Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Elske Leenaars

Public Attitudes towards Pensions

BSc Thesis 2010

Public Attitudes towards Pensions

by
Elske Leenaars (397869)

A thesis submitted in partial fulfillment of the requirements for the degree of
Bachelor in Econometrics and Operations Research

Faculty of Economics and Business Administration
Tilburg University

supervised by
Prof. dr. A. van Soest

June 23, 2010

ABSTRACT

In this thesis we investigate what peoples opinions are with respect to the rise in individual responsibility for retirement provisions. Moreover, we will investigate whether people with different socio-economic backgrounds might tend to have different opinions. In our investigation we make use of the ordered probit and ordered logit models and their marginal effects. We will investigate if the results of the ordered probit model are robust against its distributional assumption, by extending the ordered probit model with a semi-nonparametric estimation of the distribution of the error terms. We find that the characteristics age, number of children and whether an individual lives in a rural or urban area are of influence to their opinion with respect to the rise in individual responsibility for retirement provision. By extending the ordered probit model, we come to the conclusion that the net income of the individual is of influence as well.

CONTENTS

1. INTRODUCTION

Because of population ageing, a lively debate has developed on public and private pension provisions, both in society and in the social sciences. For the government, the opinion of the population and of specific socio-economic groups can be important to find support for policy measures aimed at reducing the cost of public pension provision or increasing the standard retirement age.

Page and Shapiro (1983) presented evidence that public opinion is often a proximate cause of policy, affecting policy more than policy influences opinion. Also, Hayo and Ono (2009) argue that "the success of economic reforms, at least in a democracy, is dependent upon peoples opinions. In this sense, investigations of the expected objective effects do not suffice. Research on subjective aspects of social security reform is essential for shedding light on whether a reform plan would be accepted by various segments of society".

In the literature, a lot of papers have been written about economic reforms and in particular about pension reforms. Only a few of these papers deal with public attitudes towards economic and pension reforms. A paper that focusses on public attitudes with respect to the livelihood of the elderly is written by Hayo and Ono (2007). Using data from Germany and Japan from the year 2003, it is found that both a higher income and a higher age, inclines people more toward private responsibility for the livelihood of the elderly. Also, part-time work status, pensioner status and political orientation significantly affect the opinion of citizens of either Germany or Japan. Hayo and Ono (2009) analyze public attitudes towards the livelihood and care of the elderly in Japan, using data over the period 2000-2005. They show that many significant factors are common for both the livelihood and care of the elderly. The main difference with their previous paper, is that now they have used data over a period of 5 years. From this, they were able to conclude that there is an increase in support of a government-based system over time. Corneo and Grüner (2000) have written a paper about individual preferences for political redistribution. In this paper, people's support of governmental reduction of income inequality is being evaluated by means of three effects. The results indicate that not only the economic motivation drives individual preferences for governmental redistribution, but also the desire to act in accordance with public values and the effect of the redistribution on the relative living standard of the individual.

In this thesis, we are going to investigate what the opinion of the Dutch population with respect to individual responsibility for retirement provision is. Also, we will investigate whether people with different socio-economic backgrounds might tend to have different opinions.

An overview of the data is given in section 2.1. We will start section 2.2 with a discussion of the ordered probit and ordered logit models, which are being used in our investigation. Also, we will discuss the marginal effects of the ordered probit and ordered logit models. We conclude section 2.2 with the theory of semi-nonparametric estimation and in particular the semi-nonparametric estimator of Gallant and Nychka. Theoretical hypotheses regarding a priori expectations will be presented in section 2.3. In section 2.4 we will show the results of the ordered probit and ordered logit models, the marginal effects as well as the semi-nonparametric estimation. Also, we will discuss and compare the different results.

2. INVESTIGATION

2.1 Data

The survey we will use in this thesis is included in the Dutch CentERpanel, administrated by CentERdata at Tilburg University. The CentERpanel covers the population in the Netherlands of ages 16 and older. It is composed of over 2000 households in which one or more adults complete questionnaires at home every week through the internet. The CentERpanel is not restricted to households with prior access to the internet: households without access are given a settop box that can be connected to their television set or phone line. There is rich background information about the panel respondents. More information about the CentERpanel can be found at www.centerdata.nl.

The survey is presented to the members of the CentERpanel that are aged 25 and older. It is taken in march 2008 and 1318 members responded. The survey consists of multiple questions, the question we will investigate in particular is 'Some people think that in the future, individual responsibility for retirement provision will increase and people will more than now have to save themselves for their old age. Would you consider this a desirable or undesirable development?'. Respondents have to give an answer on a scale from 1 to 7. Here, 1 means that the respondent thinks the development is very undesirable, while 7 means that the respondent thinks the development is very desirable. The histogram below shows the number of respondents that have given a certain answer.

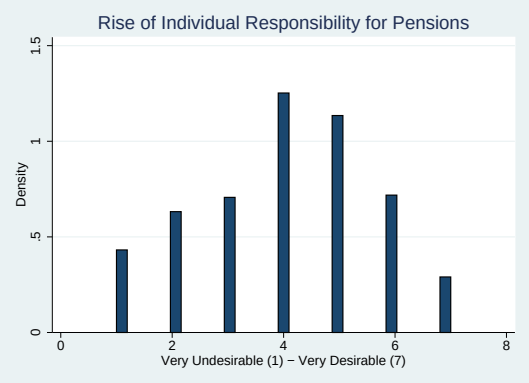

Fig. 2.1: Histogram of the answers to the question about individual responsibility for retirement provision.

When calculating the number of respondents that believe that the development is desirable, i.e. counting the number of people that answered either the number 1, 2 or 3, we see that 34.3% of the respondents thinks that the rise in individual responsibility for retirement provision is an undesirable development. On the other hand, 41.5% of the respondents think the rise in individual responsibility is a desirable development. Even though none of these is a majority, we do see that there are more people who think the development is desirable, than there are people who think the development is undesirable.

For the investigation whether people with different socio-economic backgrounds might tend to have different opinions, we will make use of explanatory variables. First, a short description of these explanatory variables is given.

Age The age of the respondent.

Net Income The net income of the respondent per month in 1.000 euros.

Sex Women Assigns the value 1 if the respondent is a women and the value 0 if the respondent is a man.

Occupation Labor Assigns the value 1 if the respondent has a paid job, 0 otherwise.

Occupation Pension Assigns the value 1 if the respondent is retired, 0 otherwise.

Occupation Social Security Assigns the value 1 if the respondent is in social security, 0 otherwise.

Occupation Other Assigns the value 1 if the respondent does not have a paid job, is not retired and is not in social security, 0 otherwise.

Urban Assigns a value on a scale from 1 to 5, in which 1 means the respondent lives in a very urban surrounding and 5 means the respondent lives in a very rural surrounding.

Number of Children The number of children the respondent has.

Partner Assigns the value 1 if the head of the household lives together with a partner, 0 otherwise.

Education Category Assigns a value on a scale from 1 to 6, in which:
1 means the respondent has completed primary school,
2 means the respondent has completed 'VMBO',
3 means the respondent has completed 'HAVO/VWO',
4 means the respondent has completed 'MBO',
5 means the respondent has completed 'HBO',
6 means the respondent has completed 'WO'.

Some summary statistics of the explanatory variables are given in the following table:

Variable	Mean	St. dev.
Age	54.32	0.39
Net Income	1.53	.28
Sex Women	0.44	0.01
Occupation Labor	0.51	0.50
Occupation Pension	0.27	0.01
Occupation Social Security	0.07	0.01
Occupation Other	0.16	0.01
Urban	3.01	0.04
Number of Children	0.66	0.03
Partner	0.78	0.01
Education Category	3.64	0.04

Tab. 2.1: Summary statistics of the data.

It is remarkable that the mean age of the respondents is 54.32 years, as this is quite high. This high age is probably due to the fact that the respondents are no younger than 25 years old.

2.2 Model

2.2.1 Ordered Probit and Ordered Logit

In this thesis, we deal with a dependent variable which can only take finitely many different values. When trying to use a linear regression model to explain a dependent variable in which only finitely many values can be taken, the (weak) assumptions of the linear regression model are not satisfied. Because of this and the fact that in the dependent variable a meaningful ordering can be made, the ordered probit model and ordered logit model will be used.

The ordered probit model is defined by:

$$y_i^* = \beta' x_i + \varepsilon_i, \tag{2.1}$$

$$\varepsilon_i \mid x_i \sim N(0,1)\ iid, \tag{2.2}$$

$$y_i = \begin{cases} 1 & \text{if} \quad \alpha_0 < y_i^* < \alpha_1, \\ 2 & \text{if} \quad \alpha_1 < y_i^* < \alpha_2, \\ \vdots & \qquad\qquad \vdots \\ J & \text{if } \alpha_{J-1} < y_i^* < \alpha_J. \end{cases} \tag{2.3}$$

In equation 2.3, y_i is the answer of the respondents to the question about individual responsibility for retirement provision. The y_i^*'s and α_i's in equations 2.1 and 2.3 can be interpreted as

$$P(y_i = j|x_i) = P(\alpha_{j-1} < y_i^* \leq \alpha_j|x_i)$$

Thus, y_i^* can be seen as person i's desirability for individual responsibility for retirement provision on a scale from $-\infty$ to ∞. The higher y_i^*, the higher this desirability. Because y_i^* can take any real value and each i must be classified into one category, we have $-\infty = \alpha_0 \leq \alpha_1 \leq ... \leq \alpha_J = \infty$.
Using equation 2.2, we can rewrite $P(y_i = j|x_i)$ in the following way:

$$\begin{aligned} P(y_i = j|x_i) &= P(\alpha_{j-1} < y_i^* \leq \alpha_j|x_i) \\ &= P(\alpha_{j-1} - \beta' x_i < \varepsilon_i \leq \alpha_j - \beta' x_i|x_i) \\ &= \Phi(\alpha_j - \beta' x_i) - \Phi(\alpha_{j-1} - \beta' x_i) \end{aligned}$$

For identification, we set the intercept parameter at 0, essentially meaning that we eliminate the constant term from the model. Also, we set the standard deviation σ of $\varepsilon_i|x_i$ at 1.

To specify the ordered probit model, we used the standard normal distribution. However, an alternative transformation is also quite popular. When we change equation 2.2 into $\varepsilon_i|x_i \sim LOG(0,1)$, we arrive at the ordered logit model. The ordered logit model is defined by:

$$y_i^* = \beta' x_i + \varepsilon_i, \tag{2.4}$$

$$\varepsilon_i \mid x_i \sim LOG(0,1)\ iid, \tag{2.5}$$

$$y_i = \begin{cases} 1 & \text{if} \quad \alpha_0 < y_i^* < \alpha_1, \\ 2 & \text{if} \quad \alpha_1 < y_i^* < \alpha_2, \\ \vdots & \qquad\qquad \vdots \\ J & \text{if } \alpha_{J-1} < y_i^* < \alpha_J. \end{cases} \tag{2.6}$$

In the ordered logit model, y_i, y_i^* and α_i can be interpreted as in the ordered probit model.

The α and β coefficients in the ordered probit and ordered logit models are calculated using the method of (conditional) Maximum Likelihood (ML).

2.2.2 Marginal Effects

In a linear regression model, β_k can be interpreted as the amount by which the dependent variable y_i on average increases, if the kth explanatory variable increases by one unit. This holds, because we have $E(y_i|x_i) = \beta' x_i$ and hence,

$$\frac{\partial E(y_i|x_i)}{\partial x_i} = \beta.$$

In the ordered probit and ordered logit model this does not hold, as these models are nonlinear. We have $E(y_i|x_i) = \sum_{j=1}^{J} j \cdot P(y_i = j|x_i) = \sum_{j=1}^{J} P(y_i \geq j|x_i)$ and hence,

$$\frac{\partial E(y_i|x_i)}{\partial x_i} = \sum_{j=1}^{J} \frac{\partial P(y_i \geq j|x_i)}{\partial x_i} \neq \beta.$$

Because of the lack of interpretation of the β_k's in the ordered probit and ordered logit models, we will investigate the marginal effects. The marginal effects estimate the change in the probability $P(y_i = j|x_i)$ when x_k increases by one unit. Thus, the marginal effects are defined by

$$\frac{\partial P(y_i = j|x_i)}{\partial x_i} = \beta(\varphi(\alpha_{j-1} - \beta' x_i) - \varphi(\alpha_j - \beta' x_i)).$$

Here, φ is the derivative of the standard normal distribution or the standard logistic distribution in case of the ordered probit model or ordered logit model respectively.

Although the marginal effects already have an interpretation, it might be more interesting to look at the marginal effect of $P(y_i \leq j|x_i)$, specifically for j=3. This marginal effect captures the change in the probability that a person thinks the rise in individual responsibility for retirement provisions is undesirable when increasing x_k by one unit. The marginal effect of this probability equals

$$\frac{\partial P(y_i \leq j|x_i)}{\partial x_i} = -\beta\varphi(\alpha_j - \beta' x_i).$$

Again, φ is the derivative of the standard normal distribution or the standard logistic distribution in case of the ordered probit model or ordered logit model respectively.

2.2.3 Semi-nonparametric Estimation

The ordered probit and ordered logit models require specific distributional assumptions, namely the assumptions $\varepsilon_i \mid x_i \sim N(0,1)$ *iid* and $\varepsilon_i \mid x_i \sim LOG(0,1)$ *iid* respectively. But actually, the distribution of the error terms is unknown meaning that the assumption might affect the results significantly. Therefore, we will investigate whether this distributional assumption indeed influences the outcome of ordered probit. We will use a semi-nonparametric method to estimate the unknown density of the conditional error terms. Semi-nonparametric estimations partly impose a parametric form on the density while keeping other parts of the density nonparametric. Wolfowitz (1942) defines the parametric case as the case where

> The distribution functions of the various stochastic variables ... are assumed to be of known functional form, and the theories of estimation and of testing hypotheses are theories of estimation of and of testing hypotheses about, one or more parameters, finite in number, the knowledge of which would completely determine the various distribution functions involved.

The nonparametric case on the other hand, is defined as the case where "the functional forms of the distributions are unknown".

The semi-nonparametric estimator relaxes the distributional assumptions of the ordered probit and ordered logit models. We will compare the results of

the ordered probit model with the extended ordered probit model, in which the distributional assumption is replaced with the estimate of the unknown density, to see if the distributional assumption influences the outcome of ordered probit and ordered logit.

Gallant and Nychka (1987) approximate an unknown density using the product of a squared polynomial and a normal density. Here, the squared polynomial is nonparametric while the normal density is parametric, the approximation therefore is semi-nonparametric. The approximation is specified as

$$f_K(\epsilon) = \frac{1}{\theta}(\sum_{k=0}^{K} \gamma_k \epsilon^k)^2 \phi(\epsilon),$$

where $\phi(\epsilon)$ is the standard normal density function, $\gamma_0 = 1$ for identification and where

$$\theta = \int_{-\infty}^{\infty} (\sum_{k=0}^{K} \gamma_k \epsilon^k)^2 \phi(\epsilon) d\epsilon,$$

to make sure that $f_K(\epsilon)$ integrates to 1. From this, it follows that the cumulative density function equals

$$F_K(u) = \int_{-\infty}^{u} \frac{1}{\theta}(\sum_{k=0}^{K} \gamma_k \epsilon^k)^2 \phi(\epsilon) d\epsilon. \tag{2.7}$$

We will obtain estimates for β, α_i and the parameters of the approximation of the unknown density by using the method of maximum likelihood conditional on K. After that, we will use standard model selection procedures to choose between different (integer) values of K. The pseudo-likelihood function we will maximize is

$$\log L(\alpha, \beta) = \sum_{i=1}^{N} \sum_{j=1}^{J} y_{ij} \log[F_K(\alpha_j - \beta' x_i) - F_K(\alpha_{j-1} - \beta' x_i)],$$

where

$$y_{ij} = \begin{cases} 1 & \text{if} \quad y_i = j, \\ 0 & \text{otherwise.} \end{cases}$$

and F_K the approximation of the unknown density given by equation 2.7.

Gabler et al (1993) use the semi-nonparametric estimator of Gallant and Nychka to approximate the distribution of the error terms in a binary choice model. They found that semi-nonparametric estimation is rather efficient and outperforms the probit estimation in non-normal samples.

Stewart (2005) compares, among other semiparametric estimators, the semi-nonparametric estimator of Gallant and Nychka with the standard ordered probit estimator. Using an application about job satisfaction and an application about the willingness to pay for environmental protection, it is found that the semi-nonparametric estimator outperforms the normal ordered probit estimate

for the application about job satisfaction, but does not do so for the application about the willingness to pay for environmental protection. Furthermore, a simulation is used to conclude that the semi-nonparametric estimator performs very well in many circumstances (in particular in so called single-index models).

2.3 Theoretical Hypotheses

Before we will use econometric methods to investigate whether people with different socio-economic backgrounds might tend to have different opinions about individual responsibility for retirement provision, we will construct theoretical hypotheses. These hypotheses explain why and how the explanatory variables might influence peoples opinions. The hypotheses are based on a ceteris paribus situation.

2.3.1 Age

The Dutch pension system consists of three pillars. One of these pillars, namely the first, is a pay-as-you-go system. This means that the costs of the Dutch state pension benefits are paid by the workforce in the form of contributions. Additional funding for the Dutch state pension comes from government public funds. From the first pillar it follows that people who are older, have paid a higher total of contributions. In return, they expect a state pension. A rise in individual responsibility for pension provisions might suggest that the state pension will decrease. This suggests that older people think that the rise in individual responsibility is undesirable.

Also, people who are older have less time left to save for their individual pension provisions. So, when we assume that people are selfish and only consider what is good for them, older people will incline more towards governmental responsibility for pension provisions.

2.3.2 Income

People with a higher income are more likely to be able to save money. These savings can be used for their old-age livelihood. On the other hand, people with a lower income might not be able to save money for their old-age livelihood. Therefore, we expect people with a higher income to incline more towards individual responsibility of pension provisions.

2.3.3 Gender

A paper by Carnaghan and Bahry (1990) states that surveys have found American women to be more likely than men to be compassionate, more in favor of governmental involvement in social welfare, health, and the promotion of equality. As the responsibility of pension provisions is a governmental involvement in social welfare, this suggests that women will find the rise in individual responsibility for pension provisions more undesirable than men do.

Another argument that suggests that women, more than men, favor governmental responsibility is the fact that women live longer and thus receive more pension provision from the government, while they do not pay a higher total of contributions.

2.3.4 Occupation

The explanatory variable Occupation consists of four dummy variables. From now on we will use the first dummy variable, namely Occupation Labor, as the benchmark. We will therefore look at the effects of Occupation Pension, Occupation Social Security and Occupation Other with respect to this benchmark.

We expect people who already receive pension payments to be more in favor of government responsibility for pension provisions. People who already receive pension payments do not have the opportunity to save for their pensions anymore. Also they have paid contributions to the Dutch state pension their entire lives.

We expect people who are in social security to be more in favor of government responsibility for pension provisions. People who are in social security usually have a lower income. As we have argued before, we expect people with a lower income to incline more towards government responsibility for pension provisions.

2.3.5 Urban

Rural family values have been identified as stronger in rural areas. Also, there are stronger feelings of responsibility to old parents in rural areas (Wenger, 2001). Therefore, we expect people that live in rural areas to incline more towards individual responsibility for retirement provision.

2.3.6 Number of children

According to Breyer and Schulenburg (1987), families that bear the financial burden of raising (more than the average number of) children are most reluctant to pay for the retirement benefits to the childless singles.

When the government is responsible for pension provisions, everyone who is not retired has to pay an amount of contributions and everyone who is retired receives an amount of pension benefits. Thus, in a sense everyone pays for the pension benefits of others.

Combining these two arguments, we expect people with more children to find the rise in individual responsibility for pension provisions more desirable.

2.3.7 Partner

We do not have expectations concerning the variable Partner in relation to the question about individual responsibility for retirement provisions.

2.3.8 Education Category

More educated people know better where and how their taxes and contributions are used than those not. Therefore, they may tend to view the government more critically (Hayo and Ono, 2009). This may mean that better educated people tend to be more critical about government responsibility for pension provisions as well.

2.4 Results

2.4.1 Ordered Probit and Ordered Logit

As we have argued in section 2.2, we cannot use a linear regression model to explain the question about individual responsibility for retirement provision. Therefore, we will use the ordered probit and ordered logit models.

When regressing all explanatory variables on the question about individual responsibility for retirement provision using the ordered probit model in Stata, we obtained the results reported in table 2.2. Table 2.2 shows that the variables 'Net Income', 'Urban' and 'Number of Children' are of significant influence on a 5% level.

	Coef.	St. dev.	p-value
Age	.003	.003	0.371
Net Income	.083	.038	0.031
Sex Women	.102	.070	0.144
Occupation Pension	.034	.100	0.733
Occupation social security	-.110	.123	0.373
Occupation other	.028	.102	0.782
Urban	.068	.022	0.002
Number of children	-.133	.033	0.000
Partner	-.041	.073	0.578
Education Category	-.030	.021	0.147
cut1	-1.027	.215	
cut2	-.462	.214	
cut3	-.042	.213	
cut4	.590	.213	
cut5	1.252	.215	
cut6	1.992	.219	
Number of observations	1301		
Log Likelihood	-2382.0		
LR test	$\chi^2(10) = 51.48$		0.000
Pseudo-R^2	0.01		

Tab. 2.2: Ordered probit model: explaining public attitudes towards individual responsibility for retirement provision.

The variable 'Age' is not significant, we have investigated if there is a

non-linear relation between the variable 'Age' and the question about individual responsibility for retirement provision by adding the variables 'Age^2' and 'log(Age)'. However, these variables where not significant suggesting that there is no quadratic relation nor logarithmic relation. Also, replacing the variable 'Age' by two dummy variables for respondents that are between 45 and 65 years old and respondents that are older than 65 years does not reveal a relation between the age of the respondents and the question about individual responsibility for retirement provision.

Finally, we have investigated whether there is a non-linear relation between the variable 'Net Income' and the question about individual responsibility for retirement provision by adding the variables 'Net Income^2' and 'log(Net Income)'. The variables where not significant, suggesting that there is no quadratic nor logarithmic relation.

The LR test statistic of the model equals 51.48 and is $\chi^2(10)$ distributed under the null hypothesis that all variables together have no influence on the answer to the question about individual responsibility for retirement provision. The critical value for this distribution with $\alpha = 0.05$ is 18.31, which means that we can reject the null hypothesis. Thus, we can conclude that all the variables together do have a significant influence on the answer to the question about individual responsibility for retirement provision.

The pseudo-R^2 reported is the goodness-of-fit measure LRI. This measure equals

$$LRI = 1 - \frac{\ell(\widehat{\beta}^{ML}, \widehat{\alpha}^{ML})}{\ell(\widehat{\beta}_0^{ML}, \widehat{\alpha}_0^{ML})}$$

In which $\widehat{\beta}^{ML}$ and $\widehat{\alpha}^{ML}$ are the maximum likelihood estimators for β and α in the full model, while $\widehat{\beta}_0^{ML}$ and $\widehat{\alpha}_0^{ML}$ are the maximum likelihood estimators in the restricted model in which only the constant term is left as explanatory variable. The LRI has value 0.01. Note that the value of the LRI in general decreases as the number of choices in the dependent variable increases. This is due to the fact that the LRI is a ratio of two log-likelihoods, which are products of probabilities. In a model with more choices, the probabilities involved are smaller than in a model with fewer choices.

Due to possible correlation, some variables might have lower p-values when we regress fewer explanatory variables on the question about individual responsibility for retirement provision. Therefore, we will use a testing-down procedure meaning that we delete the variable with the lowest p-value from the set of variables and perform the ordered probit regression again with our new set of variables. We will continue deleting the variable with the lowest p-value until all variables are significant at a 10% level or until the Likelihood Ratio (LR) test rejects the null. We use a LR test to check whether all variables we deleted together are of significant influence on the answer to the question about individual responsibility for retirement provision. We found that all LR test statistics where lower than the critical value, meaning that we can not reject the null hypotheses. Therefore it seems justified to delete the variable with the lowest p-value until all variables are significant at a 10% level. This leads to the results

in table 2.3.

	Coef.	St. dev.	p-value
Age	.004	.002	0.051
Urban	.068	.022	0.002
Number of children	-.132	.031	0.000
cut1	-1.036	.149	
cut2	-.473	.146	
cut3	-.054	.146	
cut4	.576	.146	
cut5	1.236	.148	
cut6	1.973	.154	
Number of observations	1301		
Log Likelihood	-2385.9		
LR test	$\chi^2(5) = 43.67$		0.000
Pseudo-R^2	0.01		

Tab. 2.3: Ordered probit model: explaining public attitudes towards individual responsibility for retirement provision

Now we see that the variables 'Age', 'Urban' and 'Number of Children' are significant on a 10% level, while only the variables 'Urban' and 'Number of Children' are significant on a 5% level. Again, from the value of the LR test statistic we are able to conclude that all variables together do have a significant influence on the answer to the question about individual responsibility for retirement provision. The pseudo-R^2 (LRI) is the same as in the previous regression.

So far, in our investigation we have been using the ordered probit model. Also, we have regressed all explanatory variables on the question about individual responsibility for retirement provision using the ordered logit model. However, the log likelihood of this model, including all variables, is -2382.3, which is slightly smaller than the log likelihood of the ordered probit model (see table 2.2). This means that the ordered probit model gives a better fit and therefore we will only report the results of the ordered probit model.

2.4.2 Marginal Effects

As we have argued in section 2.2, the interpretation of coefficients of ordered probit models is not straightforward. Therefore, we will report the marginal effects of $P(y_i \leq 3|x_i)$, $P(y_i = 4|x_i)$ and $P(y_i \geq 5|x_i)$ for interpretation and comparison to the theoretical hypotheses. We will only report the marginal effects of the variables that survived the testing-down procedure.

In our theoretical hypotheses we have argued that we expect both 'Urban' and 'Number of Children' to have a positive influence on the respondents answer with respect to the question about individual responsibility for pension provisions. Thus, we expect that a rise in these variables (ceteris paribus) causes

	Government	Intermediate	Individual
Age	-.002*	-.000	.002*
Urban	-.025**	-.002**	.027**
Number of Children	.048**	.003**	-.051**

Tab. 2.4: Marginal effects of the results in table 2.3. Government captures the marginal effect of $P(y_i \leq 3|x_i)$, Intermediate captures the marginal effect of $P(y_i = 4|x_i)$ and Individual captures the marginal effect of $P(y_i \geq 5|x_i)$. ** means the marginal effect is significant at a 5% level, while * means the marginal effect is significant at a 10% level.

a respondent to incline more towards individual responsibility for retirement provisions. On the other hand, we expect older people to incline more towards government responsibility for retirement provisions.

Table 2.8 shows that the marginal effects of 'Urban' are in line with the theoretical hypothesis. The marginal effects of 'Age' and 'Number of Children' contradict the theoretical hypotheses, although the effects of 'Age' are not significant at a 5% level.

A possible explanation for the result that the variable 'Number of Children' has a positive influence on the respondents answer with respect to the question about individual responsibility for pension provisions is the following. People who have more children and thus a higher burden because of raising their children, maybe don't have money left to save for their own pensions while raising their children. Therefore, they might incline more towards government responsibility for retirement provision.

Also, as argued by Hayo and Ono (2009) the variable 'Age' might capture so called cohort and life-cycle effects. The cohort effect is the effect of the respondent being part of a particular cohort. The cohort members are influenced by the dominant norms of the society at their time of socialisation. The life-cycle effect reflects the changes in preferences occurring over a life-time. These effects might explain the contradiction of the theoretical hypothesis with respect to 'Age'.

2.4.3 Semi-nonparametric Estimation

As pointed out in section 2.2.3, the ordered probit and ordered logit models require distributional assumptions which might influence the outcome of the models. Therefore, we will estimate an extended ordered probit model in which the distributional assumption is replaced with the estimate of the unknown density of the error terms. The estimation of the extended ordered probit model is being performed in Stata, using the 'sneop' command proposed by Stewart (2003). The model was estimated for different values of K[1], table 2.5 shows the values of the Akaike Information Criterion (AIC), the Bayesian Information

[1] It is worth mentioning that the log-likelihood of the extended ordered probit model is not globally concave. It is therefore possible that maxima found during the maximization of the log-likelihood function are local maxima rather than global maxima.

Criterion (BIC) and the likelihood ratio tests for the different values of K. Both the AIC, BIC and likelihood ratio tests can be used to select the appropriate model. In case of the information criteria, the different models can be ranked according to their AIC or BIC values, with the model having the lowest AIC or BIC value being the best. Note that in the BIC, the penalty for additional parameters is stronger than in the AIC.

K	AIC	BIC	LR-test of K against K-1	p-value LR-test
OP	4795.94	4878.67		
3	4785.05	4878.13	14.88	0.000
4	4779.93	4878.18	7.12	0.008
5	4779.83	4883.24	2.10	0.147
6	4781.61	4890.20	0.21	0.643
7	4783.59	4897.35	0.02	0.900
8	4778.28	4897.21	7.31	0.007

Tab. 2.5: values of AIC, BIC and likelihood ratio tests for different values of K.

Table 2.5 shows that the LR-test rejects the ordered probit model. Also, both the AIC and BIC do not suggest selection of the ordered probit model. Therefore, it seems justified to conclude that the extended ordered probit model outperforms the ordered probit model.

The AIC chooses the extended ordered probit model with K=8 as the best alternative, while the BIC chooses the model with K=3 as the best alternative. The likelihood ratio test rejects the null hypothesis for $K \leq 4$, meaning that increasing K by one is of significant influence for $K \leq 4$. This suggests selection of the K=4 model. Because the AIC and BIC for K=4 are only slightly bigger than for K=8 and K=3 respectively, we decided to select the extended ordered probit model with K=4 as the best alternative.

Even though we selected the model with K=4, it can be shown that the models with different values for K coincide with the selected model. Except for the variable 'Partner', all models have the same significant variables. Also the sign of the significant variables coincides.

Table 2.6 shows the results of the extended ordered probit model for K=4. When comparing the results of the ordered probit model (table 2.2) with the results of the extended ordered probit model (table 2.6), we see a couple of differences. The main difference is the fact that the variable 'Partner' has a positive and significant influence at a 10% level, while its influence is negative and not significant at a 10% level in the ordered probit model. The log likelihood of the extended ordered probit model equals -2372.4, while the log likelihood of the ordered probit model equals -2382.0. The latter is smaller, meaning that the extended ordered probit model gives a better fit. The LR test confirms this, as the LR test statistic of the ordered probit model against the extended ordered probit model indicates that the extended ordered probit model outperforms the ordered probit model.

	Coef.	St. dev.	p-value
Age	.004	.004	0.264
Net Income	.128	.056	0.022
Sex Women	.020	.082	0.811
Occupation Pension	.054	.112	0.626
Occupation Social Security	-.060	.143	0.672
Occupation Other	.056	.120	0.642
Urban	.083	.026	0.001
Number of Children	-.195	.041	0.000
Partner	.137	.080	0.087
Education Category	-.020	.026	0.443
cut1	-1.027	Fixed	
cut2	0.297	.229	
cut3	1.772	.325	
cut4	2.966	.258	
cut5	3.673	.242	
cut6	4.358	.239	
Number of observations	1301		
Log Likelihood	-2371.0		
LR test (OP versus SNEOP)	$\chi^2(2) = 22.0$		0.000
Wald test	$\chi^2(10) = 51.2$		0.000

Tab. 2.6: Semi-nonparametric extended ordered probit model: explaining public attitudes towards individual responsibility for retirement provision

Finally, the Wald test statistic of the model equals 51.2 and is $\chi^2(10)$ distributed under the null hypothesis that all variables together have no influence on the answer to the question about individual responsibility for retirement provision. The critical value for this distribution with $\alpha = 0.05$ is 18.31, which means that we can reject the null hypothesis. Thus, we can conclude that all the variables together are of significant influence on the answer to the question about individual responsibility for retirement provision. Note that the Wald test statistic of the model is asymptotically equivalent to the test we used when discussing the ordered probit model, the LR test statistic of the model.

In the extended ordered probit model, the assumption $\varepsilon_i \mid x_i \sim N(0,1)$ *iid* is relaxed by estimating the real density of $\varepsilon_i \mid x_i$. The estimated density for K=4 exhibits a negative skewness of -0.875 and a kurtosis of 3.008, compared to a skewness of 0 and a kurtosis of 3 in the normal distribution. A negative skewness means that the peak of the distribution is at the right of the center, while a higher kurtosis means that the distribution is less flat. The standard deviation equals 1.908. Figure 2.2 gives a plot of the estimated density for the K=4 model.

As we have argued before, possible correlation of the variables might influence their p-values. Therefore we will use the testing-down procedure as described in section 2.4.1. All LR test statistics where smaller than their critical

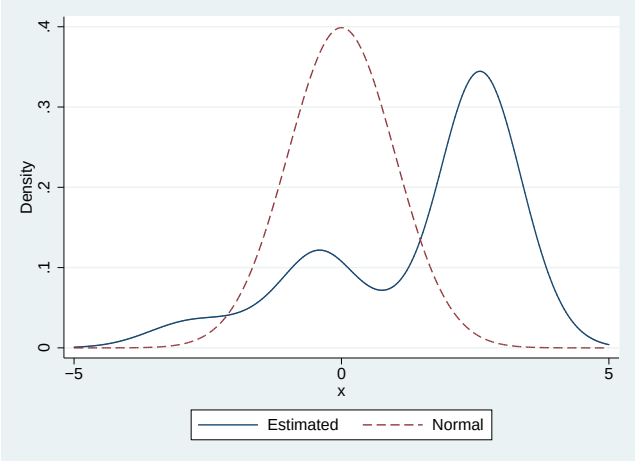

Fig. 2.2: Estimated density of the error terms for K = 4

values. This justifies deleting the variable with the lowest p-value until all variables are significant at a 10% level. The results of the testing-down procedure are shown in table 2.7.

	Coef.	St. dev.	p-value
Age	.006	.002	0.021
Net Income	.052	.031	0.095
Urban	.070	.024	0.004
Number of Children	-.157	.041	0.000
cut1	-.958	Fixed	
cut2	.115	.420	
cut3	2.160	.571	
cut4	3.136	.299	
cut5	3.759	.219	
cut6	4.393	.181	
Number of Observations	1301		
Log Likelihood	-2372.8		
LR test (OP versus SNEOP)	$\chi^2(2) = 24.2$		0.000
Wald test	$\chi^2(4) = 22.5$		0.000

Tab. 2.7: Semi-nonparametric extended ordered probit model: explaining public attitudes towards individual responsibility for retirement provision

The results of the testing-down procedure indicate that, at a 5% level, the variables 'Age', 'Urban' and 'Number of Children' are of significant influence on the answer to the question about individual responsibility for retirement provision. These results coincide with the results of the testing-down procedure of the ordered probit model shown in table 2.3. However, the variable 'Net Income' is

significant at a 10% level, while it was not significant in table 2.3. Furthermore, the extended ordered probit model obtains a higher log likelihood than the ordered probit model, which indicates a better fit. Again, from the values of the LR test and Wald test we are able to conclude that the extended ordered probit model outperforms the ordered probit model and that all variables together are of significant influence.

2.4.4 Semi-nonparametric Estimation and Marginal Effects

The β coefficients in the extended ordered probit model do not have a clear interpretation, neither can they be compared to the β coefficients in the ordered probit model. Therefore we will calculate the marginal effects of $P(y_i \leq 3|x_i)$, $P(y_i = 4|x_i)$ and $P(y_i \geq 5|x_i)$ for interpretation, comparison to the marginal effects of the ordered probit model and comparison to the theoretical hypotheses.

Because the marginal effects are not implemented in the 'snoep' command, they are calculated using the programming language in Stata. As described in section 2.2.2, the marginal effects can be calculated by multiplying the β coefficients with the value of the density in the point $\alpha_j - \beta' x_i$. We were able to obtain the density function from the 'sneop' file. Using this density function and the command predictnl in Stata, we obtained the marginal effects, their standard errors and p-values. The marginal effects are calculated for an individual with average characteristics.

The marginal effects are shown in table 2.8.

	Government	Intermediate	Individual
Age	-.001**	-.001**	.002**
Net Income	-.005*	-.015	.020*
Urban	-.007**	-.020**	.027**
Number of Children	.016**	.045**	-.061**

Tab. 2.8: Marginal effects of the results in table 2.7. Government captures the marginal effect of $P(y_i \leq 3|x_i)$, Intermediate captures the marginal effect of $P(y_i = 4|x_i)$ and Individual captures the marginal effect of $P(y_i \geq 5|x_i)$. ** means the marginal effect is significant at a 5% level, while * means the marginal effect is significant at a 10% level.

When looking at the marginal effects, it is remarkable that the marginal effects for Government are, in general, lower than in the ordered probit case, while the marginal effects for Intermediate are higher. This is due to the different slope of the estimated density in relation to the standard normal density.

The marginal effects for the variable 'Age' are significant at a 5% level, while they where not significant in the marginal effects of ordered probit. Besides this, the marginal effects of the extended ordered probit model are similar to the marginal effects of the ordered probit model.

In the theoretical hypotheses we have argued that we expect the variables 'Net Income', 'Urban' and 'Number of Children' to have a positive influence

on the answer to the question about individual responsibility for retirement provision. We expect the variable 'Age' to have a negative influence on the answer. The marginal effects of both 'Net Income' and 'Urban' coincide with the theoretical hypotheses. The marginal effects of 'Age' and 'Number of Children' contradict the theoretical hypotheses, as was also the case with the marginal effects of the ordered probit model.

3. CONCLUSIONS

In this thesis, we have investigated whether people with different socio-economic backgrounds tend to have different opinions about individual responsibility for pension provision. In the investigation, we made use of data from 1318 respondents of the Dutch CentERpanel as well as several different econometric models. The data consists of, among rich background information, answers to the question 'Some people think that in the future, individual responsibility for retirement provision will increase and people will more than now have to save themselves for their old age. Would you consider this a desirable or undesirable development?'. Respondents had to give an answer on a scale from 1 (very undesirable) to 7 (very desirable).

Because of the ordered nature of the question, we have used an ordered probit model to see if different characteristics have a significant influence on the answer to the question. The results of this model indicate that the characteristics age, number of children and whether the respondent lives in a rural or urban surrounding relate in a linear way to the respondents answer.

The marginal effects indicate that at a 5% level, the characteristic age is not of significant influence to the respondents answer. At a 10% level, the characteristic has a positive relation to the respondents answer, meaning that an older person is more likely to believe that an increase in individual responsibility for retirement provision is desirable, than a younger person. This is in contradiction with our a priori expectations, which might be due to the cohort and life-cycle effects.

According to the marginal effects, the number of children of the respondent have, at a 5% level, a negative effect on the respondents answer to the question about individual responsibility for retirement provision. This is in contradiction with our a priori expectations, but can be explained by the fact that people with more children have a higher burden and thus less money left to save for their pension provisions.

The area the respondent lives in is of significant influence at a 5% level. The marginal effects suggest that people living in rural areas incline more towards individual responsibility for retirement provision. This is in line with the a priori expectations.

From the extended ordered probit model we again find that the characteristics age, number of children and whether the respondent lives in a rural or urban surrounding are of significant influence to the question about individual responsibility for retirement provision. In addition to what we found in the ordered probit model, also the net income of the respondent is of significant influence.

A priori we expect respondents with a higher net income to be more likely to believe that an increase in individual responsibility for retirement provision is desirable. This is in line with what the marginal effects from the extended ordered probit model indicate.

Although several tests rejected the ordered probit model over the extended ordered probit model, the results of both models are more or less similar. We can therefore conclude that the results of the ordered probit model are robust against the distributional assumption.

BIBLIOGRAPHY

[1] Boeri, T. and G. Tabellini, 2005, "Does Information Increase Political Support for Pension Reform?", *CEPR Discussion Papers*, discussion paper no. 5319.

[2] Breyer, F. and J. Matthias Graf v. d. Schulenburg, 1987, "Voting on Social Security: The Family as Desicion-Making Unit", *Kyklos 40*, 529-547.

[3] Carnaghan, E. and D. Bahry, 1990, "Political Attitudes and the Gender Gap in the USSR", *Comparative Politics 22*, 379-399.

[4] Corneo, G. and H.P. Grüner, 2002, "Individual preferences for political redistribution", *Journal of Public Economics 89*, 897-931.

[5] Gabler, S., F. Laisney and M. Lechner, 1993, "Seminonparametric Estimation of Binary-Choice Models With an Application to Labor-Force Participation", *Journal of Business and Economic Statistics*, Vol. 11, No. 1, 61-80.

[6] Gallant, A. Ronald and Douglas N. Nychka, 1987, "Semi-Nonparametric Maximum Likelihood Estimation", *Econometrica 55*, 363-390.

[7] Hayo, B. and H. Ono, 2007, "Comparing Public Attitudes Towards Providing for the Livelihood of the Elderly in Two Aging Societies: Germany and Japan", *Marburg Papers on Economics 03/2007*.

[8] Hayo, B. and H. Ono, 2009, "Livelihood and Care of the Elderly: Determinants of Public Attitudes in Japan", *MAGKS Papers on Economics 12/2009*.

[9] Hendrickx, R.L.P., "Orientation ME/EM: Course Notes", 2008.

[10] Stewart, Mark B., 2003, "Semi-Nonparametric Estimation of Extended Ordered Probit Models", *Stata Journal 4*, 27-39.

[11] Stewart, Mark B., 2005, "A comparison of semiparametric estimators for the ordered response model", *Computational Statistics and Data Analysis 49*, 555-573.

[12] Page, B. I. and R. Y. Shapiro, 1983, "Effects of Public Opinion on Policy", *The American Political Science Review 77*, 175-190.

[13] Welch, S. and J. Hibbing, 1992, "Financial Conditions, Gender and Voting in American National Elections", *Journal of Politics 54*, 197-213.

[14] Wenger, G. C., 2001, "Introduction: intergenerational relationships in rural areas", *Ageing and Society 21*, 537-545.

[15] Wolfowitz, J., 1942, "Additive Partition Functions and a Class of Statistical Hypotheses", *The Annals of Mathematical Statistics 13*, 247-279.