

GETTING AHEAD OF OURSELVES - CAN INTERNATIONAL COMPARATIVE RESEARCH TELL US HOW SOCIAL POLICY AFFECTS SUBJECTIVE WELL-BEING?

KOLBEINN STEFANSSON¹

It seems something of a truism that the efforts of the welfare state ought to affect the well-being of their citizens. We might thus regard social policy as something of a failure if we were to find that it didn't. This was Ruut Veenhoven's conclusion in two papers (1995 and 2000a) where he presented his findings from an analysis of the relationship between various indicators of welfare state effort and well-being. In the current analysis I take issue with Veenhoven's conclusions and argue that they were premature. This conclusion does not so much arise from the falsification of his results as from methodological and theoretical issues rising in the course of the analysis. These suggest that it may not be viable to address this issue from an international comparative perspective, at least not given the present state of knowledge.

Veenhoven's analysis is novel in that it emphasizes that well-being is the property of individuals and that it ought to be measured as such. This is an argument with which I agree in principle. Veenhoven also takes a multidimensional view of human well-being, measuring it in terms of life satisfaction, self-reported happiness, longevity and subjective health. While it could hardly be claimed that those indicators reflect all aspects of human well-being Veenhoven does well to resist the temptation of regarding human well-being as reducible to a single ultimate good. While applauding this apparent pluralist conception of well-being I will limit my analysis to subjective well-being (SWB). This should not be taken to imply any form of subjectivist monist position on my behalf (for discussion of subjectivism and value monism see Sumner, 1996). On the contrary, I choose to limit my analysis in this way to highlight some problems with the ambitious extensions of what has become known as happiness economics into the area of policy research (see for instance van Praag, 2007) Veenhoven himself elsewhere (2000b) seems to argue for granting measures of happiness a privileged position as measures of well-being, though his argument seems to be based on pragmatic considerations rather than on value monism.

There are two substantial reasons why I am skeptical about the validity of Veenhoven's results. The first of these is that like so much of cross-sectional international comparative research Veenhoven's analysis suffers from a small-N problem. In his two papers the number of observations behind specific statistical relationships ranges from 18 to 38. With so few observations correlations must be quite strong to be judged statistically significant. Given the complexity of subjective well-being as well as the number of things that affect it which lie outside the reach of the welfare state, we have little reason to expect such strong correlations to appear at the international level. This does not necessarily imply that such welfare state effects are trivial or negligible. I suspect that by rejecting the correlations he nevertheless found on the grounds of statistical significance may have lead Veenhoven to commit a type II error, i.e. mistakenly accepting the null hypothesis. In this analysis I attempt to overcome the small-N problem by using pooled time-series cross-section data (TSCS hereafter).

The second reason is Veenhoven's choice of indicators of welfare state effort, though fairness requires that I acknowledge that the measurement of welfare state activity is not uncontroversial (Andersen, 2007) and Veenhoven does well to employ a range of indicators. Two of these seem rejectable out of hand. The first is the number of years since countries passed their first laws on social security, which is a very crude proxy for welfare state development as there are some striking counter-examples (as Veenhoven notes himself; 1995, p. 9). The second is government expenditures in various forms. The reason for rejecting these is that government expenditures, disbursements and consumption contain various components that have no obvious bearing on the issue of human well-being.

The third set of indicators used by Veenhoven is indicators on "entitlements". This is Esping-Andersen's decommodification scale (Esping-Andersen, 1990). Like Veenhoven I shall make use of them in the present analysis as they reflect the extent and quality of entitlements they provide citizens with rather

¹ I would like to thank professor Ruut Veenhoven for generously taking time to comment on the original version of this paper. I would also like to thank participants at the Fourteenth International Research Seminar on Issues in Social Security in Siguna on 15-17th of June 2007 for their comments and advice. Finally I would like to thank Professor Stefan Olafsson and Thora Kristin Thorsdottir for their helpful critique and advice on matters of theory, method and presentation.

than how much they spend on welfare related issues (some of which are better indicators of social problems, such as unemployment, rather than of people's welfare). This is in line with current thinking in much of the welfare state literature, though much of it is critical of this particular scale (See for instance O'Connor, 1993; and Bambra, 2006). One must keep in mind, however, that decommodification is only one among many possible objectives of welfare states and one would need a far wider range of objectives (e.g. defamilialization; Lister, 1995) in order to test conclusively whether welfare states affect the well-being of their citizens.

The fourth set of indicators used by Veenhoven are on social expenditures. The use of such indicators in welfare state research has been the subject of some criticism (following from Titmuss, 1958). Yet indicators of social expenditures can be useful proxies of welfare state effort as long as one is sensitive to the fact that these are highly aggregate measures that obscure some of the issues, for instance by confounding expenditures that arise in response to growing need with generosity of entitlements (Kangas 1991). This is in part a question of causal direction similar to that discussed quite eloquently by Ní Bhrolcháin (2001) in the context of "divorce effects". The real question isn't really whether countries with high social expenditures fare better in terms of well-being than those with lower levels of such expenditures, but rather how each and every country would have fared had it spent more or less on social policy than it in fact did. This issue, however, will not be dealt with directly in this paper.

Another issue I take with Veenhoven's use of social expenditure data which I will address directly is that he only examines indicators of social expenditures *as a proportion of the GDP*. It must be acknowledged that social expenditures as a proportion of the GDP is a classical measure of welfare state effort and I concede that using it makes a great deal of sense if well-being is a relative matter, e.g. function of social comparisons. There is ample evidence that this is the case for subjective well-being (e.g. Ferrer-I-Carbonell, 2005) and some evidence that this is also the case for physical health and longevity (Marmot, 2005). Nevertheless it also seems plausible that people's well-being is a function of the resources devoted to their well-being in some absolute sense and taking that into consideration certainly adds to the completeness of the picture. I therefore consider social expenditures per capita rather than as a proportion of the GDP.

1.2. DATA

This analysis relies on data from three sources, i.e. the OECD,ⁱⁱ the World Happiness Database,ⁱⁱⁱ and a data-set on entitlements compiled by Scruggs and Allan (2006).^{iv} In this section I review indicators by source and discuss theoretical and methodological issues that they give rise to. The appendix to this chapter shows a table giving overview of all the indicators that are used in this study and in which years they were observed in each country.

1.2.1. ORGANIZATION FOR CO-OPERATION AND DEVELOPMENT (OECD)

The OECD gathers information on social and economic issues in 30 member countries. The indicators are standardized as far as possible to facilitate international comparisons. The OECD indicators used in this study are GDP per capita and social expenditures per capita. Both information on social expenditures and GDP are set at fixed prices and purchasing power parities to increase comparability between countries and over time.

ⁱⁱ <http://stats.oecd.org/wbos/default.aspx>.

ⁱⁱⁱ <http://www1.eur.nl/fsw/happiness/>.

^{iv} <http://www.sp.uconn.edu/~scruggs/wp.htm>.

1.2.2. WORLD HAPPINESS DATABASE

The World Happiness Database is directed by Ruut Veenhoven. The database contains information on means, distributions and statistical association on a vast range of subjective well-being indicators in 90 nations from various years, including both happiness and life-satisfaction. For the present study I chose two indicators, one on life-satisfaction and one on happiness. The first indicator is a 4-step indicator on life-satisfaction. People were asked “How satisfied are you with the life you lead?” and given the following response options: “1) Very satisfied, 2) fairly satisfied, 3) not very satisfied, and 4) not satisfied at all”. The second question was a 4-step indicator on happiness. People were asked “Taking all things together, would you say you are?” and were given the following response option: “1) Very happy, 2) quite happy, 3) not very happy, and 4) not happy at all.” The choice of the first measure was dictated by the fact that it seemed to be the most frequently used measure. Thus it provided the highest number of country/time-points which was instrumental to my strategy for overcoming the small-N problem. However, there were no observations for three countries included in this study, namely Australia, New Zealand and Switzerland.

An indicator on happiness is included for conceptual reasons. In much research on subjective well-being, indicators on satisfaction and happiness are used interchangeably (e.g. Easterlin, 2005). According to some of the literature, however, happiness taps more of the affective aspects of subjective well-being, e.g. the balance of pleasant and unpleasant experiences, while satisfaction is thought to be cognitive and evaluative stemming from the gap between aspirations and achievement. If this is the case then results referring to one might not be applicable to the other. The four step indicator of happiness was chosen because of its similarity to the 4-step measure of life-satisfaction to reduce the probability that differences in association might be attributable to scale differences rather than in differences in content.

1.2.3. ENTITLEMENTS

Given the well known limitation of social expenditure as indicator of welfare state development and generosity it seemed sensible to include indicators of the qualities of different welfare state, i.e. what they do. To gather and process the necessary data to construct indexes and indicators of welfare state functions is both time consuming and cumbersome. Consequently one must rely on the efforts of others. Unfortunately it would seem that very few scholars have had both the time and the inclination to compile such data and consequently information on the specific workings of welfare states is scarce, especially data that covers extended periods of time.

Recently Scruggs and Allan (2006) attempted to reconstruct Esping-Andersen's decommodification index. They computed the reconstructed decommodification index for the 18 countries included in Esping-Andersen's original study for the years between 1971 and 2002, though indicators for Germany were only available for the years between 1973 and 2002 (see footnote 3). This allows us to do more sophisticated analysis of the relationship between decommodification and subjective well-being than was previously possible.

There are obviously many aspects of welfare states other than decommodification that might be of relevance to this analysis, not the least the contributions of feminist scholars emphasizing the three-way interaction between families, markets and states (e.g. Sainsbury, 1996). Many of these aspects could reasonably be expected to affect people's subjective well-being and will undoubtedly be studied as more time series of relevant indicators become available. Unfortunately I do not have access to indicators of such aspects and must make do with incomplete information.

1.2.4. CASE SELECTION

Countries are included in the study on account of data availability rather than theory (which is conspicuously absent for reasons that will be discussed in the concluding section). The determining factor was the availability of information on entitlements. This poses a serious problem for generalizability of results. Firstly, measures of statistical significance assume that the sample is somehow representative of a wider population. This representativeness is usually guaranteed via random sampling from the wider

population. As this is obviously not the case we cannot rely on measures of statistical significance to guide the interpretation of the results.

This might not be a problem if the countries in the analysis comprised a theoretically defined population of nations as the slope coefficients could be interpreted as representing “true” relationships (assuming of course that there is no measurement error, that the model is a reasonable approximation of a “true model”, and that the theory defining the population of nations is accurate - which is probably assuming a lot). This is obviously not the case in the present analysis since the data availability criteria is obviously atheoretical. These problems are all shared with Veenhoven’s analysis. I will nevertheless use measures of statistical significance as heuristic devices when interpreting results as doing so helps advance the general point of this paper.

The criteria for inclusion in the analysis was firstly that repeated measures of all relevant variables were available over time, as this was a necessary condition for the use of TSCS analysis to increase the number of observations. This is a strategy for reducing the risk of making Type 2 errors on account of the small-N problem (Goldthorpe, 2000, chapter 3). This left us with 14 countries. All analyses were conducted for these 14 countries and them alone. This was necessary as preliminary analysis that is not reported here indicated that the inclusion and exclusion of particular countries could affect results substantially. Thus it was important to guarantee that if different measures and methods produced different results that those differences did not arise from sample composition but rather from the methods and the measures themselves.

Although the countries included in this study have repeated measures on all relevant variables the total number of observations varies from country to country. In addition different years are missing for different countries. In terms of the TSCS analysis this means that the analyses are based on non-balanced time-series with imbedded missing values (i.e. the strings of observations on specific countries are not complete).

1.3. CROSS-SECTIONAL ANALYSIS

In the next two sections we examine the empirical evidence bearing on the question at hand, i.e. whether welfare states affect the level and the distribution of the subjective well-being of their people. We begin by examining simple cross-sections and then move on to TSCS analysis.

1.3.1. THE LEVEL OF SUBJECTIVE WELL-BEING

For this first part of the analysis I gathered observations for each country for years in which the availability of all relevant indicators coincided. This means that observations for different countries do not necessarily coincide in time. The earliest observations are drawn from 1995 and the latest from 2002. Furthermore, the analyses of happiness and life-satisfaction are based on observations drawn from the same years to ensure that any differences that might arise result from differences between the measures rather than from differences between occasions of observation. The details can be seen in the appendix.

Table 1.1: Bivariate and partial correlations (controlling for GDP) between indicators of welfare state effort and indicators of subjective well-being

	Life satisfaction	Happiness
Social expenditures (per capita)	0.338(.237)	-0.092(.755)
Control for GDP	0.417(.156)	-0.107(.728)
Decommodification	0.339(.236)	0.109(.711)
Control for GDP	0.646(.017)	0.316(.292)
n	14	14

Source: OECD, World Happiness Database, Scruggs and Allan (2006)

Table 1.1 shows how indicators of social policy correlate with indicators of SWB (the significance levels statistics that are reported in the brackets are p-values). I report both bivariate correlations and partial correlations holding GDP constant. None of the bivariate correlations are statistically significant though some of them are fairly strong (above .3). This suggests that small-N is a problem and that being conservative about significant levels risks a Type 2 error. Once the GDP has been controlled for the positive correlations between life-satisfaction and the indicators of welfare state effort is strengthened and the correlation between decommodification and life-satisfaction becomes statistically significant. This is the exact opposite to Veenhoven's findings, as he detected a bivariate association between welfare state effort and well-being that disappeared once national wealth was controlled for. The correlations between happiness and our policy indicators are noticeably smaller and fall, for the most part, far short of statistical significance. In the case of social expenditures the direction of the relationship is in the opposite direction from what it is for life-satisfaction. This is indicative that satisfaction and happiness are different constructs.

Table 1.2: Regression analysis (OLS), regressing indicators of subjective well-being on indicators of welfare state effort and GDP

	Life	
	satisfaction	Happiness
Social expenditures	0.0000529(.156)	0.0000092(.729)
GDP	0.0000449(.024)	0.0000261(.064)
R²	0,454	0,284
Decommodification	0.0247477(.017)	0.0081669(.292)
GDP	0.0000569(.003)	0.0000301(.036)
R²	0,615	0,348
n	14	14

Source: OECD, World Happiness Database, Scruggs and Allan (2006)

In Table 1.2 indicators of subjective well-being have been regressed on the indicators of welfare state effort and GDP. The results mirror those in Table 1.2 except they allow us to assess the "size" of the effects of welfare state effort and GDP relative to each other. The reader should not be alarmed by the smallness of the coefficients for financial indicators (social expenditures and GDP). The reason is that these coefficients cannot be anything but small unbounded variables are being regressed on very compact bounded variables (4-step measures of SWB; see Johns and Ormerod, 2007, p. 33 for a discussion of this issue). The results are mixed depending on which indicators of welfare state effort and subjective well-being are being used. The results suggest that at least some of the association of welfare state effort and well-being is independent of the GDP. This is not in line with Veenhoven's conclusion that the association between welfare state effort and well-being was a spurious function of national wealth.

1.3.2. DISTRIBUTION OF SUBJECTIVE WELL-BEING IN THE CROSS-SECTION

In this section we consider whether welfare state effort has a discernible statistical association with the inequality of SWB. For this analysis I use the standard deviations of subjective well-being scores for the same countries and years as in the preceding section as this is the measure used by Veenhoven (1995 and 2000a).

Veenhoven warns that the validity of using standard deviations in this way depends in large part on there being considerable breadth of distribution and responses not being too unevenly distributed. It is obvious that our 4-point scales violate at least one of these conditions. However, analysis using indicators with a wider range (not reported here) did not produce substantially different results from those that I present here. Consequently I feel comfortable presenting results using 4-step indicators of subjective well-

being as indicative of the association between welfare state effort and inequality of well-being, which allows us to use the standard deviations for the means that were analyzed in last section. However, the reader should take the results with a pinch of salt.

Table 1.3: Bivariate and partial correlations (controlling for GDP) between indicators of welfare state effort and the standard deviations of subjective well-being indicators

	Life satisfaction	Happiness
Social expenditures	-0.0168(.566)	-0.057(.846)
Control for GDP	-0.180(.555)	-0.059(.847)
Decommodification	-0.766(.001)	-0.401(.155)
Control for GDP	-0.742(.004)	-0.502(.080)
n	14	14

Source: OECD, World Happiness Database, Scruggs and Allan (2006)

Table 1.3 shows the bivariate correlations between indicators of welfare state effort and indicators of subjective well-being and their partial correlations controlling for GDP. The correlations of all indicators of welfare state effort and SWB are negative, indicating that welfare state efforts tend to reduce inequality in SWB. Decommodification has a sizable and statistically significant association with life satisfaction while the relationship for social expenditures does not. The correlation between decommodification and inequality of happiness also approaches statistical significance once the GDP is controlled for. Given the size of this correlation and the small number of observations underlying this associations it seems likely that rejecting it on the basis of statistical significance risks a Type 2 error. The relationship between social expenditures and happiness is very weak and falls far short of statistical significance. Again, the different results obtained for the indicators of happiness and life-satisfaction indicate that these do not measure the same underlying construct.

Table 1.4: Regression analysis (OLS), regressing standard deviations of subjective well-being on indicators of welfare state effort and GDP

	Life satisfaction	Happiness
Social expenditures	-0.0000084(.555)	-0.0000015(.847)
GDP	0.0000093(.200)	-0.0000030(.436)
R ²	0,169	0,059
Decommodification	-0.0104167(.004)	-0.0037948(.080)
GDP	0.0000043(.400)	-0.0000049(.177)
R ²	0,613	0,294
n	14	14

Source: OECD, World Happiness Database, Scruggs and Allan (2006)

Table 1.4 repeats the analysis presented in Table 1.3, this time using OLS which allows us to assess the impact of growth and welfare state effort simultaneously. There is no evidence that GDP affects the inequality of happiness or life-satisfaction as such.

In this section and the last we have found that in the cross-section welfare state effort may be associated with higher levels and more equal distributions of life-satisfaction whereas there is less evidence that welfare state effort affects either the level or the distribution of happiness. Judging by the cross-sectional analysis presented in this section and the last, one might be tempted to conclude that Veenhoven's

conclusions regarding the effect of welfare state effort on well-being were wrong. Even with very small samples we have found statistically significant associations between indicators of welfare state effort and the level and the distribution of subjective well-being. Nevertheless, it is sensible to conduct further analysis before concluding this to be the case.

1.4. POOLED TIME-SERIES AND CROSS-SECTIONS

One of the advantages of having repeated observations of a number of variables for a number of units at different points in time is that these can be pooled so as to increase the number of observations for analysis in comparative research. This approach is not as straight forward as it might seem as it comes with problems of its own. Podestá (2002) points to the most important of these:

1. In many cases errors are not independent from one period to the next as observations and traits that characterize them tend to be interdependent across time.
2. Errors tend to be correlated across nations. Countries that are proximate in some sense may be affected by similar processes such that errors in those countries are linked but independent from those of more distant countries.
3. Errors tend to be heteroscedastic.
4. Errors may contain components that reflect effects specific either to units of observation (in our case countries) or periods of observations (years in the present analysis). If these effects are not incorporated into the model they will be caught up and concealed in the error term. This can result in models with observed heteroscedasticity and auto-correlation even if we started with data that was neither. This results from assuming that the intercept of the analytical model is constant for all units of observation and points in time.
5. Causal heterogeneity. Since we assume that the relationship between our dependent and independent variables is homogenous across units of observation and points in time, errors may be non-random across countries and years if parameters are in fact heterogeneous. This results assuming that the slope of the relationship is constant for all countries and years.

There are two methods that are used to deal with at least some of these issues in estimations using TSCS data. The first is the so-called Parks-Kmenta method which uses feasible generalized least squares to overcome problems 1-3 (Parks, 1967). The latter method is Ordinary Least Squares with Panel Corrected Standard Errors (PCSE). This method was proposed by Beck and Katz (1995) after they demonstrated the Parks-Kmenta method to be somewhat flawed. I use the PCSE model for the present analysis, though it should be noted that this method has recently come under critical scrutiny (e.g. Wilson and Butler, 2007).

Before the PCSE model can be estimated the serial autocorrelation of errors must be eliminated. Becks and Katz (1996) have argued that this should be done by including a lagged dependent variable on the right hand side of the equation. Maddala (1998), on the other hand, argues for OLS estimation with panel corrected covariance matrix estimation for models with no lagged dependent variables since OLS estimators with lagged dependent variables are known to be inconsistent in the presence of serial correlation of errors. I adopt the latter approach, assuming that there is first order autocorrelation of errors and that the coefficients of these processes are specific to panels as analysis indicated that this was in fact the case. Missing values were excluded case-wise. The resulting model is referred to as Prais-Winsten regression model. This model both resolves problems 1-3 and can deal with unbalanced panels with imbedded missing values.

I solve one aspect of problem 4 by incorporating fixed effects in the model by including dummy variables for the countries which contribute observations to this analysis. The use of fixed effects requires that we have at observations of each country on at least two points in time and that the value of the dependent variable differs at least at one occasion of measurement (Ferrer-I-Carbonell and Frijters, 2004). Failing that the inclusion of a dummy variable will partial out the “effect” of that country. This leaves us with 14 countries for our analysis.

It should be noted that I make no attempt to control for unobserved time-specific country-invariant processes as doing so would use up most of the degrees of freedom gained by pooling time-series and cross-sections in the first place. It may well turn out that assuming away such time-specific processes may

be unrealistic, though one is hard pressed to think of events that might influence subjective well-being equally in different countries at the same point in time during the period under study.

Finally, I make no attempt to deal with problem five, i.e. causal heterogeneity, for now, as doing so would require the abandonment of constant coefficient models in favor of a very different research strategy suited to answer very different questions from those that preoccupy us here, i.e. whether and to what extent differences in subjective well-being of the inhabitants of different countries can be traced back to the welfare state efforts of those countries. This does not mean that I disregard causal heterogeneity as a potential problem, as will become clear in section 5.

1.4.1. LEVEL AND DISTRIBUTION OF SUBJECTIVE WELL-BEING

Table 1.5 shows the results of the TSCS analysis of whether welfare state effort is associated with average levels of subjective well-being in different countries. As before significance levels are reported in the brackets. The number of observations and the number of countries used for the analysis are reported separately as they do not coincide as they do in cross-sectional analysis. In this analysis any differences between indicators life-satisfaction and happiness are not to be taken as evidence for the difference between the two constructs. This is because while the data is drawn from the same countries for both variable they are not drawn from equal number of time-points nor necessarily from the same points in time.

Table 1.5: Prais-Winsten regression with panel corrected standard errors and fixed effects – indicators of the mean levels of SWB regressed on indicators of welfare state effort and GDP

	Coefficients	n	Countries	R ²
Life satisfaction				
Social expenditures	0.00000480(.678)	213	14	0,9907
GDP	0.00000263(.596)			
Decommodification	-0.0055123(.260)	276	14	0,9997
GDP	0.00000329(.186)			
Happiness				
Social expenditures	0.000018(.015)	54	14	0,9794
GDP	-0.0000002(.934)			
Decommodification	-0.0000671(.993)	59	14	0,9636
GDP	0.00000231(.425)			

Source: OECD, World Happiness Database, Scruggs and Allan (2006)

The results in Table 1.5 are not consistent with the cross-sectional analysis presented in tables 1.1 and 1.2. Firstly, there was evidence that welfare state effort had a positive association with life-satisfaction in the cross-section. This is not the case in the TSCS analysis. Secondly, there was no evidence of an association between welfare state effort and happiness in the cross-section. In the TSCS data there is evidence that social expenditures have a small positive effect on happiness that is statistically significant at the .01 level. On the other hand, there does not appear to be evidence that decommodification is associated with happiness.

Both the cross-sectional and the TSCS analysis provide evidence that welfare state effort is in some way associated with levels of subjective well-being in different countries. Again one might feel tempted to conclude that Veenhoven was wrong if it wasn't for the fact that we reach very different conclusion about which aspects of policy are associated with which aspects of subjective well-being and in what way depending on which method is used.

Table 1.6 (below) reports the results of the TSCS analysis of the relationship between welfare state effort and inequality of SWB. The cross-sectional results showed that welfare state efforts were associated with reduced inequalities of life-satisfaction but that they had no obvious relationship with the inequality of happiness. The TSCS analysis, however, does not suggest that welfare state effort is associated with life-satisfaction in any way whereas social expenditures have a positive association with inequality of

happiness, suggesting that social expenditures lead to higher mean levels of happiness but also more unequal distributions of happiness.

Table 1.6: Prais-Winsten regression with panel corrected standard errors and fixed effects – standard deviations of subjective well-being indicators regressed on indicators of welfare state effort and GDP

	Coefficients	n	Countries	R ²
Life satisfaction				
Social expenditures	0.00000320(.431)	213	14	0,9686
<i>GDP</i>	-0.00000394(.015)			
Decommodification	-0.0003479(.881)	276	14	0,9595
<i>GDP</i>	-0.00000213(.012)			
Happiness				
Social expenditures	0.0000224(.000)	54	14	0,9336
<i>GDP</i>	-0.00000442(.000)			
Decommodification	-0.0020024(.574)	59	14	0,8936
<i>GDP</i>	0.00000266(.145)			

Source: OECD, World Happiness Database, Scruggs and Allan (2006)

There are clear parallels between the analyses presented in tables 1.5 and 1.6. Both cross-section and TSCS analysis provide evidence that that welfare state effort affects both the level and the distribution of subjective well-being. It would be premature, however, to claim that this refutes Veenhoven's claim that people's well-being is not affected by welfare state effort because in both cases the results from the TSCS analyses are inconsistent with the results from the cross-sections with regards to which aspects of policy affect which aspects of subjective well-being and in what way. What are we to make of this inconsistency? Are we to trust the cross-section or the TSCS analysis or do these inconsistencies cast doubts on both?

1.5. CAUSAL HETEROGENEITY

In the TSCS analysis above I made no attempt to resolve the problem of causal heterogeneity, i.e. that the strength of the relationship between our policy indicators and our subjective well-being indicators might vary from year to year or between countries. The reason for this omission was that causal heterogeneity could not be accommodated within the analytical framework employed there. The question is whether there are reasons to believe that such causal heterogeneity may be present in our analysis?

There may be good reasons to expect this to be the case, at least with regards to subjective well-being. Insofar as subjective well-being has a cognitive component, issues of meaning, context and culture have a bearing on how happy and satisfied people are with their lives. This means, in effect, that people's evaluations of their circumstances are not independent of those circumstances. There are therefore reasons to expect that the "effects" of different things on subjective well-being vary across countries and that these variations are related to contextual characteristics such as culture or socio-economic characteristics that also vary across countries.

Testing for causal heterogeneity requires that we abandon the use of constant coefficient models for a research design that allows us to explore the associations between welfare state effort and subjective well-being within countries over time. The model I employ for that analysis is the so-called Seemingly unrelated regression model, which is "interpretable as a series of a nation specific regression analysis that utilize contemporaneous cross-equation error correlations among the error of a system of equations to improve the efficiency of the equation's estimates" (Podestá, 2002, p. 30). This method requires unbroken strings of observation over time. It is only for eight countries that we have data that meets these requirements. The results for those countries are reported in Table 1.7 below. This analysis is also limited to life-satisfaction as long unbroken strings of observations on happiness in countries over time were not available for this analysis. Finally, while I did control for GDP in the analysis reported in Table 1.7 reports

only coefficients for welfare state effort (social expenditures and decommodification) and R² for the models as these are what is relevant to the point I am about to make.

Table 1.7: Seemingly unrelated regression: The relationship between welfare state effort and life-satisfaction in eight countries over time

	Belgium	Italy
Social expenditures	-0.0001233(.098)	-0.0001111(.006)
R ²	0.061	0,821
Decommodification	0.011(.673)	-0.002(.858)
R ²	0.067	0,851
	Denmark	Japan
Social expenditures	0.0000645(.011)	0.0000811(.009)
R ²	0.152	0,405
Decommodification	0.024(.043)	0.024(.014)
R ²	0.108	0,456
	France	Netherlands
Social expenditures	-0.0000879(.019)	-0.0000402(.083)
R ²	0.108	0,075
Decommodification	0.018(.086)	-0.009(.508)
R ²	0.074	0,172
	Ireland	United Kingdom
Constant	-0.0001325(.016)	-0.0000251(.213)
R ²	0.496	0.121
Decommodification	-0.017(.040)	-0.007(.357)
R ²	0.241	0.043

Source: OECD, World Happiness Database, Scruggs and Allan (2006)

The evidence presented in Table 1.7 is indicative of causal heterogeneity. Social expenditures have a statistically significant association with life-satisfaction in five countries of eight. The association is positive for two of those countries and negative for three of them. This means that not only is the relationship not present in all countries but the direction of the relationship varies country from country. Decommodification has a statistically significant association with life-satisfaction in three countries of eight. In two cases the association is positive and in one case it is negative. I will not attempt to explain these differences as any attempt to determine why this is the case would be speculative at this point and well beyond the scope of this analysis. Suffice it to say that this supports the assertion that causal heterogeneity is something we must be mindful of when we use subjective well-being at a high level of aggregation.

Establishing causal heterogeneity between years is somewhat problematic given the limitations of the data. I only have observations covering sufficiently extended periods of time for eight countries. Nevertheless, examining these is at least indicative. We regressed life-satisfaction on each of our indicators of welfare state effort at five year intervals using OLS, controlling for GDP. Each indicator of welfare state effort was examined separately. Since the analysis is based on only 8 countries we would expect some variations between years if only because of measurement errors and such. Consequently it is not clear how to interpret Table 1.8 (below). Nevertheless it is at least indicative that the association between welfare state effort and life-satisfaction is not time-invariant.

Table 1.8: Life-satisfaction regressed on indicators of welfare state effort at five year intervals holding GDP constant.

		Social expenditures	Decommodification
1980	β	0.0005845(.013)	0.0695434(.015)
	R ²	745	728
1985	β	0.0002277(.266)	0.0407458(.145)
	R ²	276	403
1990	β	0.0001753(.105)	0.0443829(.055)
	R ²	440	553
1995	β	0.0001211(.251)	0.0473639(.078)
	R ²	267	504
2000	β	0.0000601(.222)	0.0196336(.374)
	R ²	793	758

Source: OECD, World Happiness Database, Scruggs and Allan (2006)

While the analyses presented in this section are far from conclusive they do suggest that specific national contexts at different points in time intervene in the relationship between welfare state effort and SWB. This has a number of implications for attempts to assess the impact of social policy on subjective well-being from a comparative perspective. These will be discussed in the concluding section.

1.6. SUMMARY AND DISCUSSION

The results in the preceding sections have a number of implications. Firstly, it would seem that the relationship of indicators of welfare state effort with happiness is different from their relationship with life-satisfaction. This implies that happiness and life-satisfaction are not identical constructs and that it may not be valid to use them interchangeably.

Both the cross-section and the TSCS analysis provided evidence that welfare state effort may well be associated with at least some well-being outcomes and that these are independent of material living standards as measured by the GDP, though the association is not always positive. This was also true of the distribution of SWB. It would nevertheless be something of an overstatement to claim to have provided evidence falsifying Veenhoven's conclusion since the results from the TSCS analysis were inconsistent with the results from the cross-section. The measured conclusion is that we simply don't know and that further research is required before we can arrive at conclusive results.

The inconsistency between the cross-section and the TSCS analysis requires an explanation. The obvious explanation is that the TSCS models are misspecified. This is almost certainly the case. On the other hand it is difficult to see how we could arrive at a more accurate specification given the current state of knowledge on the relationship between social policy and people's subjective well-being. Methodologists warn that the complications involved in TSCS analysis are formidable and that one should not attempt to pool time-series without a well informed theory about the relationships one attempts to model (Stimson, 1985). We lack such a theory. To give an indication of the size of the "black box" consider that different countries face different challenges and also choose different solutions to seemingly identical or comparable problems. These then play out in various socio-economic, demographic and cultural settings, aspects of which intervene in the relationship between policy and subjective well-being outcomes. Furthermore, these

contextual differences are greatly obscured by the use of highly aggregate measures used in most comparative research. This is in essence the issue of causal heterogeneity discussed the preceding section.

These problems are not specific to the use of TSCS models. Both TSCS and cross-sectional models suffer from problems arising from drawing statistical inferences from a sample of observations that are not selected at random especially since both strategies are sensitive to which countries we include and at which points in time. Both also suffer from loss of information due to aggregation. In addition the cross-sectional method is completely insensitive to the temporal dimension that creates so many complications for TSCS analysis.

It might seem that I regard the prospect for understanding the relationship between social policy and subjective well-being to be bleak at best. That is not the case. Rather, I would propose that given the current state of knowledge it is not necessarily informative to pursue this issue at an international comparative level. It seems more promising to pursue a strategy of disaggregation. Understanding how specific policies affect different aspects of different people in different contexts is far more likely to lead to concrete policy prescriptions that are sensitive to the problems at hand. As we accumulate knowledge it may be that eventually we can make sense of these issues from an international comparative perspective, though there is much cause for doubt. Nevertheless, it seems plausible that some aspects of subjective well-being may be less sensitive to context than others and even if we will never come up with a comprehensive theory of well-being in the welfare state there is nevertheless a great deal of valuable knowledge to be had in this line of inquiry.

APPENDIX: DATA AVAILABILITY

Table 1.9: Information on the availability of indicators used in this study

Country	Decommodification	Social expenditures per capita fixed	Life-satisfaction	Happiness
Austria	1972-2001	1980,1985,1990-2001	1995-2001	1990 and 1999
Belgium	1972-2001	1980-2001	1973 and 1975-2001	1981, 1990 and 1999
Canada	1972-2001	1980-2001	1989, 1996-7 and 2000	1981, 1985-6, 1989, 1990 and 2000
Denmark	1972-2001	1980-2001	1973 and 1975-2001	1972, 1981, 1990 and 1999
Finland	1972-2001	1980-2001	1995-2001	1972, 1981, 1990, 1996 and 2000
France	1972-2001	1980-2001	1973 and 1975-2001	1981, 1990 and 1999
Germany	1973-2001	1980-2001	1990-2001	1998 and 1999
Ireland	1972-2001	1980-2001	1973 and 1975-2001	1981, 1990-1, 1999 and 2001
Italy	1972-2001	1980-2001	1973 and 1975-2001	1981, 1990-1 and 1999
Japan	1972-2001	1980-2001	1972-1999 and 2001-2001	1981, 1990, 1995 and 2000
Netherlands	1972-2001	1980-2001	1973 and 1975-2001	1981, 1990 and 1999
Sweden	1972-2001	1980-2001	1995-2001	1972, 1981, 1990, 1996, 1999 and 2000
United Kingdom	1972-2001	1980-2001	1973 and 1975-2001	1981, 1990, 1991, and 1998
United States	1972-2001	1980-2002	1991, 1997 and 2001-2	1981, 1990-1, 1995, 1998-9 and 2002

* Information on GDP was available for all years

Table 1.10: The year observations were drawn for each country for the cross-sectional analysis

Country	Year
Austria	1999
Belgium	1999
Canada	2000
Denmark	1999
Finland	2000
France	1999
Germany	1999
Ireland	2001
Italy	1999
Japan	1995
Netherlands	1999
Sweden	2000
United Kingdom	1998
United States	2002

REFERENCES:

- Andersen, J. (2007) "Conceptualizing the welfare state - The 'dependent variable problem' writ large". CCWS working paper no. 2007-51. Aalborg: Center for comparative welfare studies.
- Bambra, C. (2006) 'Decommodification and the worlds of welfare revisited' *Journal of European Social Policy* vol. 16(1), pp. 73-80.
- Beck, N. and Katz, J. (1995) 'What to do (and not to do) with time-series cross-section data' *The American Political Science Review* vol. 89(3), pp. 634-647.
- Beck, N. and Katz, J. (1996) 'Nuisance vs. substance: specifying and estimating time-series-cross-section models' *Political Analysis* vol. 6, pp. 1-36.
- Ní Bhrolcháin, M. (2001) "'Divorce effects' and causality in the social sciences' *European Sociological Review* vol. 17(1), pp. 33-57.
- Easterlin, R. (2005) 'Feeding the illusion of growth and happiness' *Social Indicators Research* vol. 74, pp. 429-443.
- Esping-Andersen, G. (1990) *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- Ferrer-I-Carbonell, A. (2005) 'Income and well-being: an empirical analysis of the comparative income effect' *Journal of Public Economics* vol. 89, pp. 997-1019.
- Ferrer-I-Carbonell, A. and Frijters, P. (2005) 'How important is methodology for the estimates of the determinants of happiness' in *The Economic Journal* vol. 114, pp. 641-659.
- Goldthorpe, J. (2000) *On Sociology*. Oxford: Oxford University Press.
- Johns, J. and Ormerod, P. (2007) *Happiness, Economics and Public Policy*. London: The Institute of Economic Affairs.
- Kangas, O. (1991) 'The bigger the better? - On the dimensions of welfare state development: Social expenditures versus social rights' *Acta Sociologica* vol. 34, pp. 33-34.
- Lister, R. (1995) 'Dilemma in engendering citizenship' *Economy and Society* vol. 24(1), pp. 1-40.
- Maddala, G. (1998) 'Recent Developments in Dynamic Econometric Modelling: A Personal Viewpoint' *Political Analysis* vol. 7, pp. 59-87.
- Marmot, M. (2005) *Status Syndrome*. London: Bloomsbury Publishin.
- O'Connor, J. (1993) 'Gender, class and citizenship in the comparative analysis of welfare state regimes' *The British Journal of Sociology* vol. 44(3), pp. 501-518.
- Parks, R. (1967) 'Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated' *Journal of the American Statistical Association* vol. 62, pp. 500-509.
- Podestá, F. (2002) "Recent developments in quantitative comparative methodology: The case of pooled time-series cross-section analysis" DSS papers no. Soc 3-02, Brescia: Università degli Studi di Brescia.
- Sainsbury, D. (1996) *Gender, equality and welfare states*. Cambridge: Cambridge University Press.
- Scruggs, L. and Allan, J. (2006) 'Welfare-state decommodification in 18 OECD countries: A replication and revision' *Journal of European Social Policy* vol. 16(1), pp. 55-72.
- Stimson, J. (1985) 'Regression in space and time: A statistical essay' *American Journal of Political Science* vol. 29(4), pp. 914-947.
- Sumner, L. (1996) *Welfare, Happiness, and Ethics*. Oxford: Oxford University Press.
- Titmuss, R. (1958) *Essays on the Welfare State*. London: Allen and Unwin.
- van Praag, B. (2007) 'Perspectives from the happiness literature and the role of new instruments for policy analysis' *CESifo Economic Studies* vol. 53(1), pp. 42-68.
- Veenhoven, R. (1995) 'Livability of the welfare state' *Social Indicators Research* vol. 36, pp.1-48.
- Veenhoven, R. (2000a) 'Well-being in the welfare state: level not higher, distribution not more equitable' *Journal of Comparative Policy Analysis* vol. 2(1), pp. 91-125.
- Veenhoven, R. (2000b) 'The four qualities of life' *Journal of Happiness Studies* vol. 1, pp. 1-39.
- Wilson, S., and Butler, D. (2007) 'A lot more to do: The sensitivity of time-series cross-section analysis to simple alternative specifications' *Political Analysis* vol. 15, pp. 101-123.