-dimensional approach will be given further implicit empirical exploration. Overall, the analysis will 'sift the wheat from the chaff' and enable certain typologies and taxonomies to be given more prominence within discussions about welfare regimes in the comparative social policy literature.

<ha>Methods</ha>

Comparative social policy in general, and the construction of welfare state typologies in particular, has seldom been underpinned by robust methodology (Kangas 1994; Ragin 1994; Shalev 1996; Pitruzello 1999; Gough 2001; Bamba 2006). This has meant that many welfare state typologies are theoretically rather than empirically informed. For example, Esping-Andersen's Three Worlds typology has been extensively criticized on the basis of its reliance on averaging and additive indexes (Fawcett and Papadopoulos 1997; Arts and Gelissen 2002; Bamba 2006). Some commentators, such as Kangas (1994) or Ragin (1994), have responded to these problems by using methods such as cluster analysis or BOOLEAN comparative analysis (see table 2) to develop alternative, more empirically based, welfare state groupings. However, while these taxonomies are clearly a methodological enhancement, they merely serve to develop new, rival, welfare state classifications. These methods are unable to test the relative merits of existing welfare state typologies and thereby serve merely to further fuel discussions about the relative placement of certain countries and the number of different welfare state regimes (see tables 1 and 2). The use of statistical techniques that enable the more extensive comparison and testing of the different welfare state typologies (and indeed, taxonomies), such as analysis of variance (ANOVA), multivariate analysis of variance (MANOVA), and discriminant analysis (DA), has long been advocated within the welfare state modelling literature (Pitruzello 1999).

MANOVA is an extension of ANOVA (analysis of variance) that enables the comparison of the mean values of different groups (in this case welfare state regimes) for a variety of dependent variables (Weinfurt 1995; Tabachnick and Fidell 2001). ANOVA calculates an F-statistic that is the ratio of the variance between the different groups (considered to be due to the independent variable – group membership) divided by the variability within each of the groups (considered to be due to chance). A large and significant F-statistic indicates that there is more variability between the groups than within them and that therefore, the group means are not equal (Pallant 2001). In other words, it shows that the

groups differ significantly in terms of the dependent variable. Post-hoc tests, such as the Scheffe test, compare each group with all other groups to determine where the differences lie. ANOVA thereby enables a comparison of whether the regimes (groups) in each of the different welfare state typologies and taxonomies actually differ from one another in terms of the chosen dependent variables. In addition, through the η^2 effect size statistic, ANOVA enables comparisons to be made between the rival classifications on the basis of how much of the variance in the dependent variable is accounted for by the grouping or independent variable. In other words, η^2 helps to determine which of the diverse welfare state typologies/taxonomies is the most useful because it accounts for the most amount of variation.

MANOVA extends this analysis to include more than one dependent variable (Field 2000). It compares the differences in the means of the group for the dependent variables both individually, through separate ANOVAs, and together, by combining the dependent variables into one new linear composite dependent variable (Pallant 2001). Descriptive discriminant analysis (DA) extends the MANOVA analysis as it explores the underlying dimensions of the data and determines which weighted combination of scores on the two variables best distinguishes between the different groups (Field 2000; Silva and Stam 2001[qu4]). The weightings of variables form a new composite variable: the discriminant function – a linear combination of the weightings and scores on this variable (Cramer 2003). In this way, DA transforms the original set of variables into one or more new functions. DA can also be used to predict group membership (Silva and Stam 2001). Reflecting the MANOVA approach, a non-stepwise method of entry was used in the DA (Field 2000).

The choice of dependent variables is therefore very important in terms of the analysis that ANOVA, MANOVA and DA can provide. Statistically, the dependent variables need to be conceptually related but not highly correlated (multicollinearity); be normally distributed (both univariate and multivariate); have a linear relationship; and exhibit homogeneity of variance and covariance (Pallant 2001). More importantly, though, the chosen variables need to reflect as far as possible the wide variety of factors and considerations used to originally establish the different typologies which, as table 1 shows, range from stratification (Esping-Andersen 1990) to poverty rates (Ferrera 1996; Korpi and Palme 1998) and benefit equality (Castles and Mitchell 1993). This is by no means an easy task especially when, in addition to these variations, the availability of cross-national data is limited. Two dependent variables are used in this article: social expenditure as a percentage of GDP; and employer and

employee contributions as a percentage of total social security receipts. These variables provide what Bonoli (1997) refers to as a two-dimensional approach to welfare state classification (see the discussion above) as together they reflect both the quantity of welfare state provision (social expenditure as a percentage of GDP) and how that provision is funded (employer and employee contributions as a percentage of total social security receipts).

Although much of the welfare state modelling literature has been about trying to get beyond aggregate measures of funding and provision (see for example Esping-Andersen 1990: 2, or indeed table 1), the social expenditure as a percentage of GDP variable is in fact highly correlated (see table 3) with a number of the other indicators that have been used to construct welfare state typologies. For example, the correlation between social expenditure as a percentage of GDP and poverty rates is -0.783 ($p < 0.0001$). Furthermore, social expenditure as a percentage of GDP is used as one of the main measures in a number of welfare state typologies, such as Castles and Mitchell (1993) or Korpi and Palme (1998), and employer and employee contributions as a percentage of total social security receipts is used by Bonoli (1997) and reflects Leibfried (1992). Therefore, although by no means exhaustive, these two robust variables do offer a fairly indicative universal overview of the measures that have been used in welfare state classifications (Bonoli 1997).**[T3]**

The data for the social expenditure as a percentage of GDP variable were obtained from the OECD (2004) and cover 21 countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Norway, the Netherlands, New Zealand, Portugal, Spain, Sweden, Switzerland, the UK, and the USA. These countries (see table 4) represent all of those used by the various different welfare state typologies (with the exception of Luxembourg, which has not been included in the analysis as it is only considered by Ferrera 1996 and Bonoli 1997). The social expenditure as a percentage of GDP variable is the mean calculated for the period 1997–2001 to minimize the influence of any year-on-year fluctuation. The data for the contributions as a percentage of total social security receipts variable were recalculated from the ILO 'Cost of Social Security Database' (1996). Subsequently, the data for this variable are older (ranging from 1990 to 1994) and generally only represent one year's worth of data. This reflects the fact that, although more recent data are available for this variable for the sixteen European countries (EU 2005), for the five non-European countries the ILO (1996) was the only source. Therefore, to ensure data comparability between the European and non-European countries this data source was used for all the countries. It should be noted that these data

relate to different time periods from those used in the original typologies and taxonomies. All analysis was carried out using SPSS version 11.0.[T4]

<ha>Results</ha>

The results of the one-way between groups ANOVAs are presented in tables 5, 6, 7 and 8.³ Overall, they show that the differences between welfare state regimes were statistically significant in the majority of the typologies and taxonomies tested. The significant ANOVAs all show large (>0.25) η^2 effect sizes (Weinfurt 1995).[T5][T6]

The results of the one-way ANOVAs for social expenditure as a percentage of GDP (tables 5 and 6) show that the Ferrera (1996) typology has the highest η^2 score of 0.67. This indicates that 67 per cent of the variance in the dependent variable – social expenditure – is accounted for by the independent variable – Ferrera’s welfare state regimes. Similarly, the results in table 6 show that the typologies of Esping-Andersen (1990), Leibfried (1992), Castles and Mitchell (1993) and Bonoli (1997) all achieve statistically significant differences ($p < 0.05$) between welfare state regimes in terms of social expenditure and they all offer large effect sizes accounting for between 57 per cent (Esping-Andersen) and 65 per cent (Leibfried, Castles and Mitchell) of variance. Indeed, only Korpi and Palme’s (1998) typology does not achieve statistical significance ($p = 0.20$), indicating that there is no significant difference between their welfare state regimes in terms of mean social expenditure.

The results for the taxonomies (table 6) are similar as all bar one (Shalev 1996) achieve statistical significance ($p < 0.05$) and the η^2 values, ranging from 0.50 to 0.84, indicate strong effect sizes. The Wildeboer Schut *et al.* (2001) taxonomy offers the largest η^2 value of 0.84, which suggests that their welfare state regime classification accounts for 84 per cent of variance in social expenditure as a percentage of GDP. Post-hoc comparisons using the Scheffe test indicated that the Liberal (Anglo-Saxon, British, etc.) regime type differed significantly ($p < 0.05$) from the Social Democratic (Nordic, Scandinavian, etc.) regime type consistently across all the typologies and taxonomies that achieved overall significance. This reflects the fact that in all the welfare state models the Liberal-style regime had the lowest value for social expenditure as a percentage of GDP (ranging from 16.15 to 20.83) while the Social Democratic-style regime had the highest (26.63 to 28.03). The post-hoc tests also suggested that in a number of cases (Esping-Andersen, Ferrera, Ragin, Wildeboer Schut *et al.*) the Liberal-style regime differed significantly from the Conservative-style (Bismarck,

Corporatist, etc.) regime. None of the welfare state models indicated a significant difference between the other regime types, with the exception of Kangas (1994), where a significant difference was also found between the Social Democratic regime type and the Radical regime. **[T7][T8]**

The results of the one-way ANOVAs for employer and employee contributions as a percentage of total social security receipts (tables 7 and 8) show a similar pattern for the typologies. Ferrera (1996) again has the highest effect size of 0.65. On this variable all the other typologies achieve statistical significance and large effect sizes (0.38 and above), including the Korpi and Palme model ($p = 0.02$; $\eta^2 = 0.49$). However, the results for the taxonomies are much less positive as only three (Kangas, Obinger and Wagschal, Pitzurello) of the six models show evidence of a statistically significant difference between welfare state regimes on this variable. Obinger and Wagschal's model has the largest effect size of 0.60, suggesting that welfare state regime membership accounts for 60 per cent of the variation in employer and employee contributions as a percentage of total social security receipts. The post-hoc Scheffe tests show a great deal of variety in the models as to which regimes differ significantly from one another, as while the taxonomy Scheffe tests (table 7) reveal significant differences between Conservative and Radical, there is no clear pattern of difference in the typologies (table 7). The typologies show four significant post-hoc differences: Liberal and Conservative (Esping-Andersen, Ferrera); Liberal and Latin (Leibfried); Conservative and Radical (Castles and Mitchell); Conservative and Social Democratic (Bonoli, Ferrera). Generally, these patterns suggest that significant differences in contributions as a percentage of total receipts exist between the higher- and the lower-scoring regimes in each typology. In all the typologies (table 7) and taxonomies (table 8), the Conservative regimes scored the highest (ranging from 60.73 to 74.33) and the Liberal (26.26 to 50.71) and Radical the lowest (8.61 to 23.53).

The MANOVA results, in which the two dependent variables are combined (table 9), reflect the patterns evident in the single dependent variable ANOVAs. Most importantly, Ferrera's typology with an η^2 of 0.66 accounts for the most variance in the combined dependent variable, just as it did for both of the dependent variables in their separate ANOVAs. Similarly, the Korpi and Palme typology once more failed to achieve statistical significance, suggesting that in the combined MANOVA analysis, the means scores of their welfare state regimes do not differ. There are also some interesting differences between the single ANOVAs and the MANOVA results. For example, Esping-Andersen's typology accounted for 57 per cent of the variance in social expenditure as a percentage of

GDP (table 5) and 38 per cent of employer and employee contributions (table 7). However, in the MANOVA analysis of the combined variable Esping-Andersen's typology only accounted for 37 per cent of the variance (table 9). Conversely, Shalev's typology, which did not achieve a statistically significant ANOVA for either variable, accounted for 42 per cent of the variance in the combined dependent variable. These results reflect the fact that the MANOVA analysis takes into account the correlation between the dependent variables. Also, the overall MANOVA results suggest that some typologies and taxonomies perhaps reflect one variable – and therefore one dimension of welfare provision – more than the other, while others reflect both variables and therefore both dimensions of the welfare state.**[T9]**

Discriminant analysis (DA) offers the opportunity to examine these underlying dimensions in more detail and determine which weightings of variables (functions) discriminate the most between the different groups in each typology and taxonomy. Table 10 shows the key statistics produced by the DA for each of the welfare state typologies and taxonomies. Around half of the typologies and taxonomies reflect two underlying dimensions (reflected by the number of statistically significant discriminant functions) but the other half reflect only one, suggesting that only one dimension distinguishes between the groups. Those typologies and taxonomies that have variate correlations which are positive for both of the variables (social expenditure as a percentage of GDP, employer and employee contributions as a percentage of total social expenditure receipts), such as function 1 for Esping-Andersen, suggest that it is the combination of the two variables that discriminates between the groups. Those typologies that have significant discriminant functions with one negative correlation and one positive correlation with the variables, such as function 2 for Leibfried, indicate that it is the difference between the variables that separates the groups. For each significant discriminant function, the variable with the higher canonical correlation coefficient contributes the most to group separation. The group variate centroids (mean function scores for each group) show which groups the discriminant function distinguishes between. So, for example, in the case of Esping-Andersen's typology (see table 10) there is only one significant discriminant function (function 1 $p = .002$), it is the combination of the variables that distinguishes between the groups as both correlations are positive ($r = .789$; $r = .509$), social expenditure contributes the most to group separation, $r = .789$ (although the contribution of the other variable is still large, $r = .509$), and the group variate centroids suggest that the discriminant function discriminates the Social Democratic group 3 from the other two groups, particularly the Liberal group 1

(as the difference between these two scores, 1.076 and -1.879, is the largest). **[T10]**

The DA shows that the welfare state classifications of Esping-Andersen, Ragin, Shalev, Pitzurello and Wildeboer Schut *et al.* all reflect only one underlying dimension: in all cases except Shalev, it is the combination of the variables that distinguishes between the groups; social expenditure contributes the most to group separation (although in the cases of Esping-Andersen and Pitzurello the contributions correlation is also large). Also, the variate centroids suggest that the function discriminates the most between the Liberal group 1 and the Social Democratic group 3 (Esping-Andersen, Shalev, Ragin); the Liberal group 1 and the Radical group 4 compared to the other groups (Pitzurello); and between the Liberal group 1 and the other groups (Wildeboer Schut *et al.*). The typologies of Leibfried, Castles and Mitchell, Ferrera, Bonoli, Kangas, and Obinger and Wagschal encapsulate a two-dimensional approach as each has two significant discriminant functions. In each case, one of the discriminant functions shows social expenditure as contributing most to the discrimination between the groups and the other shows contributions as a percentage of total social security receipts as contributing most to group discrimination. The variate centroids suggest that the social expenditure weighted discriminant function discriminates most between the Liberal group 1 and the other groups (Leibfried, Bonoli, Kangas), while the discriminant function weighted more by contributions as a percentage of total social security receipts distinguishes most between the Social Democratic group 3 and the other groups (Leibfried, Ferrera, Bonoli, Kangas, Obinger and Wagschal). However, in the fourfold regime classification of Castles and Mitchell, the contributions weighted variate centroids distinguish most between the Radical group 4 (-1.306) and the others, particularly the Conservative group 2 (.943). Again, as expected from the MANOVA, the DA of the Korpi and Palme typology was not significant.

To summarize, the results of the single ANOVAs showed that out of the typologies, Ferrera's accounted for the highest amount of variance for each of the two variables, and that differences between particular groups varied by typology. The MANOVA confirmed Ferrera's position as the highest-scoring typology and Korpi and Palme as the lowest; and the DA suggested that some welfare state classifications were two-dimensional, while those that were one-dimensional reflected the social expenditure as a percentage of GDP variable more than the contributions as a percentage of total social security receipts variable. Furthermore, the DA provided clarity, unattained via the single ANOVAs, that

both variables distinguished most between the Liberal groups and the other groups, or the Social Democratic groups and the others.

Finally, as the welfare state taxonomies were in part developed to be a more methodologically robust and empirical way of classifying welfare states than the typologies (Kangas 1994; Ragin 1994; Shalev 1996; Pitruzzello 1999), it is worth considering whether the taxonomies account for more of the variance in the two dependent variables and in the combined MANOVA variable than the typologies. One-tailed Mann-Whitney U tests were used to compare the average η^2 values of the typologies and the taxonomies for each of the variables and the combined MANOVA variable: no significant differences were found (social expenditure $p = 0.59$; employer and employee contributions $p = 0.07$; combined $p = 0.59$). This suggests that overall the taxonomies did not account for significantly more variance than the typologies.

<ha>Discussion</ha>

One of the central aims of this article was to discover which of the various competing welfare state classifications, both typologies and taxonomies, were the most currently useful and which the least. In one respect the results have enabled this aim to be fulfilled, as it is clear that some typologies account for more welfare state variance than others. However, the amount of variance in the two variables – social expenditure and employee and employer contributions – accounted for by a number of the typologies and taxonomies is in fact very similar. Ferrera's (1996) typology achieved the highest η^2 values in both of the single ANOVAs (0.67 for social expenditure and 0.65 for employer and employee contributions), and it also scored the highest η^2 value in the MANOVA. It would therefore be simple to conclude that the Ferrera typology accounts for the most variance and is therefore the most useful of the competing welfare state classifications – the 'wheat', to follow through on the metaphor. However, the results are in fact much less clear-cut than this as a number of the classifications achieved η^2 values very similar to Ferrera's (and not significantly different from one another). For example, in the social expenditure single ANOVA, the typologies of Leibfried (1992), Castles and Mitchell (1993), and Bonoli (1997) also achieved η^2 values in the 0.60s, as did the taxonomies of Kangas (1994), Ragin (1994) and Pitruzzello (1999). Indeed, on this variable Ferrera was not actually the highest-scoring classification as the Wildeboer Schut *et al.* (2001) taxonomy achieved an η^2 value of 0.82. Similarly, in the single ANOVA for employer and employee contributions, Leibfried (1992), Bonoli (1997) and

Obinger and Wagschal (1998) all obtained similar η^2 scores to Ferrera. On the whole, the MANOVA results were likewise, with a number of typologies scoring similarly. This suggests that, far from having one single typology that stands out from the others, there are in fact a number of typologies that offer comparable levels of variance explanation. This poses somewhat of a dilemma as it means that, as a result of this endeavour, the welfare state modelling literature is actually little nearer to determining which individual classification model is the most useful.

However, while the analysis has been unable to determine conclusively which classification system is the 'wheat', it has been more successful in 'sifting out the chaff': the classifications of Korpi and Palme (1998) and Shalev (1996) were non-significant for the social expenditure single ANOVA, and the taxonomies of Ragin (1994), Shalev (1996), and Wildeboer Schut *et al.* (2001) were non-significant in the employee and employer contributions single ANOVA. The MANOVA replicated the non-significant results for Korpi and Palme (1998) but it also produced another, less expected result: despite achieving high η^2 scores in each of the single ANOVAs, Esping-Andersen's 'Three Worlds' typology, by far the most prominent of all welfare state classification systems, accounted for the lowest amount of variance (apart from the non-significant results) in the MANOVA. This suggests perhaps that his typology is less useful than others when considering a two-dimensional approach. Overall, the analysis suggests that the typologies and taxonomies shown to score low η^2 values or non-significant ANOVAs can perhaps therefore be given less accord within future discussions about the number and consistency of welfare state regimes.

However, the typologies and taxonomies cannot be discounted or reinforced on the basis of the ANOVA, MANOVA and η^2 analyses alone. The results of the DA must also be considered, particularly in terms of the extent to which the different welfare state classifications measure one or two dimensions of welfare state provision. Those typologies and taxonomies shown by the DA to measure only one or other of the two underlying welfare state dimensions of 'how much' and 'how' (Bonoli 1997) should be given less prominence than those that are shown to reflect both aspects more adequately. The DA confirmed that in some cases the typologies and taxonomies reflected only one dimension. For example, in the classifications of Esping-Andersen (1990), Ragin (1994), Shalev (1996), Pizurello (1999) and Wildeboer Schut *et al.* (2001), social expenditure was much more prominent, whereas in other cases, most notably Leibfried (1992), Castles and Mitchell (1993), Ferrera (1996) and Bonoli (1997), the DA showed that typologies can encapsulate both aspects of welfare state provision.

Those that measure both aspects should surely be given more weighting in the welfare state modelling literature than those that only examine one.

This is especially the case for those typologies, such as that of Esping-Andersen, for which claims are made that they 'get beyond aggregate measures of welfare state expenditure' by creating alternative means of comparison, such as decommodification. The DA results show that for the majority of typologies and taxonomies it is the social expenditure as a percentage of GDP weighted variate that discriminates most between the regime types. When these results are taken together with the correlations between social expenditure as a percentage of GDP and other measures of welfare state provision, such as decommodification, it suggests that a number of welfare state typologies still only reflect expenditure and little more. In common with criticisms made by those such as Castles and Mitchell (1993) or Bonoli (1997), it suggests that a lot of the attempts to 'get beyond' using expenditure as a means of comparing and contrasting different welfare states have failed, and that perhaps this aspect of welfare state modelling still needs to be fully achieved. The analysis in this article suggests that the attempt to 'get beyond' aggregate expenditure comparisons should remain as one of the focuses of comparative research.

The two-dimensional analysis has also revealed which welfare state regimes differ most from one another in terms of either and both dimensions. The results of the post-hoc tests in the single ANOVAs suggested that in the majority of the different typologies and taxonomies, the social expenditure variable (the 'how much' dimension) distinguished most between the Liberal-type and the Social Democratic-type regimes, in other words between the highest and the lowest expenditure regimes in all of the typologies. Perhaps more interestingly, though, in those classifications that included a Radical fourth regime type, the single ANOVA for employer and employee contributions revealed significant differences between the Conservative and Radical regime types. It was, as expected from the literature, the Conservative regimes that scored the highest in terms of employer- and employee-based funding; however, it was the Radical rather than the Liberal regimes that scored the lowest. These results reinforce one of the common tenets of the welfare state literature - the stark differences between the Liberal and Social Democratic regime types (however constituted) - but it also lends further support to the arguments of those such as Castles and Mitchell (1993) that a distinctive Radical regime exists, one which is more clearly Beveridgian than Bismarckian (Bonoli 1997) in orientation, even more so than the Liberal-style regime. To some extent, the DA reinforces these conclusions as the discriminant functions weighted more by social expenditure were shown to

discriminate most between the Liberal regime and the other regime types. However, while the post-hoc tests in the contributions single ANOVA suggested differences between the Conservative and the other welfare state regime types, in the DA, the weighted variate was better at separating the Social Democratic regime type from the others. This perhaps suggests that it is the lack of contributions-based funding that is creating the divisions. The DA does, however, go a little further than the single ANOVAs in terms of lending support to proponents of the Radical regime type as in some of the typologies and taxonomies, such as Pitzurello (1999), the social expenditure weighted variate also discriminated well between the Liberal and Radical regimes, and in the case of Castles and Mitchell (1993), the Radical group was distinguished from the others most by the contributions weighted variate.

The two-dimensional approach proposed by Bonoli (1997) and used within this analysis has therefore produced a means of comparing and contrasting the different regime classifications that are currently circulating in the comparative welfare state literature. Although not without its own limitations, which are acknowledged by Bonoli himself, the approach does offer a productive way forward in terms of enabling the construction of a welfare state typology which reflects both aspects of the welfare state modelling literature, at least as it currently stands in terms of the income maintenance dimension. It is perhaps surprising, then, that the results in this article have not shown Bonoli's own typology, based as it is on the two dimensions alone, to be of even more utility. There are at least three reasons for this (but it needs to be reiterated that Bonoli's typology accounts for a very high amount of variance in the MANOVA with an η^2 of 0.62). On the one hand, Bonoli's analysis was much earlier than the one in this article and the data he used were from a different date and source. However, perhaps more influential was the fact that the method used by Bonoli to put countries into regime types relied merely on cross-classification (see Bonoli 1997), and therefore, unlike the MANOVA, it did not take into account any correlation between the two dimensions, and this may have led to incorrect classification. Alternatively, of course, there could be a third underlying dimension of welfare state provision that Bonoli's two-dimensional approach overlooks, such as the mix of cash benefits and welfare state services (Castles 1998; Kautto 2002; Bambra 2005a, 2005b). The two-dimensional approach could therefore be extended both methodologically, by using a more robust system of classification such as cluster analysis, and empirically, by including this third dimension – the 'how spent' aspect of welfare state provision. Perhaps then, the welfare state modelling literature will finally succeed in obtaining a holistic typology.

<ha>Limitations</ha>

The research in this article is subject to a number of limitations. Firstly, as pointed out in the discussion section, it only examines two dimensions, which are themselves based on two indicators – social expenditure as a percentage of GDP and employer and employee contributions as a percentage of total social security receipts. If other indicators, or more indicators, had been used the results may have been different. Therefore, although multivariate analysis was used, this means that it is likely that other, unaccounted-for factors are exerting influence on welfare state variation. Subsequently, caution should be applied to the results and their interpretation, not least as the statistical correlations and associations discussed do not necessarily equate with explanation or causation. Furthermore, the data used for the contributions variable were prone to possible year-on-year variation and they were limited to the years 1990–4. Also, it needs to be noted that the various typologies and taxonomies were based on data from differing years (e.g. Esping-Andersen's data were from 1980) and so this article cannot comment on the original fit of the classifications, it is only able to examine their current relevance. The typologies and taxonomies may well have been more accurate at the time they were originally constructed (e.g. in the 1980s for Esping-Andersen), and they may still be currently accurate if all their constituent variables are included (e.g. if decommodification indicators were reanalysed for the 1990s for Esping-Andersen, or institutional characteristics were somehow included in the reassessment of the Korpi and Palme model).⁴ It is also important to acknowledge that welfare regime patterns change over time and that this article has only examined the applicability of welfare state classifications at one point in time. It is quite possible that different typologies and taxonomies will be accurate at different time points. Finally, most typologies were perhaps intending to measure something other than the two dimensions, and indeed some of them are based on a variety of factors, not all of which are quantifiable. The analysis in this article is perhaps therefore limited by overly quantifying and thereby marginalizing the more qualitative and theoretical aspects of typology construction.

<ha>Conclusion</ha>

Although the analysis has been unable to determine with any certainty which individual classification is the most useful, it has provided evidence that one

subgroup of the diverse regime classifications is more useful than the others. This is particularly the case when examining whether the classifications encapsulate one dimension or two dimensions of the welfare state. It is proposed in this article that those typologies that, in the DA, reflect the two underlying dimensions should be given more prominence than those that only reflect one. Furthermore, the η^2 results for the ANOVAs and the MANOVA should also be taken into consideration, with those classifications obtaining higher η^2 values accounting for more variation. This means that the four typologies of Leibfried (1992), Castles and Mitchell (1993), Ferrera (1996) and Bonoli (1997), and the two taxonomies of Kangas (1994) and Obinger and Wagschal (1998) emerge as the 'wheat' while the others, particularly those that have non-significant differences between regime types, particularly Korpi and Palme (1998) and to a lesser extent Shalev (1996), or which only reflect one dimension of the welfare state, such as Ragin (1994), Shalev (1996), and Wildeboer Schut *et al.* (2001) and to a lesser extent Esping-Andersen (1990) and Pitzurello (1999), are the 'chaff'. In addition, the analysis has reinforced the fact that the differences between certain welfare state regime types are more pronounced, particularly those between the Liberal and the Social Democratic types. However, it has also highlighted and upheld arguments for the existence of a distinctive Radical regime type. Methodologically, this article has pioneered the use of ANOVA, MANOVA and DA in terms of the comparative welfare state literature and has demonstrated the value of these methods in taking the modelling debate forward. Future research should rely more on such sophisticated analytical approaches and expand empirically upon the two-dimensional approach of 'how much' and 'how' by adding the third dimension of 'how spent' in terms of the cash benefits and welfare state services mix (Bambra 2005a, 2005b).

<ha>Notes </ha>

1. <not>Esping-Andersen also reflects upon social stratification and the public-private mix in welfare state provision. However, it is his decommodification index typology that is widely cited (see, for example, Arts and Gelissen 2002).
2. It should also be noted that the Korpi and Palme typology is also based on more theoretical and qualitative data about the institutional characteristic of social insurance.
3. It is more usual to run the MANOVA first and then follow up with the single ANOVAs.

4. An updated analysis of the full decommodification indicator set by Bamba (2006) suggests, however, that in the case of Esping-Andersen the threefold typology is no longer accurate. Furthermore, research by Scruggs and Allan (2005) has questioned the original 1980 validity of the threefold classification.

References

- Abrahamson, P. (1999), The welfare modelling business, *Social Policy & Administration*, 33, 4: 394–415.
- Alber, J. (1995), A framework for the comparative study of social services, *Journal of European Social Policy*, 5, 2: 131–49.
- Arts, W. and Gelissen, J. (2002), Three worlds of welfare or more? *Journal of European Social Policy*, 12, 2: 137–58.
- Bamba, C. (2004), The worlds of welfare: Illusory and gender-blind?, *Social Policy and Society*, 3, 3: 201–12.
- Bamba, C. (2005a), Worlds of welfare and the health care discrepancy, *Social Policy and Society*, 4, 1: 31–41.
- Bamba, C. (2005b), Cash versus services: worlds of welfare and the decommodification of cash benefits and health care services, *Journal of Social Policy*, 34, 2: 195–213.
- Bamba, C. (2006), Decommodification and the worlds of welfare: revisited, *Journal of European Social Policy*, 16, 1: 73–80.
- Bonoli, G. (1997), Classifying welfare states: a two-dimension approach, *Journal of Social Policy*, 26, 3: 351–72.
- Castles, F. G. and Mitchell, D. (1993), Worlds of welfare and families of nations. In F. G. Castles (ed.), *Families of Nations: Patterns of Public Policy in Western Democracies*, Aldershot: Dartmouth.
- Castles, F. G. (1998), *Comparative Public Policy: Patterns of Post-war Transformation*, Cheltenham, Edward Elgar.
- Cramer, D. (2003), *Advanced Quantitative Data Analysis*, Maidenhead: McGraw-Hill.
- Cramer, D. and Howitt, D. (2004), *The Sage Dictionary of Statistics*, London: Sage.
- Cutright, P. (1965), Political structure, economic development, and national social security programs, *American Journal of Sociology*, 70: 537–50.
- Esping-Andersen, G. (1987), Citizenship and socialism: decommodification and solidarity in the welfare state. In G. Esping-Andersen and L. Rainwater

- (eds), *Stagnation and Renewal in Social Policy: The Rise and Fall of Policy Regimes*, London: Sharpe.
- Esping-Andersen, G. (1990), *The Three Worlds of Welfare Capitalism*, London: Polity.
- Esping-Andersen, G. (1999), *Social Foundations of Post-industrial Economies*, Oxford: Oxford University Press.
- European Union (EU) (2005), *European Social Statistics: Social Protection Expenditure and Receipts*, Luxembourg: Eurostat.
- Fawcett, H. and Papadopoulos, T. N. (1997), Social exclusion, social citizenship and decommodification: an evaluation of the adequacy of support for the unemployed in the European Union, *West European Politics*, 20, 3: 1–30.
- Ferrera, M. (1996), The southern model of welfare in social Europe, *Journal of European Social Policy*, 6, 1: 17–37.
- Field, A. (2000), *Discovering Statistics Using SPSS for Windows*, London: Sage.
- Ginsburg, N. (1992), *Divisions of Welfare: A Critical Introduction to Comparative Social Policy*, London: Sage.
- Gough, I. (2001), Social assistance regimes: a cluster analysis, *Journal of European Social Policy*, 11, 2: 165–70.
- ILO (International Labour Organisation) (1996), *The Cost of Social Security*. Available at: <http://www.ilo.org/public/english/protection/socfas/research/css/cssindex.htm> (accessed 3 April 2005).
- Kangas, O. (1994), The politics of social security: on regressions, qualitative comparisons and cluster analysis. In T. Janoski and A. Hicks (eds), *The Comparative Political Economy of the Welfare State*, Cambridge: Cambridge University Press.
- Kautto, M. (2002), Investing in services in West European welfare states, *Journal of European Social Policy*, 12, 1: 53–65.
- Korpi, W. (2000), Faces of inequality: gender, class and patterns of inequalities in different types of welfare states, *Social Politics*, 7, 2: 127–91.
- Korpi, W. and Palme, J. (1998), The paradox of redistribution and the strategy of equality: welfare state institutions, inequality and poverty in the Western countries, *American Sociological Review*, 63: 662–87.
- Leibfried, S. (1992), Towards a European welfare state. In Z. Ferge and J. E. Kolberg (eds), *Social Policy in a Changing Europe*, Frankfurt: Campus-Verlag.
- Lewis, J. (1992), Gender and the Development of Welfare regimes, *Journal of European Social Policy*, 2, 3: 195–211.

- Lewis, J. and Ostner, I. (1995), Gender and the evolution of European social policy. In S. Leibfried and P. Pierson (eds), *European Social Policy*, Washington, DC: Brookings Institution.
- Obinger, H. and Wagschal, U. (1998), Das Stratifizierungskonzept in der Clusteranalytischen Überprüfung. In S. Lessenich and I. Ostner (eds), *Welten des Wohlfahrtskapitalismus: Der Sozialstaat in vergleichender Perspektive*, Frankfurt: Campus Verlag, pp. 109–35.
- OECD (Organization for Economic Cooperation and Development) (2004), *Social Expenditure Database*. Available at: www.oecd.org/els/social/expenditure (accessed 5 March 2005).
- Orloff, A. (1993), Gender and the social rights of citizenship: the comparative analysis of gender relations and welfare states, *American Sociological Review*, 58: 303–28.
- Pallant, J. (2001), *SPSS Survival Manual*, Maidenhead: Mcgraw-Hill.
- Pascall, G. and Lewis, J. (2004), Emerging gender regimes and policies for gender equality in a wider Europe, *Journal of Social Policy*, 33, 3: 373–94.
- Pierson, C. (1998), *Beyond the Welfare State*, London: Polity.
- Pitruzello, S. (1999), *Decommodification and the Worlds of Welfare Capitalism: a Cluster Analysis*, Florence: European University Institute.
- Ragin, C. (1994), A qualitative comparative analysis of pension systems. In T. Janoski and A. Hicks (eds), *The Comparative Political Economy of the Welfare State*, Cambridge: Cambridge University Press.
- Sainsbury, D. (1999), Gender, policy regimes and politics. In D. Sainsbury (ed.), *Gender and Welfare State Regimes*, Oxford: Oxford University Press.
- Scruggs, L. and Allan, J. (2005), Welfare state decommodification in eighteen OECD countries: a replication and revision. Available at: <http://www.polisci.uconn.edu/Pages/People/Faculty/percent20Pages/Scruggs/scurggspapers.htm> (accessed 18 January 2006).
- Shalev, M. (1996), *The Privatisation of Social Policy? Occupational Welfare and the Welfare State in America, Scandinavia and Japan*, London: Macmillan.
- Siaroff, A. (1994), Work, welfare and gender equality: a new typology. In D. Sainsbury (ed.), *Gendering Welfare States*, London: Sage.
- Silva, A. and Stam, A. (1995), Discriminant analysis. In L. Grimm and P. Yarnold (eds), *Reading and Understanding Multivariate Statistics*, Washington, DC: American Psychological Association.
- Tabachnick, B. and Fidell, L. (2001), *Using Multivariate Statistics*, London: Allyn and Bacon.
- Titmuss, R. (1974), *Social Policy*, London: Allen and Unwin.

- Weinfurt, K. (1995), Multivariate analysis of variance. In L. Grimm and P. Yarnold (eds), *Reading and Understanding Multivariate Statistics*, Washington, DC: American Psychological Association.
- Wildeboer Schut, J., Vrooman, J. and de Beer, P. (2001), *On Worlds of Welfare: Institutions and their Effects in 11 Welfare States*, The Hague: Social Office of the Netherlands.
- Wilensky, H. and Lebaux, C. (1958), *Industrial Society and Social Welfare*, New York: Russell Sage.
- Wilensky, H. (1975), *The Welfare State and Equality*, Berkeley: University of California Press.</ref>