Nijmegen School of Management
Department of Economics and Business Economics
Master's Thesis Economics (MAN-MTHEC)

Complexity of social protection and wellbeing

By Olli Helanen (s1029568)

Nijmegen, 8 August 2022

Program: Master's Program in Economics
Specialisation: Economics, Behavior, and policy
Supervisor: Dr. F. Bohn

Radboud Universiteit

Abstract: This research considers the role that social protection complexity plays in determining wellbeing at the country level. A panel data regression analysis is conducted with a data set consisting of 35 countries and 12 years. The results of this analysis are mixed, one out of three regressions being significant and complexity having a negative effect on wellbeing. Additionally, the determinants of complexity are investigated, with the finding that the age dependency ratio plays a role in determining the complexity of a welfare system. Lastly, this research investigates whether complexity plays a mediating role between age dependency ratio, and wellbeing, finding a small significant indirect effect through complexity. While this indicates that complexity might play a role in determining wellbeing, more research is needed to infer a causal linkage.

Table of Contents

1. Introduction

"I am not a client, a customer, nor a service user. I am not a shirker, a scrounger, a beggar nor a thief. I am not a national insurance number, nor a blip on a screen. I paid my dues, never a penny short, and was proud to do so. I don't tug the forelock but look my neighbour in the eye. I don't accept or seek charity. My name is Daniel Blake, I am a man, not a dog. As such I demand my rights. I demand you treat me with respect. I, Daniel Blake, am a citizen, nothing more, nothing less "(Loach, 2016). The final line from a 2016 film criticising the social protection system in the UK. This film manages to show the viewer, the frustration and hardship the main characters feel when trying to navigate a cold, increasingly complex, and bureaucratic maze of paperwork and social workers. While somewhat exaggerated, the film captures the feeling people might have when dealing with a highly complex social protection system. While the stated goal of these systems is to offer the necessities for life, they can often be difficult to navigate, leading to people not having access to the support they so desperately need. All at a time when social protection in western societies has become more extensive than ever, leading to many people being fully dependent on it.

Societal benefit of the study:

Social protection influences many aspects of life for numerous people. Thus, it is important to consider the effectiveness of the social protection system, since a more efficient social protection system is better for everyone inside the country, as less resources go to waste, and it is able to serve people better. Furthermore, the people who are often most dependent on welfare are those in our societies who are the most vulnerable. And even if governments act opportunistically, it is still in their interest to have a functioning social protection system, as it may improve their chances in staying in power. Taydas & Peksen, (2012) conducted a cross-sectional study, using 158 countries and 30 years. They identified that investment into social support systems can lower the chance of civil conflict. This result can be interpreted from a policy maker/ politicians' point of view to indicate that investment into the populous is beneficial.

Academic contribution

The phenomenon of complexity of the social system is not widely studied. The studies which do look at a type of complexity tend to be case studies relating to the reformation of the social protection system (Pak, 2020). While these studies can offer important insight in how a simplification of the system in a single country affects the wellbeing of its citizens, it does not provide us with the underlying mechanisms which contribute to the relationship between complexity and wellbeing. Other papers such as Niblett, Begg, & Mushövel, (2015) discuss what drives the call for reform of a social protection system, identifying aging population demographic change, and technological advancements as key drivers. However, it remains unclear what contributes to complexity and what effect does it have on wellbeing This is the research gap this master's thesis tries to fill, contributing to the current literature concerning social protection.

With this motivation in mind this master's thesis will consider the following research question:

Does the complexity of the social protection system affect the wellbeing in European countries? And what are the drivers of this relationship?

This is done by conducting panel data analysis, looking at the relationship between complexity and wellbeing. As well as the underlying factors which drive a social protection system in becoming more complex. This is done using a data set which comprises 35 European countries, over a period of 12 years. The reason for using European countries is, firstly, that European countries have a highly developed social protection system. Niblett, Begg, & Mushövel, (2015) outlines the current welfare system in Europe, identifying that EU social protection made up 39.6% of all spending on social protection in 2015. This goes to show that European countries tend to have more extensive social protection systems on the global scale. While there are other western countries with similar and in some cases higher spending per capita, (United States and Japan, being notable ones). Secondly, there is a high degree of data available for the chosen countries, for each of the key variables. (For a complete list of countries see appendix 1)

2. Literature Review & Hypothesis

2.1 Happiness economics

Happiness economics is a growing field of economics which focuses on studying wellbeing. Identifying certain things, contributing to the wider field of economics. An example of these contributions is the idea of the Esterling paradox. Noting that after a certain point, increased income does not seem to contribute much to the happiness of a country. A common measurement tool used in happiness economics is "Subjective wellbeing". Diener et al, (2003) outlines the development and history of subjective wellbeing, as well as gives us the following definition. "The field of subjective well-being (SWB) comprises the scientific analysis of how people evaluate their lives—both at the moment and for longer periods, such as for the past year. These evaluations include people's emotional reactions to events, their moods, and judgments they form about their life satisfaction, fulfilment, and satisfaction with domains such as marriage and work. (Diener et al, 2003, pp. 404). The key take away from this is that SWB can be used to study not only the fluctuations in the emotions of people, but also the long-term life satisfaction of people. This new approach serves as a alternative approach to measuring utility in economics. Ferrer-i-Carbonell, (2012) states that this new measure can be useful for public policy recommendations, as happiness economics can point to empirical results, in order to make policy recommendations. However, as with many with a lot of new measures, SWB has also come under some criticism. Johns & Ormerod (2008) make the argument that happiness is a discrete measure, so you need a large number of people changing from one category to another, in order to see movements within the aggregate. Thus, they argue that the evidence from happiness economics is not robust enough to be used in policy making.

2.2 Social protection

Social protection systems have become more and more common around the world. This is especially true in developed countries. According to International Labour Organization, (2021). most countries have some form of social protection in place. And naturally there has been a large body of literature dedicated to social protection. Vinci Gassmann, & Mohnen, (2022) gives an overview of literature concerning the goals of the social protection system. While the exact goals of systems are defined differently depending on the study, a common theme across literature seems to be that social protection aims to alleviate poverty, provide income security, and address the causes of poverty. Furthermore, while not explicitly the goal for most social protection programmes, some social protection has a redistributive effect. (Wang, Caminada, & Goudswaard, 2012). However, the degree to which social protection systems have redistribution effects varies. Esping-Andersen & Myles, (2011) identified that while welfare has redistributive effects, the degree of which varies from state to state. The Nordic countries have systems with a large degree of redistribution, while the systems in Anglo-Saxon and, to a lesser degree, Southern-European countries, have less of a redistributive streak to them. This indicates that the design of the welfare system matters and is intrinsically linked to the culture and the goals of societies. The key take away is that social protection systems differ per cultures and in order to understand them, cultural differences need to be taken into account, in our research design.

There is also evidence that social protection may help long term growth and efficiency. Dercon, (2011) argued that if social protection fixes existing market failures, it can contribute to growth in the long run. However, there is no consensus on whether social protection is good for growth. As there is also literature which found that social protection programmes have a negative effect on growth (Arjona et al., 2003). While the exact goals of welfare systems differ, the exact effects on things such as growth are not clear. It should be fair to say that most systems would value a more efficient system, as (ignoring things such as governmental corruption), this would benefit all groups in society. Furthermore, Estevez-Abe et al, (2001) conducted a study on the role social protection played in skill development, finding that unemployed people only develop firm/ industry specific skills when they have security in their employment and/or income, arguing that social protection can offer this security. This shows that social protection can open learning opportunities for the lower classes, which in turn can help them find employment opportunities in the future. However, this is contingent on the programme being well designed.

While there is a lot of research dedicated to social protection systems, there seems to be a cap in the literature when it comes to researching the effectiveness of a social protection system, especially when it comes to how it affects welfare outcomes. Studies have been conducted on subjective wellbeing (SWB) and how it relates to social protection systems, with the consensus being that higher social protection has an association of higher SWB. Kolev & Tassot, (2016) conducted research using data from 150 countries and 1000 individuals, identifying that while the social protection system overall had a positive effect on SWB across all indicators of SWB. However, this relationship did not hold when it came to the top 60% of the society. Pak, (2020) looked at pension reform in South-Korea and identified that it improved the SWB of the population. While this study only looked at one system, it still indicates that the proper functioning of different social protection systems could be crucial for welfare outcomes.

2.3 Complexity

Having a sufficiently simple social protection system, can allow for a better allocation of social support to those who actually need it, and prevents resources being wasted on a system which is over bloated. This is in the interest of the public, as research has shown that a well-functioning social protection system can contribute to improving opportunities for the lower classes (Estevez-Abe et al.,2001). This can ensure that everyone's abilities are best utilized in society, which could possibly lead to long-term growth. While there have been some previous studies on the reform of the social protection system (Pak., 2020) finding that reforming the pension system leads to a higher degree of wellbeing. Niblett, Begg, & Mushövel, (2015) identified the key drivers of reform, noting globalization, technological change, and changes in the societal/ labor force structures. There is no literature explicitly looking into the complexity of the social protection system.

2.4 Hypothesis

This brings to us to the first hypothesis that this research will consider:

H1: "The complexity of the social protection system has a negative effect on the wellbeing of a society"

And the corresponding null hypothesis to this will be

H0 "The complexity of a social protection systems does not have an effect on the wellbeing on societal level."

Schmitt, (2019) identified that an increase in dependency ratio (ADR) has a positive effect on the people who are covered by social protection. Athanasenas et al., (2015) found similar results with dependency ratio (ADR) having a positive and significant effect on social expenditure, across all their models. As the social protection system must expand, it is likely that absent full reform, the complexity of this system would go up as well, leading to a positive effect between ADR and the complexity of the social protection system. It would make sense that similar mechanisms would hold for participation rate as well, meaning that an increased participation rate would lead to a decreased need to spend on social protection, and some programs being cut, thus simplifying the system. Leading to a negative effect on complexity. Motivated by this, the second hypothesis this thesis will consider is:

` H2: *"Age dependency ratio and Participation rate, have a positive effect on the complexity of the social protection system".*

If both H1 and H2 hold up, there is a large likelihood that complexity acts as a mediating variable between ADR, participation rate and wellbeing. Both ADR and participation rate will likely have both a direct and indirect effect, through complexity on wellbeing. Previous literature has identified that ADR and participation rate, both have an effect on the wellbeing of a society (Tang, Xu, Ma, & Gao, 2021) (Mencarini & Sironi, 2010) (Shah et al., 2007) (Wu et al., 2013). While these studies did not test whether complexity played a mediating role in this relationship, it could very well be the case, in case complexity acts as a mediator. Thus, the third hypothesis is as follows

H3: Complexity acts as a mediating variable between age dependency ratio, the participation rate and wellbeing.

3. Methodology and data:

This thesis will be conducted quantitatively. This is done by using the statistical program Stata 17. A panel data analysis will be used. This has the benefit of increasing the number of observations and allowing for observation of the effects across both different countries and different years. Furthermore, using a panel data set up allows for the control of missing variable bias, by means of taking into account the unobserved variables, which stay constant in countries over time. This could include cultural factors as well as other unobserved phenomena that are specific to countries or regions. This can be either done by using country fixed or random effects.

3.1 Key variables

3.11 Subjective wellbeing (SWB)

The goal of this research is to investigate how differently complex social support systems affect the wellbeing of people. In order to do this, Subjective wellbeing (SWB) will be used. Stutzer & Frey, (2010) outline recent developments in happiness economics. Taking the idea of SWB from psychology. Furthermore, they state that SWB offers a new approach to tackling old problems, and that it has many possible applications. On the other hand, Diener, et al (2003), points out that SWB can be dependent on cultural factors as well. This will be taken into account by conducting random/fixed effects analysis. The data for SWB is gathered from the World Happiness Report, (2019), as the years used are from 2008 until 2019, from 35 European countries. However, measuring welfare is not that straightforward, as different measures for wellbeing include different aspects of life, and direct survey methods have their limitations. In order to combat this, additional robustness checks will be used. This will be done by using different measures of wellbeing and inspecting whether the relationship between complexity and wellbeing holds.

3.12 Positive affect & Human development index (HDI)

This thesis will use two alternative dependent variables as a robustness check. The first of these is “Positive affect": is simple survey data gathered by World Happiness Report, (2019), consisting of them surveying people on how frequently they felt a positive emotion, such as happiness or enjoyment, per day. This is a narrower view of wellbeing when compared to the other two dependent variables. However, while it does not offer a holistic view of wellbeing, the narrower happiness is something which highly affects people’s everyday life, and thus should not be discounted. The second alternative dependent variable is the human development index (HDI). While positive affect looks at wellbeing from a view of pure happiness, HDI takes a more developmental view on wellbeing. This data is available from the United nation human development programmes. The Human development index considers 3 different dimensions of wellbeing, health, knowledge, and standard of living. (United Nations Development Programme, 2022)

3.13 Complexity:

The independent variable this research will be looking at is the complexity of the social protection system. This is operationalized, by dividing the administrative costs of the social protection systems and dividing them with the total spending of the system. A higher degree of complexity means that more is spent on administration relative to the total spending. This is done because, naturally, more extensive welfare systems have higher spending on welfare, and as this research wants to study the complexity of the welfare system relative to the size of the system. Administrative costs for the welfare system as well as the total social protection spending are available from Eurostat. (Database - Eurostat, 2019)

3.14 Age dependency ratio (ADR)

Age dependency ratio (ADR) refers to the number of people who are working compared to the ones who are not working. Thus, a higher dependency ratio would mean more people use some form of social protection, and fewer people are paying income taxes. Previous research has found that ADR influences different wellbeing indicators. However, the results are somewhat mixed. Tang et al., (2021) looked at the dependency ratio and its impact on subjective wellbeing in China. And found a positive relationship between the two. Their reasoning being that having more children (I.E higher dependency ratio), gives rural parents a stronger sense of security in the future, as they know someone will take care of them. It is important to consider that they used a dependency ratio in general, both 0-14 and 65+, and that the study was conducted at a household level. Other studies looked at the dependency ratio of old people to working aged people and found a positive correlation between dependency ratio and (elderly) suicide rates (Shah et al., 2007) (Wu et al., 2013). The data for the ADR is gathered from (World Bank.org, 2019)

3.15 GDP per capita

GDP per capita, affects many aspects of the lives of people, so it is by no means a stretch to assume that it will have an effect on the wellbeing of people. This can work through various avenues, as GDP can dictate the opportunities people have access to, and what types of consumption opportunities people have. However, the link between wellbeing and income is somewhat of a tenuous one. A large body of research has been devoted to the Easterling paradox. First discussed in Easterling, (1974), they found positive relationship between income and happiness within country analysis, while they found much weaker or even insignificant relationships, when conducting comparisons among countries, and over time. While these considerations are important to keep in mind, it is clear that income is likely to have at least a minor effect, and thus it will be incorporated into the regression analysis. This thesis will operationalize income by using GDP per capita, as the analysis is conducted at the country level, and GDP per capita is a decent measure for the income level of households in the country. This data was gathered from “Database - Eurostat,” (2019)

3.16 Labor force participation rate

Having access to decent work tends to be seen as one of the major aspects contributing to a healthy and happy life. A higher degree of labor participation may indicate a high degree of opportunities existing in the country for a large part of the population. Looking at how it affects wellbeing, some studies have looked at the link women's labor participation rate plays in the happiness of women. Mencarini & Sironi, (2010) identified a highly significant and positive relationship between labor participation and women's happiness. Beja, (2013) found similar results. However, they argue this is more of a proxy for the openness of the society. While these studies mostly look at women and their happiness, this effect of labor force participation rate and happiness is expected to hold with the rest of society as well. Intuitively, if a large part of the society is working, it indicates that they are taking part in the society. Whereas if a large part of the working-age population is not working, it is likely that they are somewhat isolated from the society, which may lead to unhappiness. While this effect might be larger, with certain groups, due to systematic exclusion from the labor market, this is outside of the scope of this study. Thus, this thesis will look at labor participation as a whole. This is operationalized by taking the labor force participation rate of the working age population (15-64), (World Bank.org, 2021.)

3.2 Models considered

In order to test the first hypothesis following regressions will be used:

1. SWB_{it}=β_0+ β_1 *$Complexity_{it}$ + ADR_{it} + GDP_{it}+ $Participation\ rate_{it}$ + U_{it}

2. POS_{it}=β_0+ β_1 *$Complexity_{it}$ + ADR_{it} + GDP_{it} $Participation\ rate_{it}$+U_{it}

3. HDI_Score_{it}=β_0+ β_1 *$Complexity_{it}$ + ADR_{it} + GDP_{it} + $Participation\ rate_{it}$+U_{it}

The difference in these three regressions is the way they operationalize wellbeing. Reg 1 uses subjective wellbeing, which is the most holistic measure of wellbeing, taking into account different aspects of wellbeing.

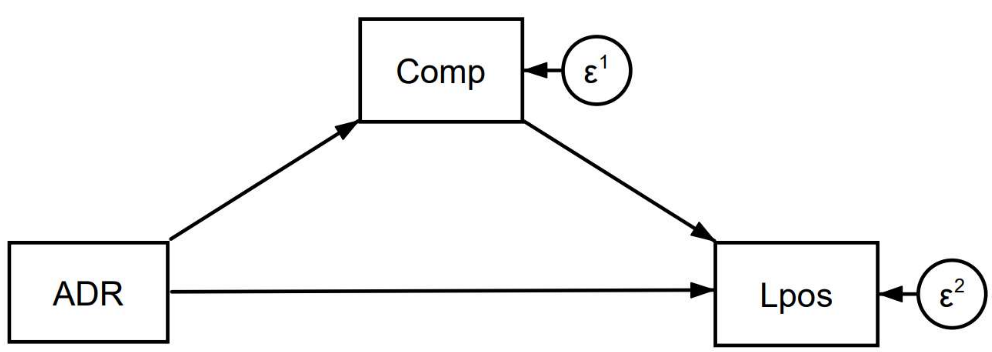
Reg 2 on the other hand uses positive affect, which is more of a simple measure of happiness, simply taking into account the number of times the person felt emotions associated with happiness in the last days. Lastly, we have the operationalization using HDI score. This type of wellbeing is more strongly correlated with development. While all of these measures have their differences, they, all represent different aspects of wellbeing. The idea behind running all three is to get as good of an understanding of wellbeing as possible, as well as increasing the internal validity of the research

In order to test the second hypothesis, the following OLS regression model is used

4. $Complexity_t = \beta_0 + \beta_1 * ADR_t + PartRate_t + e_{it}$

The second hypothesis considers the determinants of complexity, I.E what factors contribute to a social protection system becoming more complex. In order to test this model a normal OLS regression is ran between ADR , participation rate and complexity. It is expected that ADR will have a have a positive effect on complexity, while participation rate will have a negative effect on complexity.

In Order to test the third hypothesis, the following SEM model is used:



The third hypothesis will be tested by using structural equation modelling (SEM) in Stata. This will separate the effect of ADR and participation rate will have, into a direct effect, directly effecting happiness, and an indirect effect, effecting happiness through complexity.

3.3 Checking the assumptions of OLS

3.31 Heteroskedasticity:

OLS assumption 5, states that the error terms should have constant variance, I.E be homoscedastic. This is done due to heteroskedastic standard errors causing the estimations to be biased. This is done for all of the models considered. Heteroskedasticity is checked by conducting the Preusch pagan test, results of which are shown in Appendix 2 A, B, C and D. Looking at Appendix 1a, we have the result of the Preusch pagan test for model 1. the P-value is 0.0001 indicating that we must reject the null hypothesis of constant variance in the error terms. Meaning that heteroskedasticity is likely present. As such, assumption 5 is violated and the estimations may be biased. This will be dealt with by running the regression analysis using robust standard errors. Appendix 1b shows the result of the heteroskedasticity test for model 2. The resulting P-value is 0.1509, thus we cannot reject the null hypothesis of homoskedasticity, and robust standard errors are not needed. Appendix 2c and 2d show the results of models 3 and 4. Both resulting in P-values of lower than the 95th confidence level. P-values of 0.0435 and 0.000 respectively. Hence, both of these regressions will be ran using robust standard errors.

3.32 Multicollinearity

OLS assumption 6 states that there should be “no perfect multicollinearity”. In order to check for these assumptions, a variance inflation test (VIF) is conducted. Models 1-3 all use the same independent and control variables, thus it is only necessary to check one of these models for multicollinearity. The result of this test can be found in Appendix 3a. With a mean VIF score of 1.263, perfect multicollinearity will not be a problem, as only scores above 5 might have a high degree of multicollinearity. On the other hand, we have model 4. Running the VIF test on it gives us results presented in appendix 3b. With the resulting mean Vif of 1.068, there is no need to worry about perfect multicollinearity.

3.3 Normal distribution of the error terms

OLS assumption 7 states that error terms of the regression should follow a normal distribution. The variables were transformed so they followed a normal distribution to ensure a normal distribution of the error terms. The only variable which did not follow a normal distribution was GDP per capita, thus a log version of that variable was taken. In order to test for the distribution of the error terms, an additional variable is constructed out of the residuals of the main regression (Reg 1). And the distribution of it is tested. Firstly, a numerical test is conducted, in the form of Skewness and Kurtosis tests for normality. The result of which can be seen in appendix 4a-4d. From this table we see that while there is a small amount of skewness, we fail to reject the null hypothesis of non-normal distribution. Thus, we can assume that the distribution of the error terms is normal. For both of the models 2 and 3 we must reject the null hypothesis of normally distributed standard errors. For model 4 we have the following results. We see that prop chi2 is 0.087 which is higher than our confidence level, thus we cannot reject the null hypothesis of normal distribution. Thus, for model 4 the OLS assumption 7 is not violated.

Additionally, a graphical analysis was conducted on each of the models. This was done by taking the residuals and plotting them on a histogram and comparing them to the normal distribution. These results can be found in appendixes 5a – d. Looking at 5a we have the result for model 1, we can see that it follows that it follows very closely to a normal distribution. For model 2 (appendix 5b) also seems to follow mostly a normal distribution, while slightly skewed to the right. Model 3 has a similar issue. It is likely that this issue causes the numerical measures to give a significant result. In order to deal with this, the dependent variables positive affect, and HDI score are transformed by taking their log values, which normalizes their distribution. Lastly, looking at appendix 5d we see that the distribution follows normal distribution.

3.4 Dataset:

The dataset which was gathered for this research, consists of 35 European countries, and 12 years, from 2008 to 2019. This dataset is not perfectly balanced, I.E there are some missing observations, with a total of 381. However, there is no indication that the missing variables would have an underlying reason for them, hence this research assumes that the missing's are completely at random. As most of the missing stem from the variable SWB and are likely due to a survey not being conducted in that particular country for that year. A variety of databases was used, see below for a complete breakdown of where the data was gathered.

3.41 Possible problem with unbalanced panel

Wooldridge, (2016) Defines unbalanced panel as a panel data set where all of the units do not appear for each time period. A distinction is made with data missing at random and missing at completely random (MCAR). Missing at random means there may be an underlying reason for the missing, which is correlated with the explanatory variable, but not with the error terms. Whertheas with MCAR there is no underlying reason for the missing. It is stated that with both fixed and random effects models it is possible to use an unbalanced panel, however there are certain considerations we must keep in mind. For random effects, if the time averages are properly defined, by the means of random effects, there should not be a problem with unbalanced dataset for the analysis in this thesis. Thus, we are able to use both fixed and random effects estimation, as the statistical program Stata 17 takes this into account, and thus further transformation of the data set is not needed. And as we have no reason to believe the missing would be due to any underlying reason, we are good to continue

3.5: Fixed or random effects.

While conducting panel data analysis, we must keep in mind we have two different dimensions we are looking at, cross sectional, in this case different countries, and the time dimension. This makes it possible to control check for unobserved differences between countries, which stay constant over time, without the need to add variables to each possible aspect which could play a role. (Wooldridge, 2016) Random effects model is systematically more efficient than fixed effects model. However, if the difference in coefficients is systematic, we cannot use the random effects model, but must use the fixed effects model. (Wooldridge, 2016) In order to test whether we can use random effects, Hausman test must be conducted. With the following results (Table 1)

Table 1- Hausman (1978) specification test (Model 1)

	Coef.
Chi-square test value	8.439
P-value	.077

The P-value is over 0.05, thus we cannot reject the null hypothesis, of differences in coefficient are nonsystematic. In turn, this means that we can use random effects estimates. For the main regression. For the robustness checks using alternative dependent variables, the results are as follows, reg 2- p = 0.0000, thus we need to use fixed effects estimations. For reg 3 p=0.98 indicating that we can use random effects. The tables with the other two Hausman tests can be found in the tables 2 and 3 respectively. While for regressions 1&3 random effects can be used, the p values are quite on the border, hence the regressions will also be ran using fixed effects estimation.

Table 2- Hausman (1978) specification test (Model 2)

	Coef.
Chi-square test value	75.427
P-value	0

Table 3: Hausman (1978) specification test (Model 3)

	Coef.
Chi-square test value	7.823
P-value	.098

3.6 Further methodological considerations

There may be other aspects we must consider when it comes to methodology. First of these is whether there could be endogeneity or simultaneity present within the model, I.E the dependent variable (Wellbeing) might have an effect on the independent variable (Complexity). However, in this particular research, it is highly unlikely that the wellbeing of a society would have a direct effect on the complexity of the social protection, as changing the social protection is a long process, and a sudden drop in the happiness would not cause a revamp of the social protection system. Lastly, we will be considering a confidence level of 95%, this is quite usual in the field, as it minimizes the possibility of both type 1 and type 2 errors.

4. Results & Discussion:

Correlation between the dependent variables is checked via Pairwise correlation test presented in Table 4. As we see, all three variables have a high degree of correlation with each other.

Table 4: Pairwise correlations dependent variables

Variables	(1)	(2)	(3)
(1) SWB	1.000		
(2) LPos	0.791	1.000	
(3) HDI	0.867	0.770	1.000

In order to test the first hypothesis, the first 3 regressions are run. The first one uses subjective wellbeing (SWB) as a measure of wellbeing, the second one uses positive affect, and the third one uses Human development index score (HDI) as a measure of wellbeing. An independent variable for all of the regression is the complexity of the social protection system. And all the regressions use age dependency ratio (ADR), participation rate, and log GDP per capita as control variables. The results of these regressions are presented in table 5, and the full regression can be found in appendixes 6a – 6c.

Table 5: Regression models 1,2,3

	(1) SWB	(2) Lpos	(3) HDI
Comp	-.016	-.011***	.002
	(.024)	(.004)	(.002)
ADR	.014	.007***	.002**
	(.01)	(.002)	(.001)
PartRate	.001**	0**	0
	(.001)	(0)	(0)
LogGDPpercapita	1.682***	-.107*	.111***
	(.269)	(.061)	(.028)
_cons	-12.154***	.465	-1.427***
	(2.73)	(.55)	(.265)
Observations	381	381	381
R^2	0.787	0.047	0.821

Robust standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

Complexity is only significant in the 95% confidence level in one of the three regressions, model 2. For model 1, only participation rate and GDP per capita have an effect on wellbeing. The coefficients being as expected, a positive effect of both participation rate and GDP per capita on SWB. While only 2 out of 4 variables are significant in model 1, it still has a considerably high r-squared of 0.7,89, indicating that around 79 percent of the variation in SWB is explained by the four variables included.

For the second regression (model 2), complexity, ADR, and participation rate are all significant on the 95% confidence level, and GDP per capita is significant on the 90% confidence level. Despite this, the r-squared is surprisingly low at 0.047%. Complexity has a negative coefficient of 0.011, indicating that if complexity increases by 1, positive affect is expected to drop by 0.011. This is in line with the first hypothesis. However, there are 2 coefficients which are not as expected. ADR having a small, but positive effect. Furthermore, while log GDP per capita is not significant at the 95% confidence level, it does have a negative coefficient, which is not as expected. The third regression (model 3), complexity is yet again insignificant. In fact, only the participation rate and Log GDP per capita seem to have an effect, they both have a significant and positive effect. Due to the possibility of a mediation effect, all of the regressions are ran with only one of the key variables (Complexity, ADR, participation rate) giving us the following result (Table 6)

Table 6 – (Regressiong 1,2,3 nested)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	SWB	SWB	SWB	Lpos	Lpos	Lpos	HDI	HDI	HDI
Comp	-.024			-.015***			.001		
	(.025)			(.004)			(.002)		
LogGDPpercapita	2.025***	1.809***	1.882***	.058	.015	-.119**	.144***	.147***	.119***
	(.288)	(.29)	(.293)	(.041)	(.042)	(.054)	(.015)	(.018)	(.032)
PartRate		.001*			0			0	
		(.001)			(0)			(0)	
ADR			.01			.008***			.002*
			(.011)			(.002)			(.001)
_cons	-14.724***	-12.931***	-13.934***	-.805*	-.533	.516	-1.65***	-1.674***	-1.469***
	(3.044)	(2.934)	(2.951)	(.423)	(.435)	(.507)	(.169)	(.19)	(.294)
Observations	381	381	381	381	381	381	381	381	381
R^2	.7265	.7765	.7328	.2493	.4833	.3314	.7879	.7722	.8178

Robust standard errors are in parentheses

*** $p<.01$, ** $p<.05$, * $p<.1$

There does not seem to be any major changes from the original. The most striking difference can be seen in regression 6, where ADR is now significant at the 99% confidence level. This somewhat indicates that there might be a mediating effect where complexity acts as a mediator, between ADR and Positive affect. For the robustness checks, we see that for Positive affect, complexity has a highly significant effect, with a small negative coefficient of -0.008. The interpretation of this is that a 1 point increase in complexity would lead to a 0.008 decrease in people's wellbeing. This is in line with the first hypothesis and gives some support to it. Both Log GDP per capita and participation rate still seem to be highly significant, both having a positive effect on positive affect. If we look at the 2nd robustness check using Countries Human Development Index score as the stand in for wellbeing, complexity does not have a significant effect on wellbeing. However, the age dependency ratio suddenly becomes highly significant, and log GDP per capita remains significant.

Overall, only for 1 out of 3, regression did we find a significant effect from an increase in complexity on the wellbeing on a national level, thus we fail to reject the first Null hypothesis of complexity of the social protection system not having an effect. The regression where complexity was significant used the variable “positive affect” as the dependent variable. If we compare this to the main dependent variable “SWB” the difference is that SWB considers both positive and negative feelings the respondents. It could be that complexity has a stronger effect on the positive feelings, whereas a simpler system does not necessarily reduce the negative feelings people feel, leading to an insignificant effect on SWB. As SWB is more holistic measure for wellbeing, it naturally weights more, and thus we cannot reject the null hypothesis. Furthermore, the third dependent variable Human development index (HDI) is insignificant, and the direction is opposite to what was hypothesized. This reversal of the relationship could only be explained by the fact that only a highly developed countries, could have a complex social protection system. However, as the regression is insignificant, we cannot be sure that there even is a relationship.

Models 1 and 3, could be run using random effects, according to the Hausman test. However, as an added measure of robustness check, those two models will also be ran using fixed effects. With the following results (Table 7)

Table 7: Models 1 & 3 (Fixed Effects)

	(1) SWB	(2) HDI
Comp	-.027	.002
	(.025)	(.001)
ADR	.003	.002**
	(.015)	(.001)
PartRate	.001	0
	(.001)	(0)
LogGDPpercapita	2.105***	.09***
	(.703)	(.032)
_cons	-15.837**	-.181
	(6.673)	(.3)
Observations	381	381
R-squared	.364	.71

Robust standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

If we compare Tables 6 and 7 the only striking difference is that participation rate is no longer significant, other variables do not change in a significant manner. For regression 3 there is no major change between using fixed or random effects estimations. The full regression tables can be found in appendix 7a and 7b

As for the second hypothesis, considering the determinants of complexity, the following regression model was ran, and the results are found in Table 8:

$$Complexity_{it} = \beta_0 + \beta_1 * ADR_t + PartRate_t + e_t$$

Table 8: Regression (Model 4)

Comp	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ADR	-.045	.017	-2.65	.008	-.078	-.012	***
PartRate	.001	.001	1.56	.119	0	.003	
Constant	11.994	.781	15.35	0	10.458	13.531	***

Mean dependent var	9.971	SD dependent var	1.434
R-squared	0.022	Number of obs	381
F-test	3.675	Prob > F	0.026
Akaike crit. (AIC)	1352.241	Bayesian crit. (BIC)	1364.070

**** p<.01, ** p<.05, * p<.1*

Age dependency ratio (ADR) seems to have a negative coefficient, indicating a negative relationship between it and complexity. On the other hand, the participation rate does not seem to have an influence on complexity, with it not even being significant on the 90% confidence level. ADR having a coefficient of -.046, a 1 increase in ADR ratio, causes the complexity of the social protection system to drop by 0.046. This goes against the 2nd hypothesis, as the expected direction was positive.

This can be explained by the idea that countries with high ADR, or more accurately, high projected future ADR, are more willing to reform their social protection systems. This is due to there being a future need for a less complex system in order to ensure the proper functioning of the system. Intuitively, if policy makers see that population projections show that there will likely be a more dependent population, their only options are to either cut the amount spent on people, or try and reform the system, reducing complexity, and thus the overhead. However, this explanation is contingent on policy makers being forward looking, which may not be the case, due to them having an incentive to improve people's lives now, in order to ensure reelection. And as reforming the social protection system may damage their popularity in the following election, which in turn may lead to them losing power. Another possible explanation is that, as people get older, they are more dependent on their retirement, which, by its function, is simpler than for example, unemployment, and thus the complexity of the system is reduced. This is a far more likely reason for ADR having a negative effect on complexity, although future research is needed in order to infer the exact mechanisms of this relationship. However, we must keep in mind the incredibly low R-squared of 0.021, meaning that only around 2% of the variation within the dependent variable is explained. This is a likely indicator that this regression suffers from a high degree of omitted variable bias. Adding log GDP per capita to the regression, increases the r-squared to around 0,029. Log GDP per capita also being highly significant with a coefficient of 2.156. (See appendix 8 for full regression.)

The third hypothesis considers mediation effect through complexity. For the mediation effect to exist, it would be imperative that complexity actually has an effect on the dependent variable. Hence, in order to test this hypothesis, the mediation analysis will be done using the variable positive affect, as it was the only significant one. Furthermore, as only the age dependency ratio (ADR) was significant in complexity, on the 95% confidence level, mediation analysis would only use that variable. The results of the mediation analysis can be found in Table 9.

Table 9 – Indirect effect of ADR through complexity.

Indirect effects

	Coefficient	std. err	z	P>z	[95%	Conf.	interval]
Lpos	0						
Compexity	0						
ADR	-0.001	0.003	-2.02	0.044	- 0.000	-0.000	-.0000202

Running this analysis shows that while there is an indirect effect through complexity, the effect is very small, as the coefficient is -0.001, while we see that it is just significant in the 95% confidence level. Furthermore, we can look at the goodness of fit (table 10)

Table:10 - Goodness of fit of SEM analysis

Equation-level goodness of fit

	Variance					
Dependent variable	Fitted	Predicted	Residual	R-squared	mc	mc2
Complexity	115.661	0.025	115.636	.0002	.0148	.0002
Lpos	.0227	0.002	.02112	.0711	.2666	.0711

mc = Correlation between dependent variable and its prediction.

mc2 = mc^2 is the Bentlerâ€“Raykov squared multiple correlation coefficient.

The r-squared for complexity is only 0.0002, meaning that ADR only explains 0.02% of the variations in complexity. While the r-squared is slightly higher for positive affect (Lpos) at 0.71, it means the direct and indirect effect of ADR only explains around 7 percent in variation of Lpos. Thus, while there might be small negative mediation effect through complexity, it does not explain much of the variation of positive affect and is almost negligible. This, together with the fact that complexity was only significant in 1 out of the 3 initial regressions, points to the direction that it is unlikely that complexity would play a key mediating role in determining wellbeing.

5. Conclusions:

Conducting wellbeing research always has its problems. This mainly stems from there not being a clear consensus on which aspects of wellbeing should be looked at. As Johns & Oremerod (2008) points out, subjective wellbeing (SWB) has its problems as a measure, and this holds true for many other measures of wellbeing. This thesis tries to control this by using three different variables of wellbeing. While this is not a perfect solution to the problem of modelling wellbeing, it should give a more holistic look into wellbeing. For the first hypothesis, the research found only one significant result, with the variable “complexity” only being significant and negative when it came to the dependent variable “positive affect”. With this result, we fail to reject the null hypothesis of there not being an effect on wellbeing, at least if we measure wellbeing with SWB or Human development index (HDI). On the other hand, the second hypothesis provided significant results, with ADR having a significant and negative effect on complexity. This goes against the 2nd hypothesis which assumed that an increase in ADR would lead to an increase in complexity. However, these results still show that the ADR plays a significant role in determining the complexity of the social protection system. As for the third hypothesis, the mediation analysis was run only with regression 2, as it was the only regression where complexity was significant. The analysis indicated a small significant indirect effect of ADR on positive affect through complexity. This indicates that complexity might play somewhat of a mediating role when it comes to wellbeing. However, the evidence for this is not clear, and further research is needed.

5.1 Limitations

While this research tries to implement robustness checks and tests to ensure a high degree of both internal and external validity, there are some limiting factors in this research. First, one of these limitations is related to internal validity. The measure for complexity might be somewhat limited. A more extensive measure for complexity could have been used. Going over each individual social protection system in detail, would have established a more holistic view of the relationship. However, this would have limited the generalizability of the results, I.E the external validity. Because the results would only have been applicable to countries with the same social protection system, or a system which would be remarkably similar. Furthermore, it would have been outside of the scope of this thesis, as this method would have required extensive knowledge of each social protection system which was analysed. On the other hand, if we look at external validity while there is a decently large amount of cultural diversity within the data set, it simply remains the fact that this analysis was conducted in European countries, and thus there is no way of knowing for sure if the results are also applicable to other societies. Things such as attitudes towards social protection may play a large role in the relationship of social protection complexity and wellbeing. Hence, it could be possible that the findings may be different in a society like the United States, due to their widely differing views on the role the state should play in providing societal protection. Thus, further research looking at the role complexity plays when it comes to wellbeing is called for, especially in different types of societies.

5.2 Further research:

Overall, the study of the complexity of the social protection system can still provide a fruitful area of study, and further research into this phenomenon should be conducted. As stated before, more in depth investigation of social protection systems is called for. This could be done either as a case study looking at specific countries and how the internal workings of their social protection system affect the wellbeing of their citizens. Another way would be to conduct an event study, using a country which reformed their social protection system, by simplifying it and seeing how that affects wellbeing. While there are some studies which look at the effect of pension reform (Pak, 2020) (de Mesa, et al, 2006) a more comprehensive look at social protection could benefit policymakers in seeing the benefits and potential unintended consequences of these reforms.

6. Bibliography.

Age dependency ratio (% of working-age population) | Data. (2019). Worldbank.org. https://data.worldbank.org/indicator/SP.POP.DPND

Arjona, R., Ladaique, M., & Pearson, M. (2003). Growth, Inequality and Social \Protection. *Canadian Public Policy / Analyse de Politiques*, *29*, S119. https://doi.org/10.2307/3552279

Athanasenas, A., Chapsa, X., & Michailidis, A. (2015). Investigating social protection convergence in the EU-15 : a panel data analysis. European Research Studies Journal, 18(2), 79-96.

Beja, E. L. (2013). Who is Happier: Housewife or Working Wife? *Applied Research in Quality of Life*, *9*(2), 157–177. https://doi.org/10.1007/s11482-013-9235-9

Database - Eurostat. (2019). Europa.eu. https://ec.europa.eu/eurostat/data/database

de Mesa, A. A., Bravo, D., Behrman, J., Mitchell, O., & Todd, P. (2006). *The Chilean Pension Reform Turns 25: Lessons From the Social Protection Survey*. https://doi.org/10.3386/w12401

Dercon, S. (2011) Social Protection, Efficiency and Growth, Oxford, UK: University of Oxford, CSAE Working Paper WPS/2011–17.

Diener, E., Oishi, S., & Lucas, R. E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual review of psychology*, *54*(1), 403-425.

Easterlin, R. A. (1974). Does economic growth improve the human lot? Some empirical evidence. In *Nations and households in economic growth* (pp. 89-125). Academic Press. %

Esping-Andersen, G., & Myles, J. (2011). Economic Inequality and the Welfare State. In *Oxford Handbooks Online*. https://doi.org/10.1093/oxfordhb/9780199606061.013.0025

Estevez-Abe, M., Iversen, T., & Soskice, D. (2001). Social protection and the formation of skills: A reinterpretation of the welfare state. *Varieties of capitalism: The institutional foundations of comparative advantage*, *145*, 145-183.

Ferrer-i-Carbonell, A. (2012). Happiness economics. *SERIEs*, *4*(1), 35–60. https://doi.org/10.1007/s13209-012-0086-7

International Labour Organization. (2021). World Social Protection Report 2020–22: Social Protection at the Crossroads–In Pursuit of a Better Future.

Johns, H., & Ormerod, P. (2008). The unhappy thing about happiness economics. *Real-World Economics Review, Issue*, (46), 139-146.

Kolev, A., & Tassot, C. (2016). Can investments in social protection contribute to subjective well-being? *OECD Development Center Working Papers*. https://doi.org/10.1787/18151949

Loach, K. (Director). (2016). *I, Daniel Blake*. British Film Institute.

Mencarini, L., & Sironi, M. (2010). Happiness, Housework and Gender Inequality in Europe. *European Sociological Review*, *28*(2), 203–219. https://doi.org/10.1093/esr/jcq059

Niblett, R., Begg, I., & Mushövel, F. (2015). The Welfare State in Europe: Visions for reform.

Nikolaev, B., & Rusakov, P. (2015). Education and happiness: an alternative hypothesis. *Applied Economics Letters*, *23*(12), 827–830. https://doi.org/10.1080/13504851.2015.1111982

Pak, T.-Y. (2020). Social protection for happiness? The impact of social pension reform on subjective well-being of the Korean elderly. *Journal of Policy Modeling*, *42*(2), 349–366. https://doi.org/10.1016/j.jpolmod.2019.12.001

Shah, A., Padayatchi, M., & Das, K. (2007). The relationship between elderly suicide rates and elderly dependency ratios: a cross-national study using data from the WHO data bank. *International Psychogeriatrics*, *20*(03). https://doi.org/10.1017/s104161020700628x

Stutzer, A., & Frey, B. S. (2010). Recent Advances in the Economics of Individual Subjective Well-Being. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1579572

Tang, J., Xu, Y., Ma, W., & Gao, S. (2021). Does participation in poverty alleviation programmes increase subjective well-being? Results from a survey of rural residents in Shanxi, China. *Habitat International*, *118*, 102455. https://doi.org/10.1016/j.habitatint.2021.102455

Taydas, Z., & Peksen, D. (2012). Can states buy peace? Social welfare spending and civil conflicts. *Journal of Peace Research*, *49*(2), 273–287. https://doi.org/10.1177/0022343311431286

United Nations Development Programme. (2022). *Human Development Index (HDI) | Human Development Reports*. Hdr.undp.org. https://hdr.undp.org/en/content/human-development-index-hdi

Vinci, V., Gassmann, F., & Mohnen, P. (2022). The effect of institutional factors and people's preferences on expenditure for social protection. *International Social Security Review*, *75*(1), 107–154. https://doi.org/10.1111/issr.12290

Wang, C., Caminada, K., & Goudswaard, K. (2012). The redistributive effect of social transfer programmes and taxes: A decomposition across countries. *International Social Security Review*, *65*(3), 27–48. https://doi.org/10.1111/j.1468-246x.2012.01435.

Wooldridge, J. M. (2016). *Introductory econometrics : a modern approach* (6th ed.). Boston: Cengage Learning, Cop.

World Bank.org. (2021). Labor force participation rate, total (% of total population ages 15-64) (modeled ILO estimate) | Data. Retrieved from data.worldbank.org website: https://data.worldbank.org/indicator/SL.TLF.ACTI.ZS

Wu, J., Värnik, A., Tooding, L.-M., Värnik, P., & Kasearu, K. (2013). Suicide among older people in relation to their subjective and objective well-being in different European regions. *European Journal of Ageing*, *11*(2), 131–140. https://doi.org/10.1007/s10433-013-0297-1.

7. Appendixes

Appendix 1: List of countries

Austria
Belgium
Bulgaria
Croatia
Cyprus
Czechia
Denmark
Estonia
Finland
France
Germany (until 1990 former territory of the FRG)
Greece
Hungary
Iceland
Ireland
Italy
Latvia
Lithuania
Luxembourg
Malta
Montenegro
Netherlands
North Macedonia
Norway
Poland
Portugal
Romania
Serbia
Slovakia
Slovenia
Spain
Sweden
Switzerland
Turkey
United Kingdom

Appendix 2 a: Breusch-Pagan/Cook-Weisberg test for heteroskedasticity (Model 1)

Assumption: Normal error terms

Variable: Fitted values of SWB

H0: Constant variance

chi2(1) = 15.83

Prob > chi2 = 0.0001

Appendix 2 b: Breusch-Pagan/Cook-Weisberg test for heteroskedasticity (Model 2)

Breusch “Pagan/Cook“Weisberg test for heteroskedasticity

Assumption: Normal error terms

Variable: Fitted values of Pos

H0: Constant variance

chi2(1) = 2.06

Prob > chi2 = 0.1509

Appendix 2 c: Breusch-Pagan/Cook-Weisberg test for heteroskedasticity (Model 3)

Breusch “Pagan/Cook “Weisberg test for heteroskedasticity

Assumption: Normal error terms

Variable: Fitted values of HDI

H0: Constant variance

chi2(1) = 4.07

Prob > chi2 = 0.0435

Appendix 2 d: Breusch-Pagan/Cook-Weisberg test for heteroskedasticity (Model 4)

Breusch “Pagan/Cook“Weisberg test for heteroskedasticity

Assumption: Normal error terms

Variable: Fitted values of Comp

H0: Constant variance

chi2(1) = 25.42

Prob > chi2 = 0.0000

Appendix 3a: Variance inflation test (Model 1,2,3)

Variance inflation factor

	VIF	1/VIF
LogGDPpercapita	1.467	.682
PartRate	1.391	.719
ADR	1.152	.868
Comp	1.043	.959
Mean VIF	1.263	.

Appendix 3b: Variance inflation test (Model 4)

Variance inflation factor

	VIF	1/VIF
ADR	1.068	.936
PartRate	1.068	.936
Mean VIF	1.068	.

Appendix 4a: Skewness and kurtosis tests for normality (Model 1)

				----- Joint test -----	
Variable	Obs	Pr(skewness)	Pr(kurtosis)	Adj chi2(2)	Prob>chi2
resid	381	0.028	0.843	4.890	0.087

Appendix 4b: Skewness and kurtosis tests for normality Model 2

				----- Joint test -----		
Variable	Obs	Pr(skewness)	Pr(kurtosis)	Adj	chi2(2)	Prob>chi2
resid2	381	0.000	0.000	57.880	0.000	

Appendix 4c: Skewness and kurtosis tests for normality Model 3

----- Joint test -----

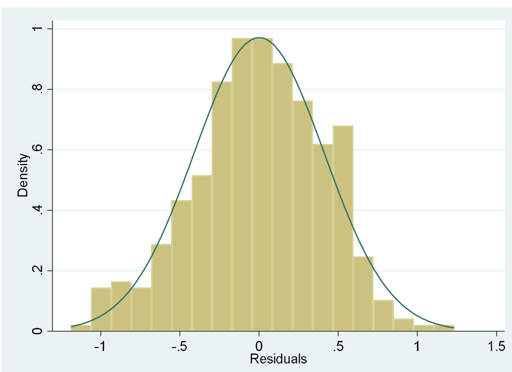
Variable	Obs	Pr(skewness)	Pr(kurtosis)	Adj	chi2(2)	Prob>chi2
resid3	381	0.000	0.000	97.310	0.000	

Appendix 4d: Skewness and kurtosis tests for normality Model 4

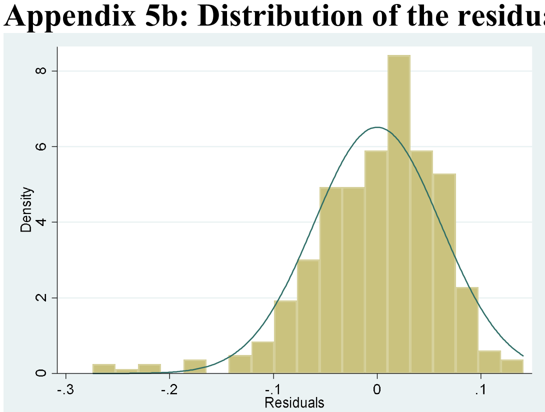
----- Joint test -----

Variable	Obs	Pr(skewness)	Pr(kurtosis)	Adj	chi2(2)	Prob>chi2
resid4	381	0.931	0.023	5.180	0.075	

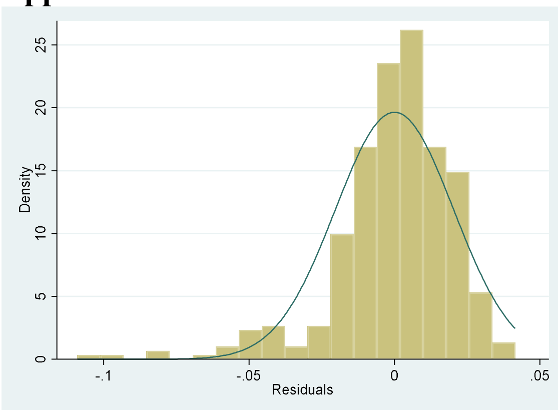
Appendix 5:a Distribution of the residuals model 1



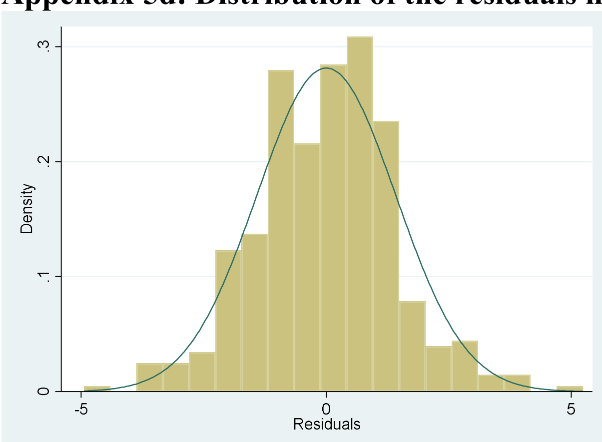
Appendix 5b: Distribution of the residuals model 2



Appendix 5c: Distribution of the residuals model 3



Appendix 5d: Distribution of the residuals model 4



Appendix 6a: Regression 1, (random effects)

SWB	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Comp	-.016	.024	-0.67	.503	-.062	.03	
ADR	.014	.01	1.41	.157	-.005	.032	
PartRate	.001	.001	1.98	.048	0	.003	**
LogGDPpercapita	1.682	.269	6.24	0	1.154	2.21	***
Constant	-12.154	2.73	-4.45	0	-17.505	-6.804	***

Mean dependent var	6.279	SD dependent var	0.922
Overall r-squared	0.787	Number of obs	381
Chi-square	163.218	Prob > chi2	0.000
R-squared within	0.358	R-squared between	0.841

**** p<.01, ** p<.05, * p<.1*

Appendix 6b: Regression 2, (fixed effects)

Regression results

Pos	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Comp	-.006	.003	-2.45	.015	-.011	-.001	**
ADR	.004	.001	2.93	.004	.001	.007	***
PartRate	0	0	2.48	.014	0	0	**
LogGDPpercapita	-.056	.04	-1.40	.162	-.134	.023	
Constant	1.122	.357	3.14	.002	.42	1.824	***

Mean dependent var	0.710	SD dependent var	0.102
R-squared	0.069	Number of obs	381
F-test	6.386	Prob > F	0.000
Akaike crit. (AIC)	-1386.703	Bayesian crit. (BIC)	-1366.989

**** p<.01, ** p<.05, * p<.1*

Appendix 6c: Regression 3, (random effects) robust

Regression results

HDI_Score	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Comp	.002	.001	1.27	.205	-.001	.004	
ADR	.002	.001	2.99	.003	.001	.003	***
PartRate	0	0	-0.03	.975	0	0	
LogGDPpercapita	.092	.022	4.27	0	.05	.134	***
Constant	-.202	.204	-0.99	.322	-.603	.198	

Mean dependent var	0.875	SD dependent var	0.052
Overall r-squared	0.827	Number of obs	381
Chi-square	366.392	Prob > chi2	0.000
R-squared within	0.709	R-squared between	0.842

**** p<.01, ** p<.05, * p<.1*

Appendix 7a: Regression 1, fixed effects

SWB	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Comp	-.027	.025	-1.07	.291	-.077	.024	
ADR	.003	.015	0.21	.837	-.028	.034	
PartRate	.001	.001	0.62	.536	-.002	.003	
LogGDPpercapita	2.105	.703	2.99	.005	.676	3.534	***
Constant	-15.837	6.673	-2.37	.023	-29.398	-2.276	**

Mean dependent var	6.279	SD dependent var	0.922
R-squared	0.364	Number of obs	381
F-test	9.809	Prob > F	0.000
Akaike crit. (AIC)	-8.743	Bayesian crit. (BIC)	7.029

**** p<.01, ** p<.05, * p<.1*

Appendix 7b : Regression 3, fixed effects

Regression results

HDI_Score	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Comp	.002	.001	1.59	.122	0	.004	
ADR	.002	.001	2.18	.036	0	.004	**
PartRate	0	0	-0.56	.582	0	0	
LogGDPpercapita	.09	.032	2.80	.008	.025	.155	***
Constant	-.181	.3	-0.60	.551	b-.792	.43	

Mean dependent var	0.875	SD dependent var	0.052
R-squared	0.710	Number of obs	381
F-test	84.342	Prob > F	0.000
Akaike crit. (AIC)	-2683.404	Bayesian crit. (BIC)	-2667.633

**** p<.01, ** p<.05, * p<.1*

Appendix 8:Regression 2 with gdp per capita - Full results

Regression results

Comp	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ADR	-.098	.021	-4.59	0	-.139	-.056	***
LogGDPpercapita	2.156	.457	4.72	0	1.261	3.052	***
PartRate	-.001	.001	-0.69	.492	-.004	.002	
Constant	-7.558	4.245	-1.78	.075	-15.877	.762	*

Mean dependent var	9.971	SD dependent var	1.434
Overall r-squared	0.033	Number of obs	381
Chi-square	30.674	Prob > chi2	0.000
R-squared within	0.138	R-squared between	0.039

**** p<.01, ** p<.05, * p<.1*

Appendix 9: (Goodness of fit of SEM analysis

Equation-level goodness of fit

Dependent variable	Variance Fitted	Predicted	Residual	R-squared	mc	mc2
Complexity	115.661	0.025	115.636	.0002	.0148	.0002
Lpos	.0227	0.002	.02112	.0711	.2666	.0711

mc = Correlation between dependent variable and its prediction.
mc2 = mc^2 is the Bentlerâ€“Raykov squared multiple correlation coefficient.

Appendix 10 Do file

```
*Do file master's Thesis
clear
import excel "C:\Users\ollih\Desktop\Master Thesis\Admin cost of social protection.xlsx",
firstrow clear
encode AC, gen (Admin)
encode Socialprotectionbenefits, gen (SP)
rename Time year
save Panel31, replace
clear
import excel "C:\Users\ollih\Desktop\Master Thesis\Wellbeing-Panel.xls", firstrow clear
encode Countryname, gen (Country)
merge m:1 C_ID year using "C:\Users\ollih\Desktop\Stata\Dofiles\Panel31.dta", force

save ActualPanel21, replace

clear
use "C:\Users\ollih\Desktop\Stata\Dofiles\ActualPanel21.dta"
keep if _merge== 3
gen Complexity = (Admin/SP)
hist Complexity
gen LComp = log(Complexity)
hist LComp, normal
gen Comp = (LComp + 10)
hist Comp, normal

*normally distributed
hist LifeLadder, normal
rename LifeLadder SWB
drop _merge
save Final, replace


clear
import excel "C:\Users\ollih\Desktop\Master Thesis\age dependency.xlsx", firstrow clear
```

```
rename Time year
encode CountryName, gen (Country)
drop CountryName
*drop CountryCode
drop Time
encode Laborforceparticipationratef, gen (PartRate)
rename Agedependencyratioofworki ADR



merge m:m Country year using C:\Users\ollih\Desktop\Stata\Dofiles\Final.dta
keep if _merge== 3
drop _merge

*Variace inflation test
 reg SWB Comp  ADR PartRate LogGDPpercapita PartRate, robust
asdoc vif
 *variance inlationlation test  is insignificant so no multicollinearity

 reg SWB Comp ADR PartRate LogGDPpercapita
 asdoc hettest
 *Heteroskedasticity could be present thus we have to run the regression using robust standard
errors
 reg SWB LComp  ADR PartRate LogGDPpercapita, robust
hist SWB, normal
gen LSWB = log(SWB)
hist LSWB
*SWB is sufficiently normally distributed, no need to log
drop LSWB
sktest SWB
*Not sure how to interpreted the skweness is low however the

*Set up the panel datasets
xtset Country year


ssc install xttest3, replace
xtreg SWB LComp ADR PartRate LogGDPpercapita, fe robust
 rename Positiveaffect Pos

 **First Robustness check
 reg Pos Comp PartRate ADR LogGDPpercapita
 hettest
 xtreg Pos LComp PartRate ADR LogGDPpercapita, fe


 save ActualFinal, replace
```

```
*Second dependent variable (Robustness check)
clear
import excel "C:\Users\ollih\Desktop\Human Development Index (HDI) actual.xlsx", firstrow
clear
encode CountryName, gen (Country)

merge m:m Country year using "C:\Users\ollih\Desktop\Stata\Dofiles\ActualFinal.dta"
keep if _merge== 3
drop _merge
sum HDI_Score, detail
hist HDI_Score, normal
*normally distributed
gen LHDI = log(HDI_Score)
hist LHDI, normal
drop LHDI
reg HDI_Score LComp  ADR PartRate LogGDPpercapita
hettest
*still need to use robust

//CHecking ols assumption 7 Model 1
reg SWB Comp  ADR PartRate LogGDPpercapita
predict resid, residuals
hist resid, normal
sktest resid
***errors are normnally distributed

///Model 4
reg Comp ADR PartRate
predict resid4, residuals
hist resid4, normal
asdoc sktest resid4

///Model 2
reg Pos Comp  ADR PartRate LogGDPpercapita
predict resid2, residuals
hist resid2, normal
asdoc sktest resid2

//Model 3
reg HDI_Score Comp  ADR PartRate LogGDPpercapita
predict resid3, residuals
hist resid3, normal
asdoc sktest resid3


//Hausman test model 1
xtreg SWB Comp  ADR PartRate LogGDPpercapita,  fe
estimate store fe
```

```
xtreg SWB Comp  ADR PartRate LogGDPpercapita ,  re
estimate store re
***Cheking whether Fixed or random effects are appropriate***
**Null hypothesis = random**
asdoc hausman fe re
*HEre it looks we need random Effects

drop _est_fe
drop _est_re

//Hausman test model 2
xtreg Pos Comp  ADR PartRate LogGDPpercapita,  fe
estimate store fe
xtreg Pos Comp  ADR PartRate LogGDPpercapita ,  re
estimate store re
***Cheking whether Fixed or random effects are appropriate***
asdoc hausman fe re
* FIXED Effects

drop _est_fe
drop _est_re

//Hausman test model 3
xtreg HDI_Score Comp  ADR PartRate LogGDPpercapita,  fe
estimate store fe
xtreg HDI_Score Comp  ADR PartRate LogGDPpercapita ,  re
estimate store re
***Cheking whether Fixed or random effects are appropriate***
**Null hypothesis = random**
asdoc hausman fe re

//Hausman test model 4
xtreg Comp ADR PartRate LogGDPpercapita,  fe
estimate store fe
 xtreg Comp ADR PartRate LogGDPpercapita ,  re
estimate store re
asdoc hausman fe re


//Cleaning the data
drop Administrativecost
drop SocialSpending
drop CountryName
drop _est_fe
drop _est_re
drop Laborforceparticipationratef
```

```
//Checking OLS heteroskedasticity model 2 & 3
reg  Pos Comp  ADR PartRate LogGDPpercapita
hettest
**insignifitcant
reg  HDI_Score Comp  ADR PartRate LogGDPpercapita
asdoc hettest
*Significant - we need robust

//Preliminary regressions
 xtreg  SWB Comp  ADR PartRate LogGDPpercapita ,  re robust
 xtreg  Pos Comp  ADR PartRate LogGDPpercapita,  re
 xtreg  HDI_Score Comp  ADR PartRate LogGDPpercapita, re robust
save ActualFinal37, replace


 //Checkign correlation between the dependent variables
 pwcorr SWB  Comp ADR PartRate LogGDPpercapita Healthylifeexpectancyatbirth
asdoc pwcorr SWB  Pos HDI_Score



gen Lpos = log(Pos)
gen HDI = log(HDI_Score)

ssc install xttest2.pkg, replace
//Run the hettest with this for models 1-3\\
///Nested with everything ///
 asdoc xtreg SWB Comp   LogGDPpercapita,  re robust  nest replace
 asdoc xtreg SWB   PartRate LogGDPpercapita,  re robust  nest append
  asdoc xtreg SWB ADR   LogGDPpercapita,  re robust  nest append
 asdoc xtreg Lpos Comp   LogGDPpercapita,  fe   nest append
 asdoc xtreg Lpos   PartRate LogGDPpercapita,  fe   nest append
 asdoc xtreg Lpos   ADR LogGDPpercapita,  fe  nest append
  asdoc xtreg HDI Comp   LogGDPpercapita,  re robust  nest append
 asdoc xtreg HDI   PartRate LogGDPpercapita,  re robust  nest append
  asdoc xtreg HDI ADR   LogGDPpercapita,  re robust  nest append

  ///Testing the first hypothesis
 xtreg SWB Comp ADR PartRate LogGDPpercapita,  fe robust
   xtreg Lpos Comp ADR PartRate LogGDPpercapita,  fe
    xtreg HDI_Score Comp ADR PartRate LogGDPpercapita, fe robust
        ///Getting the tables//
        asdoc xtreg SWB Comp ADR PartRate LogGDPpercapita,  re robust  nest replace
                asdoc xtreg Lpos Comp ADR PartRate LogGDPpercapita,  fe   nest append
                        asdoc xtreg HDI Comp ADR PartRate LogGDPpercapita,  re robust  nest
append
```

```
	asdoc xtreg HDI_Score Comp ADR PartRate LogGDPpercapita,  re robust

asdoc xtreg SWB Comp ADR PartRate LogGDPpercapita, fe robust nest replace
asdoc xtreg HDI_Score Comp ADR PartRate LogGDPpercapita,  fe robust nest append

//h2 */esting  H2 - The detrminants of Complexity, initial alt hypothesis is that a larger ADR will cause a higher degree of complexity

reg Comp ADR PartRate LogGDPpercapita
hettest
vif
** it would appear that there is a highly significant and negavitve correlation between Complexity and ADR



*3rd hypothesis
ssc install medsem, replace

asdoc sem (ADR -> Complexity, ) (ADR -> Lpos, ) (Complexity -> Lpos, ), nocapslatent
asdoc estat teffects
```