

Retirement savings by self-employed: a behavioral investigation

An investigation of the relation between retirement behavior and individual characteristics of self-employed.

Matthijs de Snoo

Master Economics of Management and Organization
Erasmus School of Economics, Erasmus University Rotterdam

Student number: 343732as

Supervisors:

dr. J. Delfgaauw (professor of Economics at Erasmus School of Economics)

drs. L.V. van Andel AAG (senior consultant at Towers Watson)

Abstract

The pension system in the Netherlands is based on regular employed employees. Self-employed have to arrange their pensions by themselves. The saving opportunities and incentives are much smaller for self-employed due to government regulation. Since self-employed have to fix their after retirement income by themselves, the pre-retirement behavior is very important. This thesis combines economic theories and behavioral insights to find an answer on the question if, in the case of pension provision to self-employed, an individual approach will be more profitable than a collective approach. Therefore this thesis investigates the differences in pension payments, savings and motives between different groups of self-employed based on their characteristics. Data of three large datasets, containing financial and behavioral information has been used. The savings, motivation to save and the pension payments are investigated. Personal characteristics like level of education, being the main wage earner and individual income variability affect the saving behavior and financial position of self-employed. The motivation to save for retirement differs between self-employed owing to their behavioral characteristics as well as to the perception of the personal income situation relative to previous years and relative to other self-employed. Finally, the implications for the pension providers are discussed.

Contents

1.	Introduction.....	4
1.1.	History and current situation of labor market and pension system	4
1.2.	Retirement and self-employed.....	5
1.3.	Gaps in current literature.....	6
1.4.	Research question.....	7
1.5.	Short introduction to hypotheses, data and methodology	8
2.	Theoretical framework.....	9
2.1.	Concepts	9
2.1.1.	Self-employed	9
2.1.2.	Pensions.....	10
2.1.3.	Pension plans	11
2.2.	Groups and characteristics of self-employed.....	12
2.3.	Theories	16
2.3.1.	Traditional Economic Theory	17
2.3.2.	Bounded Knowledge	18
2.3.3.	Recent Behavioral Insights.....	19
2.4.	Literature.....	20
2.4.1.	Income and assets	20
2.4.2.	Financial knowledge.....	21
2.4.3.	Case studies and policy analysis	21
3.	Hypotheses.....	22
3.1.	H1 – subgroups and the pension payments.....	22
3.2.	H2 & H3 – pension savings.....	23
3.3.	H4 – pension motives	24
3.4.	Overview hypotheses	25
4.	Data.....	28
4.1.	SHARE	28
4.2.	ZZP-panel.....	29
4.3.	DHS Data	30
5.	Methodology.....	33
6.	Results.....	35
6.1.	Hypothesis 1 – subgroups and the pension payments.....	36
6.1.1.	Comparison regular employed to self-employed.....	36

6.1.2.	Education	38
6.1.3.	Pension funds	38
6.1.4.	Industries	39
6.1.5.	Regression pension payments.....	41
6.1.6.	Conclusion hypothesis 1	41
6.2.	Hypothesis 2 – subgroups and the pension savings.....	42
6.2.1.	Education and financial literacy	42
6.2.2.	Industries	43
6.2.3.	Main wage earners	45
6.2.4.	Age	45
6.2.5.	Pension funds.....	46
6.2.6.	Conclusion hypothesis 2	48
6.3.	Hypothesis 3 – pension savings over time.....	48
6.3.1.	Measurement of savings over time	49
6.3.2.	Industries	50
6.3.3.	Regression and subgroups based on education.....	53
6.3.4.	Conclusion hypothesis 3	57
6.4.	Hypothesis 4 – pension motives	57
6.4.1.	Financial literacy	58
6.4.2.	Age	59
6.4.3.	Mutations in income	60
6.4.4.	Alternative method: relative retirement saving motivation	61
6.4.5.	Conclusion hypothesis 4	64
7.	Conclusion and remarks.....	65
7.1.	Main findings	65
7.2.	Discussion.....	66
7.3.	Limitations and further research	67
8.	Bibliography.....	69
9.	Appendices	74
9.1.	Appendix 1 - Subgroup composition	74
9.2.	Appendix 2 – different regression models paragraph 6.3.3	76
9.3.	Appendix 3 – Classification of saving motives.....	77
10.	Data disclaimers	78

1. Introduction

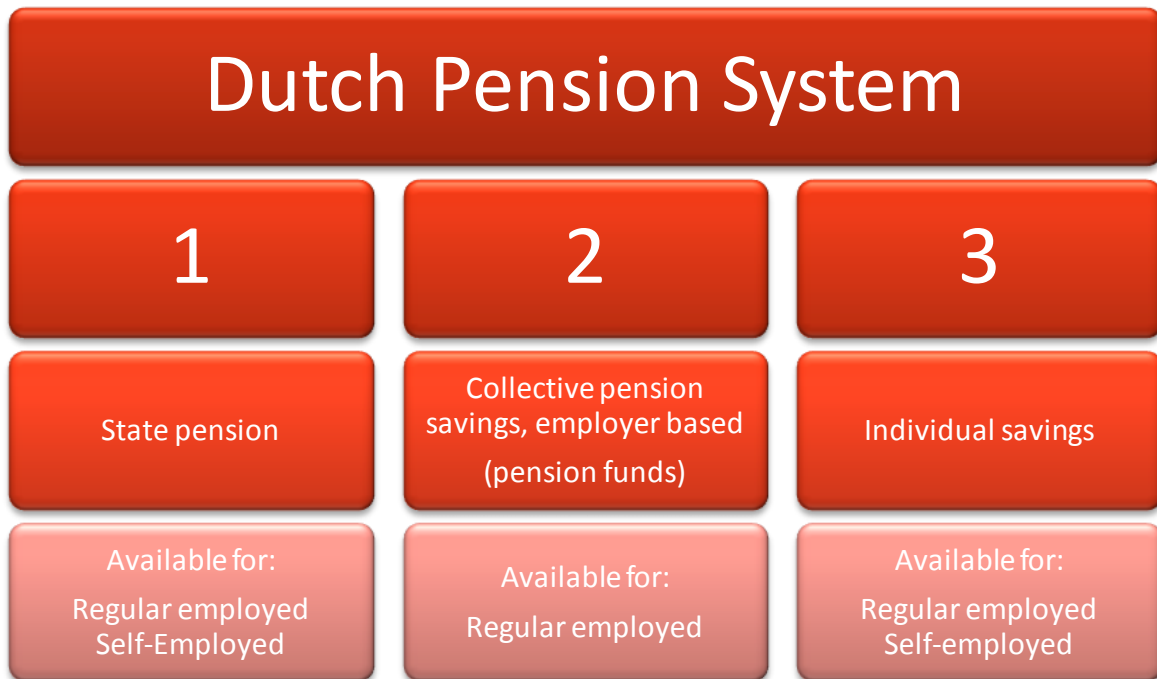
1.1. History and current situation of labor market and pension system

During the last decades the amount of self-employed people (hereafter self-employed) increased, in 2013 42,000 Dutch inhabitants became self-employed (Ministry of Finance, 2014). The Dutch government estimates the amount of self-employed above 800.000 (Josten, Vlasblom, & Vrooman, 2014). In the Nineties, 6% of the Dutch labor force was self-employed. This percentage increased to 10% nowadays and this trend will presumably continue in the near future (Ministry of Finance, 2014). Independency of individuals, specialization of employees and last but not least the financial crisis caused a decrease in the amount of regular employed employees and an increase in the amount of self-employed (SER, 2010). Due to this kind of changes in the labor market, the economy adapted. Furthermore, the characteristics of regular employed and self-employed differ (Josten, Vlasblom, & Vrooman, 2014).

Parallel to the increase in the amount of self-employed, the pension system expanded over time and became a sluggish institute that had to adapt to new situations (Van der Westen, 2014). Retirement programs evolved into more individual and complex systems. The evolution of the labor market caused and causes a lot of effects on the pension system. This research investigates the opportunities for an approach based on individual characteristics of Dutch self-employed instead of a general and collective approach.

The Dutch pension system is based on the idea of providing a certain income after retirement (Pensioenfederatie, 2010). Pension payments are the payments done by a fund, bank or other financial entity when the retirement period begins. In the pre-retirement period, people can save for the pension payments. They can save in regular accounts (banks, pension funds, insurers), but also in unconventional accounts, like investing in real estate. The government encourages saving in the regular accounts by fiscal policy. The Dutch pension system can be divided in three 'pillars', first the state pension, the second pillar contains the collective pension savings via the employer and the third pillar contains individual savings. The collective pension schemes are additional pensions to the state pension and are organized by the employers in cooperation with mainly pension funds. To hold the system in control, pension fund arrangements in the second pillar can only be supplied by these employers. However, the self-employed do not have an employer and therefore they cannot participate in the second pillar. Simplified, the Dutch self-employed have to arrange their pension payments in the third pillar autonomously.

Figure 1: Dutch pension pillar system



The funds in the second pillar collect contributions from employers and employees, invest this money on the market and pay the retired with the returns. Collectivity is an important element of this pillar. Since self-employed cannot participate in the second pillar, the scale advantages of the second pillar cannot be obtained by self-employed. The third pillar is accessible for self-employed since this pillar contains the individual pension plans. “Anyone can purchase a product in the third pillar to meet his/her requirements. In this way, people can save extra pension, often taking advantage of tax benefits” (Pensioenfederatie, 2010, p.13). An important disadvantage of the third pillar is the lack of collectivity. In the second pillar you share financial risks with a lot of other participants while you bear your own risk in the third pillar.

1.2. Retirement and self-employed

Within the group of self-employed, you can make distinctions based on the level of entrepreneurship of self-employed. For some groups of self-employed, you can doubt the entrepreneurship. For example, dentists are self-employed while the clientele is more or less constant and income does not depend of entrepreneurship. They will not be challenged to perform, like other self-employed who have to compete with peers and have to acquire customers. Furthermore, the group of dentists can be defined. The combination of these characteristics has resulted in the idea of establishing a pension fund for dentists while the law normally the establishing of pension funds for self-employed prohibits. Currently, self-employed can ask the government to allow the establishing of a specific pension fund for a precisely limitable group of self-employed. These groups are compulsory for the

self-employed who practice the defined job. This option of compulsory pension funds for self-employed can be regarded as an exception. In the Netherlands, eleven of such funds exist and around 90.000 self-employed are enrolled in these funds (Unie van Beroepspensioenfondsen, 2012). This thesis focuses on the group of self-employed who are not enrolled in a compulsory fund. Sometimes the self-employed outside these funds are compared to the self-employed who participate in such compulsory funds.

1.3.Gaps in current literature

There is a lot of literature on self-employed and their retirement savings¹. A common result in these papers is that actual savings for old age pensions are lower for self-employed compared to regular employed people since saving is not compulsory for self-employed. Some papers find also behavioral aspects as reasons for lower retirement savings, e.g. self-employed procrastinate savings. Furthermore, solidarity issues will arise if self-employed can use the same possibilities as regular employed (provided that self-employed holds a certain level of freedom, so they can cheat the regular employed by switching between funds). A lot of these papers state or assume that self-employed have on average less financial knowledge compared to regular employees. Since they have less financial knowledge, they do not fully understand the imaginable problems, if they do not save enough for an acceptable old-age pension. The differences between regular and self-employed can be stronger since the regular employed are automatically enrolled in second pillar pension funds, so the literacy effect will be restrained. These differences are behavioral differences. The effect of the behavior differs between regular and self-employed since self-employed are responsible for their pension savings while in the case of regular employed the employer bears this responsibility mainly.

An important weakness in the current literature is generalization. The papers treat the self-employed (more or less) as a homogeneous group of people. However, you can doubt the similarities between, for example, painters, plumbers and dentists and between high or low educated self-employed. Recently, Mastrogiacomo and Alessie (2014) found a disparity between saving intentions and the actual savings, the savings fall often below the intended sum. Furthermore, “finding the retirement motive important does not directly translate in additional retirement savings” (Mastrogiacomo & Alessie, 2014). While the dissimilarity between motivation and actual actions is fascinating, there is still an important limitation in research. Mastrogiacomo and Alessie took also all the self-employed together. It is conceivable that highly educated self-employed realize the necessity of savings better, while low-educated self-employed think of the matters of the day instead of planning their pension.

¹ The theoretical framework section will provide an overview of the literature and discusses these papers more deeply. For clarity reasons, the citations of the papers can be found in the theoretical framework instead of a few rows of citations in this paragraph.

Another assumption (or bias) is the idea that self-employed who earn a lot of money use the opportunity to save more compared to self-employed that earn just enough to buy 'food and housing', but cannot save for their pension. Mastrogiacomo and Alessie do not correct for this effects. Genre, Kohn and Momferatou (2011) and Dickens and Katz (1987) showed differences in wages between industries. Cole, Paulson and Shastry (2012) presented such differences due to the level of education, these effects should be incorporated in the framework of Mastrogiacomo and Alessie. This thesis investigates if there are differences between such subgroups and considers the desirability of a more individual approach instead of a general approach. Simplified, there is a lot of research about savings for retirement. These studies consider the whole group of workers or the whole group of self-employed, so the found effects are unspecific due to this generalization.

1.4. Research question

This thesis focusses on the self-employed and investigates the individual distinctions between self-employed in order to find information about their retirement behavior. Based on the behavioral insights, the desirability of individual or group arrangements for self-employed can be discussed. So the research questions are:

- 1. What are the differences in pensions between subgroups of self-employed?**
- 2. Is an individual approach more desirable than a general approach in the case of the pensions of self-employed?**

The pensions are split into the categories savings, motivation to save and payments after retirement. The groups are composed in different ways. There are groups based on industry, on level of education and based on income changes. To common sense, if you are confronted with a decreased income, you will limit your expenses and savings. Catching up of the skipped savings will possibly be absence. In the case of regular employed, the savings (second pillar) will increase if income grows due to the agreements between pension provider and employer. Since the self-employed are not enrolled in the second pillar, the effect of limited savings due to income reductions, will possibly be stronger than for regular employed. Also a distinction will be made between main wage earners and non-main wage earners. The latter can relative save more for their retirement since the non-main wage earner does not bear the household expenses fully. Furthermore, groups based on age will be compared. I expect an increase in pension savings and motivation to save if the retirement date comes sooner. I assume that a self-employed procrastinates his/her savings until the awareness arises that he/she has to arrange a pension to obstruct the situation of no pension payment at all. The groups will be more deeply discussed in the section about groups and characteristics.

1.5.Short introduction to hypotheses, data and methodology

To find an answer on the research question, four hypotheses will be tested. In short, I will investigate whether there are differences in pension payments (H1) and pension savings (H2) between subgroups. The third hypothesis investigates the saving patterns over time and the fourth hypothesis discusses the motives to save for retirement. So the root of the pension behavior of self-employed will be investigated by the third and fourth hypotheses.

Table 1: Hypotheses

H1:	There are no differences in annual gross pension payments after retirement between the subgroups.
H2:	There are no differences in pension savings (as percentage of income) between the subgroups.
H3:	The savings (assets) for pension (as percentage of income) vary with the same pattern as income over time.
H4:	There is no difference in the average pension saving motive between subgroups.

The first and second hypothesis investigate the behavior in the pre-retirement period. If the payments and savings of self-employed do not deviate from the values of the regular employed and if the deviations within the group of self-employed are small, an approach based on such subgroups will be a waste of time. However, if the self-employed in a certain subgroup (e.g. low-educated) behave in a different way compared to their regular employed counterpart, focusing on these subgroups by pension suppliers can be fruitful.

In the second pillar, the saving percentages are more or less fixed per income level, regardless the change of income of individuals. When the income increases, savings increase and when income decreases, the savings decrease. Since the second pillar is only accessible for regular employed, self-employed have to adapt their saving percentages by themselves. The third hypothesis investigates the saving behavior of the self-employed after income changes. The motivation to save for retirement does not depend necessarily on income or actual savings. However, if income changes or if the characteristics differ between self-employed, the motivation to save for retirement can differ. The fourth hypothesis tests whether the motivation differs between self-employed.

I use three databases to describe the behavior of the self-employed. First of all, the data provided by the research institute CentERdata is used. The set, called DNB Household Survey (DHS) is built up by a longitudinal panel study. Annually around 2,000 households are requested to fill in a questionnaire

about, inter alia, health, income and work. Also a lot of questions about the pensions and saving intentions are included. Second, I use the European dataset SHARE (Survey of Health, Ageing and Retirement in Europe). It contains information about health, income and pensions. More than 85.000 Europeans who passed the age of fifty filled in the survey. This dataset provides information about the value of pensions of (former) self-employed. Third, I will use the ZZP-panel, a dataset provided by Panteia/EIM. This dataset contains financial and personal information of self-employed over the years 2010-2013. Annually around 3,000 self-employed in the Netherlands participated.

SHARE is used for the first and fourth hypotheses, DHS for the third hypothesis and the ZZP-panel is used for the second and fourth hypotheses. To test these hypotheses the Mann-Whitney test will mainly be performed and sometimes regressions are used. This test provides information about the distribution of the values per group. For example, similarity of pension savings between subgroups can be tested. The regressions are used to investigate the economic significance. The details of these tests are discussed in the methodology section.

In the second section the concepts, composition of subgroups, theories and literature will be discussed. Section three elaborates the hypotheses. Section four discusses the data in more detail. The fifth section designs the methodological route to the results in the sixth section. This thesis ends with conclusions, recommendations, and remarks in the seventh section.

2. Theoretical framework

This section consists of four parts; first some concepts will be discussed. The second part explains the subgroups, the third part discusses some theories and the fourth discourses the literature. This thesis is based on several papers and theories. None of these focuses on an approximation based on subgroups within the total population of self-employed. However, these papers and theories regard elements that can explain together the potential differences between subgroups. The ideas, insights and interpretations will be used together to answer the question if the characteristics of self-employed affect retirement savings.

2.1. Concepts

2.1.1. Self-employed

The main concept of this thesis is 'self-employed'. There are a lot of definitions and criteria to be recognized as self-employed. The European institutes use the definition "persons who are the sole owners, or joint owners, of the unincorporated enterprises in which they work, excluding those unincorporated enterprises that are classified as quasi-corporations. Self-employed persons are

classified here if they are not also in a paid employment which constitutes their principal activity: in that latter case they are classified as employees” (Council European Union, 1996). “The self-employed are often taken to be individuals who earn no wage or salary but who derive their income by exercising their profession or business on their own account and at their own risk” (Parker, 2004, p.6). Compared to regular employed people, in these definitions differences appear in the payment structure, in the responsibilities and in the bore risks. Another important characteristic is autonomy, like Kalleberg (2000) mentioned. He stated that self-employed “have neither an employer nor a wage contract and are responsible for their own tax arrangements” (Kalleberg, 2000). In this thesis the self-employed will be defined by a person who runs a business by his own or with a few subordinates. Furthermore, this individual has to arrange the finance of his business himself and there is no employer relation with anyone else. Consequently, the individual self-employed bears all the risk. Since the self-employed do not have any relation to finance or legal departments, they have to arrange e.g. the pension savings and contracts by themselves while this kind of ‘beside the job’ issues will be arranged for regular employed by the employers.

2.1.2. Pensions

The disputes about the definition of pensions are smaller than the debate about the definition of a self-employed. Pension can be defined as the ‘after retirement period’. Pension payments are the payments done by a fund, bank or other financial entity in the after retirement period. In the pre-retirement period, people can save for the pension payments. They can save in regular accounts (banks, pension funds, insurers), but also unconventional accounts, like investing in real estate, can be used.

Currently in the Netherlands the retirement age will be raised; therefore data till 2013 will be used and the (common) retirement age is 65. Apart from that, more recent data shows a lot of gaps and is unusable. While the definition of pensions is more or less simple, the actual pension system is ‘a bit’ complex. To set a certain life expectancy and use this age for the calculation of the pension payments, you have to collect a lot of data about the individual. Due to (privacy) legislation, information about the health status of an individual cannot easily be obtained. Therefore the metric experience mortality will be calculated and there is a common use of this metric. Financial risks (exchange market, interest rates) as well as individual risks (accidents, exceed the life expectancy) should be included. To diminish such risks, pension funds created different pension plans and components for such plans. In some cases, participants can compose an individual additional pension plan that they desire. Individuals can also save for their pension by investing in real estate (e.g. buying a house). When pension savings are discussed in the next paragraphs, the composition of the pension will be mentioned.

Also the pension system can be divided in different components. The Dutch system contains of three pillars. The first pillar is the general old-age pension. This state pension can be seen as basic income. Everyone who passed the retirement age and lived in the Netherlands between his age of 15 and 65² received this payments, the quantity is linked to the statutory minimum wage (Pensioenfederatie, 2010). Collective pension schemes are stored in the second pillar; companies organize an additional pension to the state pension in cooperation with mainly pension funds. These funds collect contribution from employers and employees, invest this money on the market and pay the retired with the returns. The second pillar is marked by the company aspect and collectivity advantages. You cannot participate in the second pillar without a company. So self-employed cannot participate in the second pillar, excepting the situation of a by the government accepted compulsory fund for a specific group of self-employed. In contrast to the second pillar, the third pillar is accessible for self-employed since this pillar contains the individual pension plans. "Anyone can purchase a product in the third pillar to meet his/her requirements. In this way, people can save extra pension, often taking advantage of tax benefits" (Pensioenfederatie, 2010, p. 13). An important disadvantage of the third pillar is the lack of collectivity. In the second pillar you share financial risks with a lot of other participants while you bear your own risk in the third pillar.

2.1.3. Pension plans

The offered pension plans in the Netherlands can roughly be divided into two groups: the defined benefit (DB) and the defined contribution (DC) plans. Defined benefit plans ensure the participant of a certain annuity after retiring while the defined contribution plan does not. "In defined contribution plans, future benefits fluctuate on the basis of investment earnings and the employee's benefits during retirement depend on the contributions made to and the investment performance of the assets in his or her account" (Hodowanic & Rubin, 2013). So the difference between DB and DC is the structure of risks and savings. To implement international retirement plans in the Netherlands, the idea of personal pension insight (PPI³) is initiated. During the working life, the participant deposit premiums in their individual investment account. "On the retirement date, the PPI transfers the accrued pension capital to the insurance company or pension fund selected by the participant, which then purchases a life-long retirement pension (and partner pension, if required) with this capital. The pension provider then periodically pays out a pension to the retiree (or dependent)" ("Robeco", 2014). Since the financial information is presented in an easy understandable format, the participant can modify the plan based on his individual preferences. The advantages of the PPI are uncomplicated communication and the modification opportunities.

² The age of 65 will change due to the new legislation about the retirement age.

³ Dutch: 'premiepensioeninstelling', for example the Towers Watson PPI. <http://www.towerswatson.com/nl-NL/Services/Services/towers-watson-ppi-nl>

In the Netherlands, pensions plans are offered mainly by specific funds called 'Pensioenfond's'. These funds collect the member's contributions to invest on and get return on investment. Some companies have their own fund (company pension funds, e.g. Douwe Egberts and Unilever have corporate funds), other funds are specialized in certain industries (industry-wide pension fund, e.g. civil servants fund ABP). Furthermore, the compulsory funds for groups of self-employed mentioned above exists. Like insurers, a pension fund supplies a basic product and, possibly, gives the opportunity to buy additional components to get a higher pension payment. There is a lot of regulation about the minimum contribution of companies to funds, coverage ratios and communication to participants⁴. This regulation is ordered to diminish risky behavior by funds and ensure that funds can pay the promised payments. In this thesis, the presence of pension funds is essential. Specific characteristics of different funds will not be discussed and can be found in other literature.

2.2.Groups and characteristics of self-employed

During the last centuries, the labor market has changed hugely. Driven by social development, the rate of labor force participation by women in the United States increased by 40 percent points in the twentieth century (U.S. Department of Commerce, 2011). Borjas (2013) links this trend to the decreasing labor force participation of men. Furthermore, the paper by Kalleberg (2000) reviews nonstandard contracts like part-time work and independent contracting. He found an increase in the proportion of part-time workers. These changes in the society and labor force affect the increase of the amount of self-employed (Van Es & Van Vuuren, 2010). In some industries, the self-employed are more represented, mainly in the agriculture and construction sector (Hipple (2010) and SER (2010)). Van Es and Van Vuuren (2010) presented that the 'distribution' over industries differs between countries. In the case of employees, the employer arranges inter alia insurances and pension savings. Due to the independence of self-employed, they have to organize this sort of things by themselves.

For employees as well as self-employed, there is little or no 'pension awareness' (SER, 2010). A lot of regular employed (around 90%) participate in a pension plan (SER, 2010). However, in the case of self-employed, the default of such an arrangement is absent. So if they want a certain level of income after retirement, they have to fix that for themselves. The self-employed use other arrangements (assets and continuation of work) to fix their pension income (Geijtenbeek, de Rijk, & Rienstra, 2012). Due to different incomes of self-employed (industry-effects, time-effects, economic environments), due to the increasing amount of self-employed, due to the low pension awareness and due to the (little) legalized pension saving opportunities it will be interesting if different arrangements (different

⁴ Inter alia the General Old Age Pension Act, Financial Assessment Framework and Pension Act.

to groups, ages etc.) for their pension payments will be fruitful to fix the certain level of income after retirement.

A few groups will be discussed in this paragraph. One group will be discussed more extensively to show the inevitable differences between self-employed. A simple distinction can be made between the self-employed who participate (like the dentists in the introduction) and them who do not participate in compulsory pension funds. Participating will presumably result in higher pension payments since participants outsource the savings instead of postponing and of hesitating about the best saving plan (Kok, Baarsma, & Heyma, 2014).

Self-employed as well as regular employed can be divided into industry categories like service sector and agriculture. Inter alia Genre, Kohn and Momferatou (2011) and Dickens and Katz (1987) show differences in wages between industries. The selection of industries differs between several papers, but all of the papers find industry effects on wages due to, for example, the male/female ratio and the presence of part-time work. Since wages differ between industries and retirement savings and pension payments are related to wages (Steenbeek & Van der Lecq, 2007), differences between industries in retirement savings and pension payments can be expected.

Another industry effect is the establishing of pension funds for specific groups of self-employed. In the case of these pension funds for self-employed, only individual professionals within a specified field of activities (e.g. dentists, pharmacists and pastors) are enrolled. For example, the Netherlands self-employed painters decided to start a pension fund where participation is compulsory. There are eleven of such funds in the Netherlands, mostly in the healthcare industry (Unie van Beroepspensioenfondsen, 2012). The reasons behind the establishment of this kind of pension funds are based on a common sense that a collective fund diminishes risk, creates scale opportunities and gives the opportunity to get professional consults ("MinSZW", 2013). Risk can be reduced since the participants can spread risks due to larger (aggregate) and diverse investment portfolios. Administrative costs and other overhead expenses can be distributed among members and they can collectively get (expensive) professional consults instead of getting no advice due to the high hourly wages of consultants. Besides these economic reasons, there are some historic explanations. Formerly, doctors and other physicians were seen as individual practitioners. So they had to manage their own company, including the funding of their pension. As mentioned in the introduction, you can doubt the entrepreneurship of such individual practitioners since they have a certain and defined group of (returning) clients, marketing is not needed and there is few competition in those markets. Hence, you can discuss the individual (entrepreneurial) component of their work and why these self-employed are not regularly employed, including a pension plan. About 90% of the Dutch self-

employed is not enrolled in the abovementioned compulsory pension funds ("MinSZW", 2013). These employees have to seek for individual retirement programs. Van Els, van Rooij and Schuit mentioned this problem clearly: "The question is whether they [the self-employed, AMdS] are capable of doing this or whether a form of mandatory participation must be introduced for them" (Van Els, Van Rooij, & Schuit, 2007, p. 163).

The difficulty to define subgroups of individual professionals, the risk and solidarity concerns and the absence of a continuous income stream (volatility of earnings) are the main arguments against compulsory funds. If the benefits of pension funds for self-employed exceeds these costs (the counterarguments), obliging participation in these funds can probably be justified. Consequently, it can be reasonable to oblige these self-employed to participate in such compulsory funds. A counterargument is the limitation of the independency (and liberty) of the self-employed in the case of compulsory pension plans (Kok, Baarsma, & Heyma, 2014). Economic theories about the establishing of such compulsory pension funds are lacking. Hence the different wages between industries and the differences in pension agreements cannot be explained by economic theory. However, the differences are real, so this state of the world induces the question how to optimize the current situation of self-employed. An industry specific approach can be a fruitful method owing to the differences between separate industries mentioned above. Separating subgroups of self-employed based on industry can shed light on the desirability of specific nudges for separate subgroups.

The self-employed who are not the person in the household that pays the rent and foods and takes care of tax declarations due to a partner who pays these costs, can also be recognized as a different group. If the partner of this self-employed has a pension scheme, the self-employed can join the revenues of the partner's plan. If the partner saves for a certain level of collective income after retirement, this self-employed have a lower incentive to save for his pension. Another reason for the distinction between main wage earners and non-main wage earners is linked to the intention to save for retirement. On average, the uncertainty of adequate savings is much higher in the case of main wage earners than of non-main wage earners. This uncertainty results in a lower intention to start (or increase) saving (Van Schie, Donkers, & Dellaert, 2012).

Beside time issues, other characteristics can affect the saving incentives. The level of education can have an effect on savings via income. The relation between financial knowledge and financial performance is investigated (e.g. Lusardi & Mitchel (2007) and Hilgert, Hogarth & Beverly (2003)). However, there is less published about the relation between education levels and financial performance. The higher educated a person is, the more financial 'performance' they have (Cole,

Paulson, & Shastry, 2012). Cole et al. found also a positive relation between years of education and receiving a pension. This thesis tests if these results hold in the case of self-employed. Finally, higher educated people earn on average more than low educated (Borjas, 2013). So the education can indirectly (via income) affect pension savings and payments. Therefore subgroups based on education levels will be used⁵. Also variability in income over time will be measured; subgroups will be based on positive or negative income changes since I assume a decrease in savings after a decrease in income while the growth in income will not result in catching up. Finally, subgroups based on age will be added since I assume an increase of savings over age, as mentioned before.

Some characteristics are less personal since they concern the job features. Major shareholder managers can be recognized as self-employed, a lot of these managers invest in their company. After retirement, they sell their company and buy a pension annuity. So the pensions 'are in their companies', therefore the retirement savings can be totally different compared to other self-employed.

To summarize, groups based on age, industries, education, being main wage earner and income variability over time will be used. Participants can be classified in more than one group (e.g. high educated main wage earner). Hence comparing of self-employed between different subgroups will be hard. The figures 2 and 3 show the classification of the labor force and the categorizing of self-employed in the subgroups. By using these subgroups, I will investigate where pensions differ due to industry, education and so on.

⁵ In the case of regular employed, the distinction between high- and low-educated employees can be meaningless since the pension saving plans are shaped by the consultation between social partners and employers and these plans are based on level of income. So the effect of education can be meaningless in the case of actual savings. Nevertheless, differences in motivation can still hold.

Figure 2: Classification of labor force

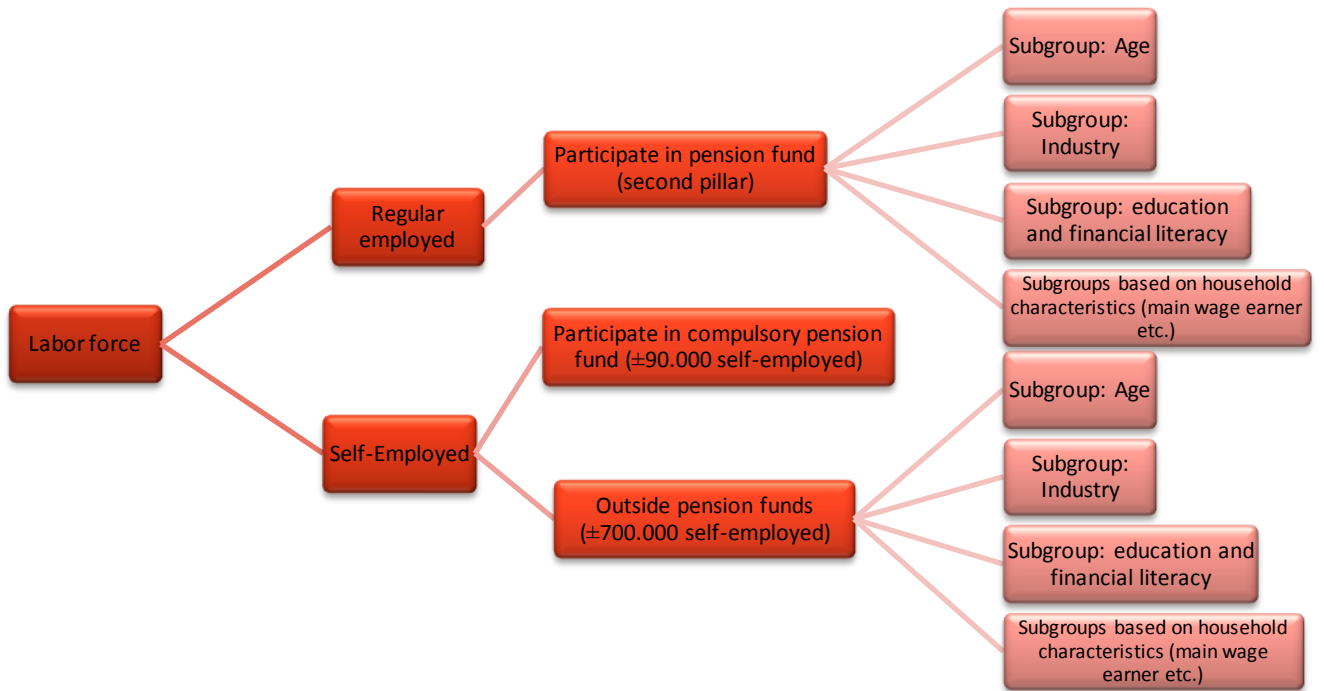
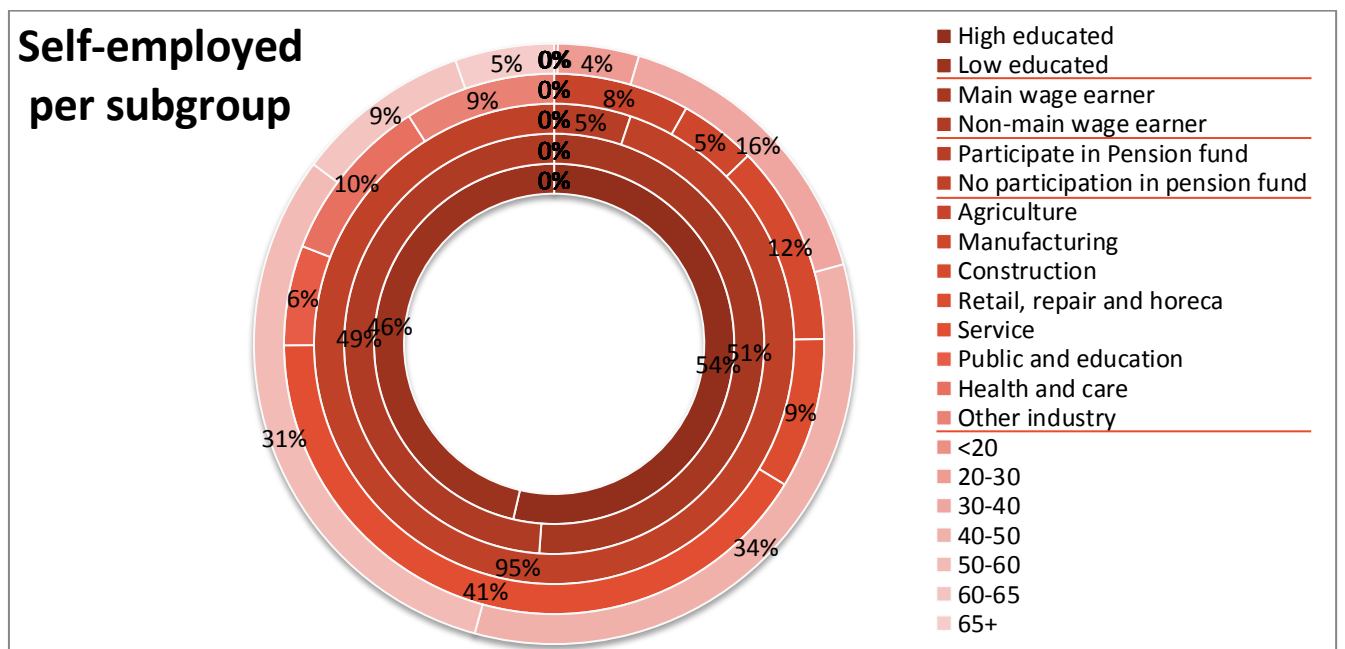


Figure 3: Percentage of self-employed per subgroup

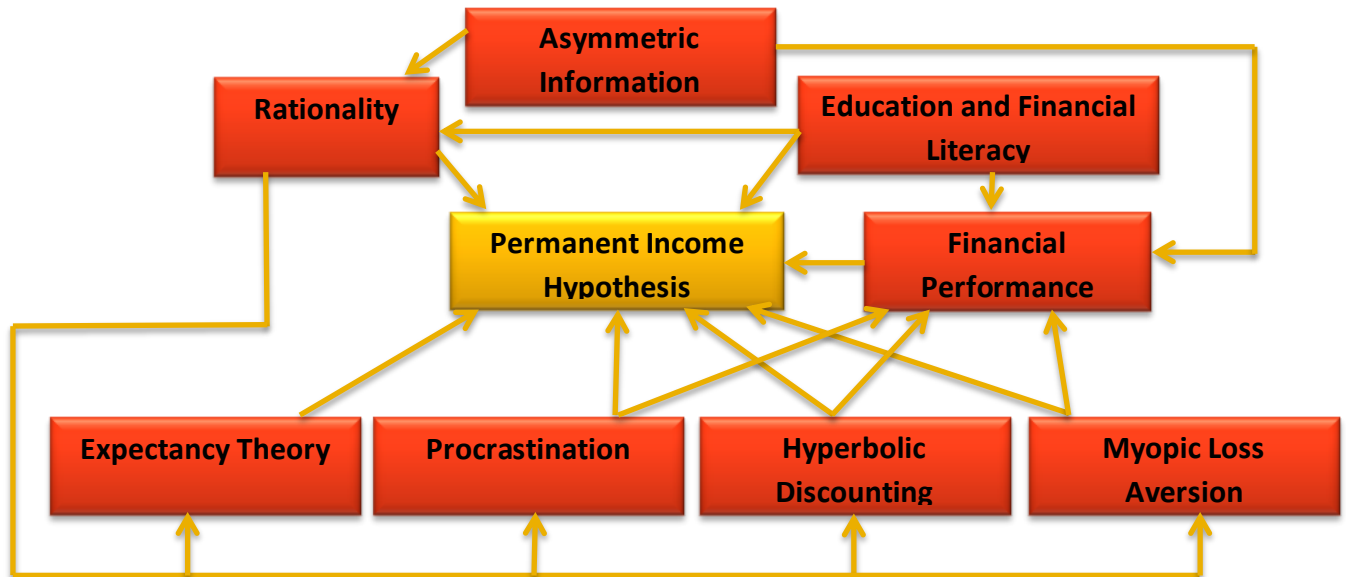


2.3.Theories

The theories used in this thesis can be split in three subsections. The traditional economic theory, the bounded knowledge ideas and recent behavioral insights. These theories are all linked to the permanent income hypothesis. This theory will be explained in the next paragraphs and can be

summarized as the idea of a more or less stable income during the whole life. Figure 4 shows the relations between the separate economic and behavioral ideas.

Figure 4: Relation between theoretical concepts



2.3.1. Traditional Economic Theory

One of the basic ideas in economics is the rationality assumption. The idea that human beings act rational is explained and used in a lot of the existing literature (Wilkinson, 2008). (In particular the game theory uses the rational choice theory, (Von Neumann & Morgenstern, 2007 (1953)). The participants of voluntary pension programs can be considered as players in such a game.) For example, if a person prefers good A over product B and good B over product C, then he prefers A over C. The idea behind this consistent ranking is the optimization; we assume that a person wants more instead of less of a certain good. Ross summarized these rationality issues as follows: “Economic rationality implicates two aspects of agency: (i) consistency of choice from one consumption or investment opportunity to another, and (ii) use of full information in arriving at beliefs about the relative expected payoffs from possible choices (so called ‘rational expectations’)” (Ross, 2010, p. 28).

In the twentieth century, Friedman introduced the ‘permanent income hypothesis’ (Friedman, 1957). This theory assumes that people prefer a stable consumption over time, the spendable income (sum of wages and income of assets) should also be more or less the same (ceteris paribus). Since people cannot work forever after retirement, they have to save before retirement. The level of savings depends on the expected consumption (and other needs) after retirement. This thesis tests if the consumption smoothing of self-employed affects the savings for income smoothing. Id est, if there is a negative shock in the income of self-employed, they can postpone their savings to smooth their

consumption nowadays while this action affects the income smoothing in the long run if the postponed savings are not compensated later⁶. If you combine the rationality to Friedman, people have to make rational choices about how to save and invest for the optimal pension payments. Hereafter, consumption smoothing as well as permanent income hypothesis refer to the longitudinal idea of savings and consumption during the consumer life cycle.

Since the sources of income differ over time (before and after retirement), Friedman's theory cannot be applied fully. In short: the lifetime of people can be divided roughly in three parts. The first part of life will be spent on learning, the middle part on working and during the last years of the life people are retired. During their working-life, people have to save for retirement. If we simplify the reality, two periods left (working and retirement). Subsequently, the Hirshleifer model can be used to calculate the needed savings during the working career (Hirshleifer, Glazer, & Hirshleifer, 2005). The interest should be taken into account due to inflation and profits of the investment. If we use Hirshleifer's mechanism, we should know the duration of the periods, the long-term preferences and (financial) opportunities of the individual. The duration of the (remaining) career is needed to calculate the minimum amount of savings per year and can be based on the experience mortality and retirement age. The preferences are needed to calculate the minimum payments to satisfy the needs of the individual after retirement. The limited opportunities can be seen as the income restriction.

If people act irrational in one of the three stages, the outcome will be positively or negatively affected (the latter was described by inter alia Mastrogiacomo and Alessie (2014)). Furthermore, due to the large periods, risks increases and therefore actions should be taken cautiously (Ross, 2010). Some theories describe the most frequent reasons of inconvenient retirement payments; these are hereafter discussed.

2.3.2. Bounded Knowledge

Commonly known is the absence of the rationality and Friedman's theory in retirement due to the lack of financial knowledge of people and the uncertainty due to the duration of pension plans. These drawbacks can be linked to some other theories about human decision making under risk, discounting and uncertainty. Some people try to invest in the right way, but decide in the wrong way. The financial knowledge of (self-) employed affects their financial performance and also affects their pension saving performance (Lusardi & Mitchell, 2007).

⁶ This idea looks like the policy recommendation of Keynes to save during booms (Keynes, 1965 (1936)). A lot of governments save (reduce public debt or cut their bills) too little during such economic periods (Stockhammer, 2012). The same holds possibly for self-employed.

In another paper by Lusardi and Mitchell, they present in the theoretical framework a list of papers concerning the relation between life cycle models, financial knowledge and the performance (Lusardi & Mitchell, 2013). Based on the research by Jappelli and Padula (2011), “financial literacy and wealth will be strongly correlated over the life cycle, with both rising until retirement and falling thereafter”⁷ (Lusardi & Mitchell, 2013, p.4). Due to financial inability of (self-) employed, their financial performance will be lower and therefore the pension payments are affected.

The literacy is related to asymmetric information. If (self-) employed are financial unknowledgeable, they are dependent of financial professionals to get optimal advices. However, Van Rooij et al. found a positive correlation between high literacy and relying on financial advisors (Van Rooij, Lusardi, & Alessie, 2011). Furthermore, they found a positive relation between low literacy and relying on family and friends for financial advices. Additionally, there are a lot of empirics concerning the drawbacks of naïve people who bought usurious products. Naïve buyers would buy the most profitable products which fits their life cycle model without knowing the details of the financial product⁸. The asymmetric information gap can be bridged by increasing financial literacy. Probably, some (self-) employed are not insured since their financial ability is at a high level. In short, the pension saving intentions of the self-employed can be admirable while the optimal actions are absent.

2.3.3. Recent Behavioral Insights

The last decades, the scientific collaboration between the economic and psychological discipline increased. There is an increasing interest in behavioral economics. One of the behavioral theories is the expectancy theory. The expectancy theory discusses the behavior of people in the way that people base their actions on the expected outcomes of those actions (Vroom, 1964). In the case of pension savings, people want a certain level of pension payments and modify their saving actions to attain the payments. People can use the expectancy theory and discount the values by using the Hirshleifer’s model. Then the person knows what he/she should do, but these theories do not predict what he/she really does. A self-evaluation is (more or less) included in the expectancy theory, but the real effects are absent. The insights of procrastination and hyperbolic discounting try to describe the real behavior.

Procrastination in economics can be described as the process of knowing which actions are needed nowadays to achieve a certain result without taking those actions. In the case of retirement savings, procrastination can occur in different ways. People procrastinate, for example, to choose the best

⁷ The fall in literacy can possibly explained by the decrease of cognitive skills over time after retirement (Korniotis & Kumar, 2011).

⁸ In the Netherlands, insurers and banks are accused of selling such financial products called ‘woekerpolissen’, these products are opaque and usurious priced. The game theory, and in particular the Decision Maker-Advisor Game, is related to this asymmetric information game (Wrasai & Swank, 2007).

saving opportunities and most optimal investments (O'Donoghue & Rabin, 1999). These problems can probably be solved by correcting the behavior by setting other defaults (e.g. compulsory minimum savings). However, such actions are paternalistic and the justification can be doubted. Especially in the case of self-employed compared to regular employed, two sides appear: firstly procrastination is easier for self-employed since they do not save compulsory for their pension payments. Secondly, paternalistic regulation impairs the idea of independency directly. Finally, if you postpone savings, the time that your savings are profitable are smaller so the final yield will be lower.

Hyperbolic discounting is related to procrastination and describes the discounting process. In the case of hyperbolic discounting, the discount ratio is higher for nearer periods than for the future (Wilkinson, 2008). So, loosely speaking, in the case of retirement savings, the 'revenue' of postponing savings will be higher than the 'costs' of the lacked yield in the end. Comparably, myopic loss aversion occurs also in the case of employees. Myopic loss aversion can be described as the idea that investors value losses more than gains and evaluate long term investments frequently. To decide about investments, the investor has to calculate a long-term rate of return to diminish the variation of the rate of return (smooth the positive and negative rates). They prefer "long-term rather than one-year rates of return" (Benartzi & Thaler, 1999).

2.4.Literature

In the past years, the number of papers about self-employed increased. Some papers discuss the income and asset composition of self-employed; others investigate the relation between characteristics like financial literacy and retirement savings in general. This section discusses also some case studies.

2.4.1. Income and assets

The starting point of the report by Geijtenbeek, De Rijk en Rienstra (2012) is the finding of lower pension incomes for self-employed compared to regular employed. They investigated the assets of self-employed and tried to answer the question if the assets compensate lower pension incomes. A lot of self-employed continued their business after retirement. The authors tested if this continuation is caused by lower pension payments. They found a difference in the variation of income growth over time. Regular employed have a more stable growth while the curve for the self-employed is very steep in some years. The consumption levels of self-employed had to adapt. To hold a comparable level of income and consumption compared to regular employed, the self-employed have to use their assets. The continuation of business is not a significant argument for avoiding 'income disasters'. Some self-employed continued to confine the income reduction after retirement. This

sheds light on the differences in the composition of the net income after retirement due to differences in the pre-retirement situation.

Another view on the income composition is mentioned in the paper by Campbell, Cocco, Gomes and Maenhout (2001). They found that “some households – particularly self-employed college graduates – are exposed to much greater volatility in their labor income than are typical households” (Campbell, Cocco, Gomes, & Maenhout, 2001, p. 471). Meager and Bates (2001) used UK data about self-employed to find patterns in income and savings. Self-employed are confronted with “losses of pension entitlements and low savings potential during self-employment” (Meager & Bates, 2001, p. 55). So the self-employed have to seek other ways to keep a certain income after retirement. Vroonhof and De Vries (2011) found that the self-employed are aware of this problem, but adequate actions are absent.

2.4.2. Financial knowledge

Cole, Paulson and Shastry (2012) study the link between knowledge and financial behavior. They found that an increase in education causes a reduction of the probability of a bankruptcy. Furthermore, “for retirement investments, an additional year of schooling increases the probability of non-zero income by about 5.9 percentage points” (Cole, Paulson, & Shastry, 2012). They found also a positive relation between years of schooling and receiving a pension. The financial situation (and opportunities) differs between regular- and self-employed. If the ‘knowledge-effect’ also occurs in the case of self-employed, probably the effect will be strengthened due to income aspects. Also Hilgert and Hogarth (2003) investigated the relation between (financial) knowledge and behavior. They found a positive relation between IQ and actual savings. Having a saving account (four-fifths of households) does not lead into regular saving (less than fifty percent saves regularly). Bernheim (1998) found for the same group the result that they save too little for retirement. A lot of other papers (e.g. the paper by Lusardi and Mitchell (2007), mentioned below) discuss the relation between knowledge, financial behavior and savings, but the combination of these three demonstrate the significant effect of education on retirement savings.

2.4.3. Case studies and policy analysis

The Dutch Ministry of Social Affairs and Employment researched why the pension prospects of self-employed compared to employees differ from each other (“MinSZW”, 2013). Due to differences in the characteristics (e.g. obliged retirement savings, various income schemes), some concerns will arise in the case of pension funds for self-employed or cooperation in the current second pillar. First, if self-employed can decide to opt-in in a pension fund, they can seek the best opportunity and diminish risk at the expense of regular employed who cannot decide in which fund they participate

(solidarity issues) (Kok, Baarsma, & Heyma, 2014). Second, opposites should be considered. On the one hand collectivity advantages of pension funds and on the other hand the existing opportunities of individual pension savings. Also the distinction between compulsory and voluntary participation should be bridged. The authors of the report by the Ministry discussed the policy implications and implied that the net income situation of self-employed after retirement is urgent, but modification of the system requires a profound discussion about the obliging of pension savings for self-employed. If the Dutch government decides to take action in the field of pension systems for self-employed (other than 'ZZP-Pensioen', cfr. De Jong (2009) and Van der Lecq & Oerlemans (2009)), discussions about incentives (health insurances and age issues) will arise (cfr. Zissimopoulos, Maestas, & Karoly, 2007).

The Dutch organization EIM (Economisch Instituut voor het Midden en Kleinbedrijf, research institute, nowadays merged into Panteia, a policy and market research institute) investigated the pension savings by self-employed (Vroonhof & de Vries, 2011). They found, like Mastrogiacomo and Alessie (2014), a difference between intended and actual savings. 47% of self-employed did not save for their retirement in 2009 and around 40% did not expected to be able to make ends meet (Vroonhof & de Vries, 2011). This comprehensive report has an important drawback: the expectations were discussed while actuarial information is lacking.

All these studies implied a comparable pension for self-employed as for employees. This statement is explicitly expressed in Choi (2009). "The main issue would be how to build equitable pension systems for the self-employed and employees, while providing the self-employed with enough retirement income security" (Choi, 2009). Another motive for researching self-employed is the suggestion by Beugelsdijk and Noorderhaven (2005) about differences in characteristics between self-employed and regular employed. They found, besides a 'more individual orientation' and an 'ethic of working hard', some social differences. For example, self-employed are convinced that individual effort should be greater incentivized (Beugelsdijk & Noorderhaven, 2005).

3. Hypotheses

3.1.H1 – subgroups and the pension payments

The regular employed people in the Netherlands are usually enrolled in pension funds. Self-employed have to organize a pension plan by themselves. Consequently, the gross pension payments per year can be different across self-employed due to the individual responsibility to save during their working life. A lower level of the payments for the self-employed is imaginable due to the expectable procrastination of savings in the case of non-compulsory saving (some literature discussed parts of this behavior, in the descriptive book by Steenbeek and Van der Lecq (2007, p. 168) an overview is

included). Furthermore the financial literacy can affect the savings and given the myopic loss aversion, the understanding of saving opportunities will be different (Kahneman & Tversky, 1979). This research investigates if different characteristics (based on the theories mentioned above) of self-employed explain the differences in pension behavior (the composition of savings, motives and payments).

Some of the papers mentioned above found differences in these payments, but none of the papers checked if there are differences between different subgroups of self-employed. Since there are differences in income between subgroups (Genre, Kohn, & Momferatou, 2011), these differences probably continue to exist after retirement. Therefore the first hypotheses will test if there are differences in the monetary value of the gross pension payments per year between subgroups of self-employed. Differences in the pension payments between subgroups can shed light on the necessity of an overall pension fund for self-employed. I expect some differences between the subgroups due to the level of financial knowledge during their working life. In particular, I expect higher levels of income for former service sector self-employed and physicians. Furthermore, the ownership of assets can cause different incentives to save for pensions. If you can sell your assets (for instance the firm) and buy therefore an annuity e.g., there is no reason to save a lot before retirement. The division of the entire group of self-employed in subgroups can shed light on these issues. So the first hypothesis is intrinsically simple: there are no differences in the average pension payments per year between the subgroups.

3.2.H2 & H3 – pension savings

While the first hypothesis tests the pension payments, the second hypothesis will test the savings. Hypotheses two is: there are no differences in the average pension savings per year between the subgroups. The second hypothesis tests a certain moment in time (snapshot), the third hypothesis is about the actual savings over time. Some papers discussed the heterogeneity of the income streams of self-employed (Campbell, Cocco, Gomes, & Maenhout, 2001). That can be seen as a fact and reason for lower savings in some periods. The input of pension funds depends on the contribution of their participants. If the participants have a relatively constant income, they can deliver a constant contribution to the funds. It is hard to maintain funds without constant contributions. Probably, the annual savings by dentists are more homogeneous due to less variation in income over time. In the case of musicians, whose business depends of a lot of factors, the savings can differ due to the presumable income heterogeneity. The costs of living can be the same over time, therefore the willingness to save for retirement will probably be the balancing item: the savings will be fully correlated to positive and negative income changes. If there is a difference in homogeneity of savings between subgroups, two main effects will arise. First the opportunity for pension funds to establish plans for

self-employed that have a more or less homogeneous income. Due to this homogeneity, the risks will be lower (the savings can be the same over time). Secondly, specific plans for the other subgroups can be invented. The savings differ over time and, perhaps, the saving motives are not fully rational. Hence, the risks for the suppliers of such pension plans are higher. Another issue for the latter group can be the demarcation of this group. Dentists and painters can be classified due to their license (if necessary) and defining of those jobs is quite simple. Defining the job of a self-employed artist can be hard. How can you point out the difference between an amateur artist and a professional?

Finally, the second hypothesis will test if there are differences in the savings between the subgroups. I expect differences due to the combination of economic state of the world (conjuncture), level of education, industry and being main wage earner. For example, if you are a low-educated and main-wage earner self-employed in the construction sector (economic sensitive sector), lower savings can be the case compared to a 'high-educated full-time non-main wage earning financial service sector consultant'. If there are no differences at all, dividing the whole group in subgroups will not be profitable.

As mentioned before, the third hypothesis tests the saving pattern over time. Probably, if a respondent faced a fall in income, he reduced his savings but omits to increase his savings after an increase in income. If this is the case, the theory about consumption smoothing will be violated. The third hypothesis tests if there is any difference at all in the homogeneity of the pension savings. I expect a violation of the economic theories since lowering the savings is very easy and penalties for non-increasing of the savings are absent. Since self-employed can switch between subgroups of self-employed, the hypothesis will be tested on the subgroup level as well as in general. Hence, the hypothesis will be that the savings (assets) for pension (as percentage of income) vary with the same pattern as income over time.

3.3.H4 – pension motives

The last hypothesis tests the saving intentions for retirement. Mastrogiacomo and Alessie (2014) found a disparity between the intentioned and necessary savings. Due to income variation over time, as mentioned above, the actual savings can differ but the intentions will probably be the same. While the minimum pension levels can be discussed, the necessity of a certain pension is common sense. If there are differences in the intentions between subgroups, then the awareness of the necessity of a certain pension can be an issue. For example, if the younger respondents have lower saving intentions compared to the older respondents, the theories about procrastination and hyperbolic discounting will be more acceptable. If some groups demonstrate such low intentions, government and pension providers can decide to invest in awareness campaigns.

Lusardi published many times about financial literacy and retirement. She mentioned repeatedly that financial literacy is endogenously related to wealth in the retirement period. “People display different levels of financial knowledge early in life, and this simple feature has important implications for how much people save” (Lusardi, Michaud, & Mitchell, 2011). The amount of savings and the motivation to save differ between people, inter alia due to the financial knowledge of the individual (self-) employed. Lusardi found a difference between the literacy and retirement behavior, “particularly among specific groups” (Lusardi, 2009). She made mainly the distinction between low- and high-educated and male and female. The fourth hypothesis tests if the saving motives differ between different groups of self-employed too. I expect differences in the motivation between (subgroups of) self-employed and regular employed based on the combination of the ideas of Beugelsdijk & Noorderhaven and Lusardi. Self-employed are more risk-seeking and ‘seize the day’ (Beugelsdijk & Noorderhaven, 2005). I expect, in the case of self-employed a lack of awareness of the necessity of retirements that affect the motivation to save. So the fourth hypothesis assumes no differences in the average saving motives between subgroups.

3.4. Overview hypotheses

To sum up, table 2 repeats the hypotheses and shows the relation between economic concepts and hypotheses. The combination of rejections of some hypotheses will shed light on the future of pension opportunities for self-employed. If some subgroups demonstrate a lack of awareness and of savings repeatedly, answering the research question will be obvious. Presumably, answering is more complex because of the combination of rejections of hypotheses. Consequently an answer like a consideration is more reasonable. This consideration can contain different options and suggestions.

Table 2: Function of concepts in the different hypotheses

Concept	H1	H2	H3	H4
	“There are no differences in annual gross pension payments after retirement between the subgroups”	“There are no differences in pension savings (as percentage of income) between the subgroups”	“The savings (assets) for pension (as percentage of income) vary with the same pattern as income over time”	“There is no difference in the average pension saving motive between subgroups”
Permanent Income Hypothesis	Is there a stable income after retirement to meet consumption needs?	Is the saving percentage adequate during the working life to continue a desired level of consumption after retirement?	Is there a correction of savings after income shocks to hold an average permanent income over time?	Are the self-employed motivated to save during their working life to continue a desired consumption level after retirement?
Rationality		People fully know and understand how much they have to save.	People fully know and understand how much they have to save.	
Education and Financial Literacy	Differences in payments due to level of education and financial literacy.	Higher literacy will presumably result in higher savings.	Higher literacy results possibly in lower variability of saving-ratio S_{it} .	Higher literacy results in smaller increase of motivation over age.
Financial Performance		Adequate saving percentage expected in case of high financial performance.	In case of high financial performance, lower variability of saving-ratio S_{it} assumed.	Higher performance results in lower increase of motivation over time.

Concept	H1 “There are no differences in annual gross pension payments after retirement between the subgroups”	H2 “There are no differences in pension savings (as percentage of income) between the subgroups”	H3 “The savings (assets) for pension (as percentage of income) vary with the same pattern as income over time”	H4 “There is no difference in the average pension saving motive between subgroups”
Expectancy Theory		Adequate saving percentage expected in case of high financial performance.	Corrections of falls in savings expected to achieve the predicted needed amount of money after retirement.	
Procrastination		Savings of respondent increase over age.		
Hyperbolic Discounting			Decrease of savings after income reduction without increase after income growth. Irrational savings.	Increasing motivation over age
Myopic Loss Aversion		Low savings expected if respondent is young		Increasing motivation over time.

4. Data

This section discusses the data used in this thesis. Since the data is mostly about financial information, the means and standard deviations lose explanatory power. The medians provide more useful information; hence the median is mentioned in the summary statistics.

4.1.SHARE

For the discussion of the first hypothesis, data provided by SHARE is used. SHARE is a European dataset, funded by inter alia the European Commission. This data contains information about the main occupation of more than 85.000 Europeans. Since 2000 four waves are published, each wave included information about extra countries and citizens. Sometimes, variables are added to get more information about the respondents. The fourth wave was published in 2013 and the data was collected in 2011 and 2012. This wave is used in this thesis. The data is split up in a subsample of European citizens who are self-employed (selecting variable *EP009_*) and are retired (selecting variable *EP005_*). Table 3 shows summary statistics about this fourth wave. Some data is modified due to incorrect answers. For example, respondent F1-693729-01 filled in an income of more than ten milliard euros, his values are set to zero. Other respondents exceeds the values of peers by far, this causes misinterpretation of the data. Therefore, I modified the data based on the boxplots so that the outliers will be changed in the value of the maximum of the 'normal' respondents. Furthermore, some respondents recorded values of '0.00' income while they had to register e.g. 'no income' or check 'not applicable'. These '0.00' values are also eliminated.

Moreover, by using these data one should take care of the cultural and economic differences between countries, especially the differences between the Northeast and the rest of Europe. Economic difference can be found in the amount of part time-workers or the average education level in a specific country. On average, retired get €6.887 state pension per year (Eichhorst, et al., 2011).

Table 3: Summary statistics of SHARE data

Subject	Wave 4		
	Mean	Median	Standard deviation
Retired self-employed	583		
Average age 2013	70	69	6
Male/female	386/197 (66.2%)		
Average years of education	11.45	12	5
Average pension payments	€39,670	€20,000	€85,923 ⁹

Country	Average age 2013	Average total annual pension payments	Median total annual pension payments	Standard deviation total annual pension payments	Total observations included
Austria	69	€91,219	€34,500	€200,765	73
Germany	71	€51,902	€50,400	€23,090	24
Sweden	71	€39,575	€35,887	€30,934	40
Netherlands ¹⁰	71	€9,922	€9,800	€4,199	32
Spain	73	€36,025	€15,450	€50,957	7
Italy	71	€33,549	€21,600	€41,364	49
France	70	€50,973	€49,850	€40,441	29
Denmark	70	€26,604	€24,154	€12,121	10
Switzerland	71	€50,567	€43,131	€42,341	55
Belgium	71	€59,569	€32,500	€97,544	39
Czechia	69	€12,025	€11,201	€9,629	106
Poland	65	€6,439	€4,873	€4,242	12
Hungary	65	€12,227	€9,114	€14,030	40
Portugal	70	€39,144	€24,000	€48,248	41
Slovenia	67	€45,258	€30,400	€34,856	19
Estonia	72	€14,227	€7,669	€19,040	7

4.2.ZZP-panel

The biggest dataset is the ZZP-panel and is used for the second hypothesis. The Dutch Ministry of Social Affairs and Employment and Panteia/EIM, a policy and market research institute, started a panel to investigate the long term trends of self-employed. In 2010 a group of self-employed was

⁹ Due to the inclusion of part-timers and of the great amount of low annual pensions the distribution is right skewed.

¹⁰ The values for the Netherlands contradict common sense values due to the relative large group of part time self-employed. The household income of the Dutch respondents does not stick out.

asked to fill in a survey. In 2013 3,000 self-employed participated and provided information about their business, income, savings et cetera. The questions about pensions (*PE00-PE15*) are specific and supply information about the percentage of savings, composition of savings and partner income after retirement. A drawback of this dataset is the lack of absolute values. To compare the data with DHS or SHARE is more complex, but still possible due to the relative values. Table 4 provides summary statistics of the self-employed in the ZZP-panel.

Table 4: Summary statistics of ZZP-Panel

Self-employed in the ZZP-panel	Panel 2010	Panel 2011	Panel 2012	Panel 2013
Average age	48	47	49	49
Number of participants	2,009	2,198	3,001	3,000
Male/female	1,436/573 (71.5%)	1,571/627 (71.5%)	2,118/883 (70.6%)	2,188/812 (72.9%)
Average net income per month respondent*	€1,455	€1,488	€1,510	€1,535
% household income earned by self-employed	N/A	58%	59%	62%
Mean percentage pension savings of (business) income**	10%	13%	13%	14%
Median percentage pension savings of (business) income	7%	10%	10%	10%
% participate in pension fund	19.2%	16.8%	14.3%	11,7%
% saves by themselves for pension income***	56.7%	64.4%	67.3%	64,2%

* Average standard deviation: €1,322

** Average standard deviation: 15%

*** The percentage of self-employed who said that they save by themselves, this does not mean that they actually save for their after retirement income.

4.3.DHS Data

Data provided by the research institute CentERdata is used for the third and fourth hypothesis. This data is built up by questionnaires. More than 2,000 households participate in this panel annually. In 2013 2,093 and in 2012 1,830 households participated. They answered questions about, inter alia,

income, health status, assets and welfare. Since this dataset contains a lot of information about labor, income and retirement, this dataset will be useful to reject or accept the third hypothesis. Table 5 shows summary statistics about the waves of the past five years. The differences between some years could be explained by the partially changed respondent group (so the difference between 2012 and 2013 exceeds 263 (2,093-1,830) households due to dropouts) as well as changes in the situation of particular household. There are small changes due to modifications in the questions because of amendments of the law.

Table 5: Summary statistics of DHS data

Subject	Wave 2009	Wave 2010	Wave 2011	Wave 2012	Wave 2013
# households	1,660	1,885	1,734	1,830	2,093
Mean net household income	€34,892	€36,974	€34,062*	€35,285	€35,513
Median net household income	€31,000	€33,000	€32,216*	€34,000	€33,000
net household income >€40.000 (percentage of respondents)	23.6%	26%	25.5%	30,1%	31,6%
Percentage of retired respondents	15%	19.8%	19%	19.1%	17%
Average percentage net pension income in comparison to the last earned net salary (of retired respondents)	76.96%	82.1%	79.71%	67,91%	68,1%
# Self-employed respondents	159	205	196	225	201
% completed university or higher professional education (percentage of respondents)	27.1%	29%	29.7%	29.5%	33.3%
Owner of current accommodation (percentage of respondents)	71%	73.9%	75.4%	73.6%	71.5%

* outliers eliminated (respondent who reported >1 Mld net income)

There are a lot of ways to select the group of self-employed. Question *BZR1* asks the respondent what their (main) occupation is. This question can be used to separate the self-employed and participants of partnerships (the Dutch Maatschap and VOF) in the whole group. A disadvantage is the amount of respondents, 1,572 (75%) filled in the assets and liabilities questionnaire that contains

question *BZR1*. 100% of the respondents answered question *BEZIGHEI* about the primary occupation, the option ‘free profession, freelance and self-employed’ can be used as separation variable. But since a part of the respondents did not fill in all questionnaires, the adequacy can be discussed. Mastrogiacomo and Alessie used *IZ1* to dissociate the self-employed of the total group of respondents. *IZ1* is part of the Health and Income questionnaire and asks if the respondent was “self-employed, working as a free professional or as a freelancer” in the last year. Mastrogiacomo and Alessie chose this variable since *IZ1* “is available for all waves” (Mastrogiacomo & Alessie, 2014). Since the third hypothesis discusses about twenty years, this argument of Mastrogiacomo and Alessie will also be used in this thesis. Table 6 provides summary statistics of the self-employed.

Table 6: Summary statistics of self-employed in the DHS data

Self-employed (<i>IZ1</i>) (percentages of self-employed respondents)	Wave 2009	Wave 2010	Wave 2011	Wave 2012	Wave 2013
Number respondents	159	205	196	225	201
Average age	51	51	55	52	51
% self-employed respondents at age >65	11.3%	12.7%	20.4%	17.3%	16.9%
Male/female	91/68 (57.2%)	120/85 (58.5%)	131/65 (66.8%)	143/82 (63.6%)	123/78 (61.2%)
% main wage earner of household	61%	66.3%	71.4%	68.4%	68.7%
% completed university or higher professional education	53.4%	52.2%	45.5%	51.5%	48.7%
Average taxable income respondent	€34,306	€32,213	€37,130	€32,309	€31,963
Average net household income	€31,400	€36,174	€34,657	€35,719	€35,708
Median net household income	€30,000	€30,000	€30,000	€33,600	€34,110
Standard deviation net household income	€27,969	€39,140	€35,891	€23,342	€24,373

The total savings can be counted by aggregating some partial variables in the dataset concerning bank accounts, deposits, savings, lifetime insurances, mutual funds, shares and other individual

saving actions. CentERdata (mistakenly) did not ask some questions (BDR3a to BDRL8) about employer-sponsored saving plans in the panel of 2013; therefore the values of 2012 are copied to the dataset of 2013. The impact of this imputation is negligible since these particular savings concern the regular employed. Only the self-employed who switched from regular employed to self-employed can probably be affected by this imputation. If respondents stated their savings in an interval, the average is used (for example, if the respondents save between 10,000 and 15,000, the expected savings are 12,500). In the case of savings above a certain minimum (“I save more than 25,000”), this minimum is the expected amount.

The DHS data provides also information about the motivation of respondents to save for inter alia, retirement. Since the motivation depends on the current and previous years, a proxy containing the average of the years 2010 till 2013 is estimated based on at least two observations per respondent. While the respondents with only one observation were eliminated, the total number of valid cases increased from 298 (only the wave of 2013) to 415 (waves 2010-2013). In the case of the third hypothesis, data of 1993-2013 has been used. The monetary values are presented in guilders for the first ten years. These are converted to euro's by the fixed rate €1 = 2.20371 (European Commission, 2011), so all values in this thesis are in euro's.

5. Methodology

To test the hypotheses, different statistical methods will be used. The number of observations differs; therefore the (commonly used) significance level of 5% loses explanatory power. However, for consistency reasons the 5%-level will still be used, nonetheless the economic significance of insignificant findings can be indicative. Before performing any test, outliers are determined and the normality of the distribution will be checked. Usually, the distribution is not normal.

The first, second and fourth hypotheses will test if there are significant differences between subgroups. The payments per subgroup will be compared to the payments of another subgroup by using a Kruskal Wallis Test since the distribution is not normal. The Kruskal Wallis Test shows if there are differences between the groups, but does not mention which groups differ. To compare two groups, the non-parametric Mann Whitney Test will be performed to test if the median of a certain group exceeds the median of another subgroup.

Before performing the Mann Whitney Test, some assumptions have to be satisfied. The groups are categorical and independent, also the observations are independent. Hence the low educated are not included in the high educated category for example. Sometimes a variable is modified into categorical variables (e.g. the main wage earners in the ZZP-panel). The dependent variables (pension

payments, pension savings) are continuous or ordinal variables, only this kind of variables can be the dependent variable in the case of Mann Whitney Tests. In the case of normal as well as non-normal distributions, the Mann Whitney Test can be used. But in the case of non-normality Mann Whitney Test fits better than the standard parametric t -test. The majority of the distributions show a skewing to the right; the mean exceeds the median due to large values. This is intuitively accountable: a negative pension payment is impossible while, in contrary, some people get very large payments.

To test if there are differences of one particular subgroup compared to all other participants, the total group will be reduced to a new group without the treated subgroup. Then you can use also the Mann Whitney Test (provided that the other assumptions are met). A table provides the p -values of equality of means. While the p -value states the statistical significance, the ratio will shed light on the economic significance. The ratio in the tables can be computed by the fraction of the mean rank of highest over the mean rank of the lowest. This Mann Whitney Test is used in the case of the first, second and fourth hypothesis.

A drawback of this calculation of the economic significance is the use of medians. Therefore, in the case of all hypotheses regressions will be used to test the *effect* of changes in income in previous years on actual savings. The OLS-method will be used and consequently the Gauss Markov assumptions should be satisfied (Verbeek, 2012). The basic assumptions are

$$(A1), \quad E\{\varepsilon_i\} = 0, \quad i = 1, \dots, N$$

$$(A2), \quad \{\varepsilon_1, \dots, \varepsilon_N\} \text{ and } \{x_1, \dots, x_N\} \text{ are independent}$$

$$(A3), \quad V\{\varepsilon_i\} = \sigma^2, \quad i = 1, \dots, N$$

$$(A4), \quad cov\{\varepsilon_i, \varepsilon_j\} = 0, \quad i, j = 1, \dots, N, \quad i \neq j$$

The fifth assumption (A5) can be used to check if the coefficient is normally distributed by checking the normality of the error terms.

$$(A5), \quad \varepsilon_i \sim NID(0, \sigma^2)$$

Since the time series are not always stationary¹¹, and normality cannot be assumed rashly, the Gauss-Markov assumptions are violated. This non-stationarity causes the problem of non-explanatory relationships and performing t -tests or F -tests will be invalid. Therefore the weak (asymptotic) assumptions will be used. So the assumption (A2) will be replaced by (A8) and (A5) by (A6).

¹¹ E.g. the income-series for the third hypothesis has a unit-root which can be eliminated by making the model autoregressive.

(A8), ε_i and \mathbf{x}_i are independent

$$(A6), \quad \lim_{n \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i' = \boldsymbol{\Sigma}_{xx}$$

By using these assumptions, it is possible to make some statements about the coefficients. In the test of the third hypothesis, the coefficients present the effect of the changes in income. So if the weak assumptions hold, conclusions about the effects of income changes on assets can be drawn. In the tables that contain the statistics, besides the (adjusted) R-squared the Bayesian Information Criteria (BIC) will be stated. The accuracy of this metric exceeds the precision of (adjusted) R-squared due to the penalty for the number of observations (Verbeek, 2012).

6. Results

This section discusses the results in the order of the hypotheses. Table 7 (next page) shows the descriptive data. Note that, for example, a respondent can be low educated as well as main wage earner and he/she is employed in the service sector. Hence, comparing respondents between incomparable subgroups is useless.

Table 7: description of data

Subgroups, average values	H1: gross annual pension payment after retirement	H2: pension savings (average percentage of business income)		H4: Pension motives (motivation on scale 1=very unimportant – 7=very important to save for retirement)	
			% high educated		% high educated
Dataset	SHARE	ZZP-Panel		DHS	
Low educated	€47,719	12.63%	0	5.6	0
High educated	€50,943	14.55%	100%	5.44	100%
Main wage earner	-	10.78%	42%	4.78	47.1%
Non-main wage earner	-	18.27%	15.6%	4.86	41,7%
< 50 years old	-	12.74%	44.1%	4.84	32.8%
≥ 50 years old	-	14.76%	48,7%	4.78	31.1%
Participate in pension fund	€98,442 (after elimination outliers €37,265)*	10.23%	40.6%	-	-
No participation in pension fund	€36,133	14.12%	21.6%	-	-
Industry: Agriculture	€41,881	11.22%	5.6%	5.06	0.0%
Industry: Manufacturing	€25,254	14.40%	15.8%	4.91	16%
Industry: Construction	€27,279	11.94%	8%	4.47	9.7%
Industry: Retail, repair and horeca	€47,686	17.72%	9.8%	5.2	16.1%
Industry: Service	€27,450	13.76%	28.4%**	4.92	23.5%
Industry: Public and education	€33,912	13.16%	**	4.7	32.2%
Industry: Health and care	€32,883	15.78%	**	4.8	22.1%
Industry: Other industry	€49,636	12.18%	33%	4.64	38.4%

* one respondent received annually more than one million euro's.

** The distinction based on level of education is hard due to overlap between the industries in the ZZP-panel, therefore the aggregated percentage of high educated service suppliers is calculated.

6.1.Hypothesis 1 – subgroups and the pension payments

6.1.1. Comparison regular employed to self-employed

The comparison of the subgroups will be split in different parts to overcome unimportant interactions. It is more fascinating to compare low educated to high educated instead of comparing low educated to e.g. 'health and care self-employed'. This eliminates also the interdependency

between subgroups. Furthermore, if the differences between regular employed and self-employed are small, a particular approach of self-employed is useless.

Table 8 shows the mean pension payments of self-employed and regular employed. Since in the case of self-employed the pension payments of the non-self-employed partner are included, the comparison is deceiving. Hence the values in the table are the sum of individual private pension payments.

Table 8: Comparison individual private pension payments (annual)

Subgroups	Self-Employed (593 respondents included)	Regular Employed (1,008 respondents included)
Financial illiteracy	€ 15,097	€ 18,100
Financial literacy	€ 16,044	€ 22,846
Low educated	€ 16,116	€ 18,402
High educated	€ 14,855	€ 14,134
Full time employed	€ 13,811	€ 14,743
Industry: Agriculture	€ 17,369	€ 8,998
Industry: Manufacturing	€ 11,676	€ 10,762
Industry: Construction	€ 19,146	€ 18,170
Industry: Retail, repair and horeca	€ 10,708	€ 23,517
Industry: Service	€ 15,426	€ 18,659
Industry: Public and education	€ 9,070	€ 15,326
Industry: Health and care	€ 37,577	€ 17,534
Industry: Other industry	€ 17,094	€ 9,571

The values in the green cells exceed the values in the red cells. The tests of significant differences between industries can be found in the next paragraphs.

The difference in private pension payments in the health and care sector can probably be explained by the difference between medical specialists (self-employed) and academic hospital personal (regular employed). The income levels of the hospital personal are regulated by the collective labor agreements while medical specialists can set the wages by themselves. The same kind of argumentation can be used to explain why the levels of payments of self-employed exceed sometimes those of the regular employed. The standard deviation in private pension payments in the case of self-employed exceeds the variance in the case of regular employed by 51.9%.

6.1.2. Education

Table 9: differences between education levels in retirement payments of self-employed

<i>p-values Mann Whitney Test ([first character of sector with highest mean rank of pension payments]; [ratio highest/lowest])</i>	Low educated (148 participants)
High educated (106 participants)	0.008*** (L; 1.22)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

The null hypothesis of equal pension levels between low educated and high educated can be rejected. The low educated earn on average less, however the median of low educated exceeds the median of the high educated. This difference can be explained by the fact that the standard error of mean in the cases of high educated is much higher, some high educated self-employed earn a lot (e.g. >€100,000). These drive the average without driving the median. The distribution of income of low educated is more compressed. The Mann Whitney Test provides evidence for the statement that the subgroups are not equal, low educated show higher levels of pension income. By checking these results in depth, other income parts like income streams from rents or sublets, alimonies and annuities compensate the lower level of high-educated. The differences in the mean pension income are more visible by testing if the respondent worked more than 30 hours per week in the last job. This full time proxy shows a lower mean pension income for low educated while the pension payments of the high educated doubled compared to the data of the full- and part-time self-employed together. High-educated can be part-time self-employed to fulfill individual needs without the necessity of earnings, so their activities are 'for fun'. Furthermore, the income from other household members is higher in the case of high educated and can explain the lower levels of their individual pension payments since e.g. their partner can supplement the income.

6.1.3. Pension funds

Being participant of a pension fund can affect the total pension payments per year. So the null hypothesis of equal levels of pensions between them who are enrolled and who are not enrolled in pension funds will be tested. In the Netherlands, all the pension funds are funds in which people can participate through the employer. So the self-employed who participate in a pension fund are either part-time regular employed or stopped their regular employment to become a self-employed (former regular employed). The Mann Whitney Test provides insignificant evidence for the statement that the subgroups are not equal. The test cannot prove the assumption that pension fund participants have a higher level of pension payments; the test shows that the median of the participants exceeds the median of the non-participants in more than sixty percent of the cases.

Table 10: differences in retirement payment levels of self-employed between pension funds participants

<i>p-values Mann Whitney Test ([first character of sector with highest mean rank of pension payments]; [ratio highest/lowest])</i>	No participation in pension fund
Participate in pension fund	0.385 (P; 1.16)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

In the Netherlands, participating in pension funds is only possible if the law allows participation. Since there is currently no option for the self-employed to participate in the second pillar, (unless compulsory funds) self-employed cannot benefit of the advantages of pension funds. Furthermore, solidarity issues will arise since self-employed can cheat the system (Bosch, de Graaf-Zijl, & van Vuuren, 2014). Therefore, a policy recommendation to self-employed to participate in pension funds is worthless.

6.1.4. Industries

In case of the comparison of the industries, respondents are assigned to the category of their last job. The assumption is that respondents did not switch between subgroups. Also should be noted that all the respondents are self-employed, so the differences between e.g. a regular employed accountant and a regular employed plumber do not need to be the same in the case of self-employed. Firstly, the Kruskal Wallis test on equality of subgroups is executed. The null of equality is not rejected since the p-value is 0.692. Table 11 shows the p-values and ratios for the industries. The self-employed that worked in another industry than the classified sectors have the highest average pension income (€49,636 annually). The Mann Whitney Test suggests that the self-employed of the public and education industry have the highest levels of pension income. The mean rank for the public and education industry is only in the cases of the comparison to the health and care sector and to other industries the smaller one. To conclude, the first hypothesis will be rejected for most of the cases. The null hypothesis will not be rejected for the comparison between other industries and classified industries. A caveat of this rejection is the generalization: the answer to the first hypothesis is based on European data. Therefore the disclaimer should be stressed that the situation in specific countries can differ (cfr. Van Es and Van Vuuren, 2010).

Table 11: differences between industries in retirement payments of self-employed

p-values Mann Whitney Test ([first character of sector with highest mean rank of pension payments]; [ratio highest/lowest])	Manufacturing	Construction	Retail, repair and horeca	Service	Public and education	Health and care	Other	Total minus subgroup
Agriculture	0.860 (M; 1.03)	0.889 (C; 1.02)	0.876 (R; 1.02)	0.822 (S; 1.03)	0.634 (P; 1.08)	0.907 (A; 1.02)	0.248 (O; 1.16)	0.466 (T; 1.09)
Manufacturing	-	0.919 (C; 1.02)	0.852 (M 1.03)	0.714 (M; 1.05)	0.897 (P; 1.02)	0.676 (M; 1.07)	0.28 (O; 1.16)	0.641 (T; 1.06)
Construction	-	-	0.922 (C; 1.01)	0.917 (C; 1.01)	0.631 (P; 1.074)	0.958 (H; 1.008)	0.342 (O; 1.12)	0.608 (T; 1.06)
Retail repair and horeca	-	-	-	0.853 (R; 1.04)	0.773 (P; 1.04)	0.729 (R; 1.05)	0.222 (O; 1.12)	0.495 (T; 1.06)
Service	-	-	-	-	0.514 (P; 1.09)	0.894 (H; 1.02)	0.141 (O; 1.15)	0.329 (T; 1.09)
Public and education	-	-	-	-	-	0.801 (P; 1.04)	0.607 (O; 1.06)	0.983 (T; 1.002)
Health and care	-	-	-	-	-	-	0.388 (O; 1.12)	0.585 (T; 1.07)
Other	-	-	-	-	-	-	-	0.037** (O; 1.12)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

6.1.5. Regression pension payments

The previous paragraphs discussed if there are differences between subgroups while the actual financial differences are not explored. Regressions can provide insight in the financial effects of e.g. being higher educated. Table 12 provides the regression statistics.

Table 12: Regression statistics pension payments (Least Squares Method)

Variables	Model 1	Model 2
Dependent variable	Private pension payments (€)	Total annual pension payments (€)
<i>C</i>	9,804***	10,473***
<i>DummyRegularEmployed</i>	2,844*	2,466
<i>DummyPensionFund</i>	17,246***	17,189***
<i>DummyHighLiteracy</i>	2,222***	2,014**
<i>DummyHighEducated</i>	1,124*	1,280*
<i>DummyIndustryAgriculture</i>	1,018	1,171
<i>DummyIndustryManufacturing</i>	-1,889	-78
<i>DummyIndustryConstruction</i>	5,192	4,841
<i>DummyIndustryRetailRepairHoreca</i>	5,452*	5,242*
<i>DummyIndustryService</i>	4,218	4,116
<i>DummyIndustryPublicAndEducation</i>	107	532
<i>DummyIndustryHealthAndCare</i>	7,838**	10,350***
<i>DummyIndustryOtherIndustry</i>	-398	-471
Observations included	21,481	21,896
R-squared	0.006874	0.006290
Adjusted R-squared	0.006319	0.005745
BIC	23.67834	23.76813

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

Drawing conclusions of these regressions is hard, since multicollinearity issues occur and significant added value of some coefficients is absent. However, the regression suggests higher pension payments in the case of regular employment, being employed in the health and care sector or in the retail. Furthermore (the combination of) high literacy and high education drive the payments positively.

6.1.6. Conclusion hypothesis 1

The actual individual pension payments differ between subgroups. Self-employed in the public and education sector get on average higher pension payments. Participants in the agriculture achieve on average the smallest amount of money. So some groups show significant higher levels of pension payments compared to other subgroups. Hence there is little evidence that the findings of Genre, Kohn and Momferatou (2011) and Dickens and Katz (1987) holds after retirement of the treated

people. Furthermore, the regression suggest a lower pension payment for self-employed. If a policy officer wants to ensure a certain level of total income after retirement, he can use this hypothesis to find¹² proof for the thesis that some groups of self-employed earn ‘too little’ since these groups are self-employed (lower payments than regular employed) and belong to a certain subgroup (industry comparison).

6.2.Hypothesis 2 – subgroups and the pension savings

The second hypothesis looks like the first hypothesis, the difference is the timing. The first hypothesis considered the after retirement period, the second hypothesis is about the pre-retirement period. The second hypothesis investigates the pension savings as percentage of income. As mentioned before, data of the Dutch ZZP-Panel will be used. All of the subgroups have a non-normal distribution, so non-parametric tests will be executed. The tables below provide the test statistics. If one subgroup saves significant less than another subgroup, the focus of campaigns to save more should probably be addressed to this subgroup.

6.2.1. Education and financial literacy

Table 13: Statistics on differences in pension savings as percentage of income due to level of education

<i>p-values Mann Whitney Test ([first character of group with highest mean rank of pension savings]; [ratio highest/lowest])</i>	Low educated (435 respondents)
High educated (533 respondents)	0.012** (H; 1.1)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

The Mann Whitney Test provides evidence for the assumption that higher educated self-employed save relatively more for their pension payments. The null hypothesis of equality between the education subgroups is rejected. The level of net monthly income of high educated exceeds the level of low educated significantly¹³. An approximation of the annual earnings can be computed by multiplying the net monthly income by 12. If you multiply the annual earnings by the saving percentage subsequently, you get an estimation of the annual pension savings of a respondent. By performing the Mann Whitney Test to check if there are differences in the savings between low and high educated, the null is rejected¹⁴. Hence, the high educated self-employed save indeed more for their pension payments. The absolute and relative weighting of the saving levels of the high educated

¹² Country specific regressions are not estimated in this thesis due to the small number of self-employed within a particular country.

¹³ Mann Whitney Test, p=0.000, ratio 1.09, mean rank of net monthly income high educated exceeds mean rank of low educated.

¹⁴ Mann Whitney Test, p=0.001, ratio 1.13, mean rank of approximation annually savings by high educated exceeds mean rank of low educated.

exceeds the levels of low educated. The level of education is, logically, positively related to the financial literacy of the respondent. By comparing the group of the financial knowledgeable (high financial literacy) to the unknowledgeable (low financial literacy), the total savings by the first group exceeds the savings of the latter¹⁵. So if a policy maker wants to increase the net savings by the self-employed, focusing on the unknowledgeable will presumably be more profitable.

6.2.2. Industries

The second hypothesis is also tested on the subgroups of industry. Table 14 provides the test statistics. Self-employed in the service industry and in the health and care industry save relative more than self-employed in the agriculture, construction or other industry. Loosely speaking, the industry comparison suggests that the self-employed in the service and health and care industries save most.

¹⁵ Mann Whitney Test, $p=0.181$, ratio 1.15, mean rank of savings by 'high literacy' exceeds mean rank of 'low literacy'.

Table 14: Statistics on differences in pension savings as percentage of income due to industry

<i>p-values Mann Whitney Test ([first character of sector with highest mean rank of pension savings]; [ratio highest/lowest])</i>	Manufacturing	Construction	Retail, repair and horeca	Service	Public and education	Health and care	Other	Total minus subgroup
Agriculture	0.293 (M; 1.12)	0.762 (C; 1.03)	0.241 (R; 1.13)	0.027** (S; 1.20)	0.99 (A; 1.001)	0.01*** (H; 1.28)	0.898 (O; 1.01)	0.084* (O; 1.15)
Manufacturing	-	0.456 (M; 1.07)	0.793 (R; 1.03)	0.496 (S; 1.06)	0.378 (M; 1.1)	0.186 (H; 1.14)	0.411 (M; 1.08)	0.858 (O; 1.01)
Construction	-	-	0.263 (R; 1.11)	0.029** (S; 1.14)	0.697 (C; 1.04)	0.01*** (H; 1.22)	0.875 (C; 1.01)	0.082* (O; 1.11)
Retail repair and horeca	-	-	-	0.816 (S; 1.02)	0.273 (R; 1.13)	0.359 (H; 1.09)	0.264 (R; 1.11)	0.786 (R; 1.02)
Service	-	-	-	-	0.062* (S; 1.19)	0.244 (H; 1.08)	0.046** (S; 1.15)	0.027** (S; 1.08)
Public and education	-	-	-	-	-	0.024** (H; 1.27)	0.856 (O; 1.02)	0.143 (O; 1.14)
Health and care	-	-	-	-	-	-	0.017** (H; 1.23)	0.039** (H; 1.13)
Other	-	-	-	-	-	-	-	0.113 (T; 1.12)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

6.2.3. Main wage earners

Around fifty percent of the Dutch self-employed is main wage earner in his or her household. The approximation of their monthly income exceeds the income of the non-main wage earners¹⁶, the same holds for the annually pension savings measured in euro's¹⁷. Table 15 shows the test results of the hypothesis about saving percentages. The table shows that the non-main wage earners save - in terms of percentage - more compared to main wage earners. This can be explained by the fact that the household expenses will be paid by the main wage earner. Main wage earners save on average 10.78% while non-main wage earners save on average 18.27%. So the residual income of the non-main wage earners can be used for e.g. pension savings. The dataset does not provide information about the composition of households, so if two self-employed share a household, one will be the main wage earner and the other can save for the pension payments. In this case, the test results are ambiguous.

Table 15: Statistics on differences in percentage of pension savings due to being main wage earner

<i>p-values Mann Whitney Test ([first character of group with highest mean rank of pension savings]; [ratio highest/lowest])</i>	Non-main wage earner (379 participants)
Main wage earner (593 participants)	0.000*** (N; 1.19)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

6.2.4. Age

Based on the hyperbolic discounting insights, an increase in the savings in terms of percentage of income can be expected if the retirement comes closer. Therefore two subgroups are created, one containing the self-employed below the age of fifty years and one containing the others¹⁸. Table 16 gives proof for this idea; the null of similar savings over ages will be rejected. This means that older self-employed saves more than the younger self-employed.

Table 16: Statistics on differences due to age in pension savings as percentage of income

<i>p-values Mann Whitney Test ([first character of group with highest mean rank of pension savings]; [ratio highest/lowest])</i>	<50 years old (512 participants)
50-65 years old (422 participants)	0.008*** (50-65; 1.1)

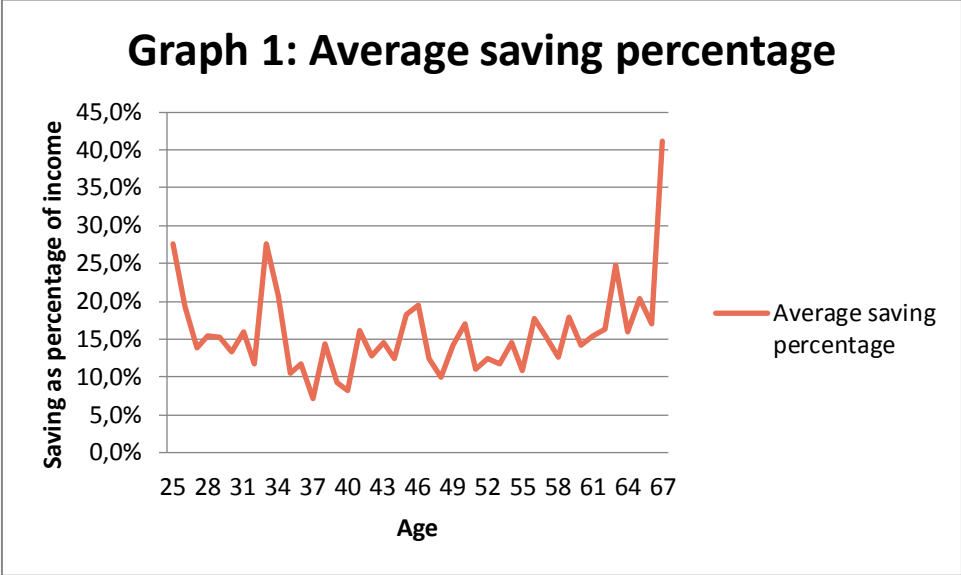
* significant at 10% level; ** significant at 5% level; *** significant at 1% level

¹⁶ Mann Whitney Test, $p=0.000$, ratio 1.76, mean rank of approximation monthly income of main wage earners exceeds mean rank of non-main wage earners.

¹⁷ Mann Whitney Test, $p=0.038$, ratio 1.08, mean rank of approximation annually savings by main wage earners exceeds mean rank of non-main wage earners.

¹⁸ The boundary at age of fifty is arbitrary, but the graph renders the absence of distortionary effects of this demarcation.

The border of age fifty can be discussed; graph 1 shows the average saving percentage based on the age of the respondent. Except some 'outliers', the red line walks around the mean of 15% and increases slightly by the age of the respondent. The graph suggests that another arranging of subgroups based on age will not result in other answers.



6.2.5. Pension funds

Just as the case of the first hypothesis, being participant of a pension fund can affect the total pension payments per year. So the null hypothesis of equal levels of pension savings between them who are enrolled and who are not enrolled in pension funds will be tested. The Mann Whitney Test does not reject the null hypothesis. The test gives proof for the assumption that pension fund participants save just as much as other self-employed.

Table 17: differences in saving levels between pension funds participants

<i>p-values Mann Whitney Test ([first character of sector with highest mean rank of pension payments]; [ratio highest/lowest])</i>	No participation in pension fund (866 respondents)
Participate in pension fund (106 respondents)	0.329 (N; 1.06)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

Since the saving levels are more or less the same while the payments differ between the subgroups of who participate in pension funds and who do not participate, it is expectable to think that participation is more efficient since the yield is higher (inter alia due to the duration as explanatory variable).

The effects mentioned above can be estimated by performing a regression. The regression shows the economic effect while the Mann Whitney Test shows whether there is an effect and it can provide

information about the medians. The factors of interest are the within effects, so a fixed model is the most apparent model. However, since the dummy for pension funds is stable over time, the problem of singularity arises. Due to this, a non-fixed model is used. The reader has to draw the conclusions carefully since the average effects are shown and generalization of causality is impossible. The regression formula is

$$(1) Y_{it} = C + \beta' \cdot dummy_x + \varepsilon_{it}$$

Where Y_{it} is the saving percentage of income, C is the constant level of savings and $dummy_x$ represents the different dummies. So you have four models:

- a. $Y_{it} = C + \beta_1 \cdot dummy_{HighEducated} + \varepsilon_{it}$
- b. $Y_{it} = C + \beta_2 \cdot dummy_{MainWageEarner} + \varepsilon_{it}$
- c. $Y_{it} = C + \beta_3 \cdot dummy_{PensionFund} + \varepsilon_{it}$
- d. $Y_{it} = C + \beta_1 \cdot dummy_{HighEducated} + \beta_2 \cdot dummy_{MainWageEarner} + \beta_3 \cdot dummy_{PensionFund} + \varepsilon_{it}$

Table 18: Regression models of formula (1) (Least Squares Method)

(2305 respondents included)	Model 3a	Model 3b	Model 3c	Model 3d
Intercept (C)	12.13429***	14.02221***	12.85487***	13.38596***
Effect of being high-educated ($\beta_1 \cdot dummy_{HighEducated}$)	1.59537**	-	-	2.648339***
Effect of being main wage earner ($\beta_2 \cdot dummy_{MainWageEarner}$)	-	-3.422368***	-	-3.945797***
Effect of participating in pension fund ($\beta_3 \cdot dummy_{PensionFund}$)	-	-	-2.465323***	-1.306701*
R-squared	0.00242	0.011636	0.001567	0.018263
Adjusted R-squared	0.00212	0.011339	0.001267	0.017378
BIC	8.317496	8.308215	8.315995	8.306356

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

The values of (adjusted) R-squared are very low. Hence, the explanatory power of these regressions are small, model 3d is of the highest explanatory power. However, these regressions indicate that

high educated self-employed save on average 2.6 percentage points more than low educated. Furthermore, the main wage earners in this model save on average approximately 4 percentage points less compared to the non-main wage earners¹⁹. Since non-main wage earners do not have to pay rents and other household payments, they can save relatively more. The participants of pension funds save less; this can be explained by the fact of the net yields. In the case of pension funds, the operating costs can be shared with other participants.

6.2.6. Conclusion hypothesis 2

Table 19: Overview hypothesis 2

Prevailing sector	Prevailed sector
High educated	Low educated
High literacy	Low literacy
Non-main wage earner	Main wage earner
≥50 years old	<50 years old
Service	Agriculture
Service	Construction
Service	Other
Service	Total minus Service
Health and care	Agriculture
Health and care	Construction
Health and care	Public and education
Health and care	Other
Health and care	Total minus Health and care

The hypothesis that the actual savings differ between subgroups of self-employed can partially be rejected. In the cases as mentioned in table 19, the left-column subgroups save significant more compared to the right-column subgroups. High educated, non-main wage earners and those who passed the age of fifty save relative more for their retirement compared to their peers. Self-employed in the service industry and in the health and care industry save relative more than self-employed in the agriculture, construction or other industry. Loosely speaking, the industry comparison suggests that the self-employed in the service and ‘health and care’ industries save most. Once again, these results are about the *savings* and do not explain the level of *pension payments*.

6.3.Hypothesis 3 – pension savings over time

The third hypothesis investigates the pattern of the savings of self-employed over time. By comparing the means, medians and deviations of pension savings as percentage of income in the ZP-Panel, some patterns can be expected. The standard deviation increased by 9.18% in the period

¹⁹ This result differs from the means in tables 15 and 17 due to time-effects. The table contains one-year data while the regressions contain data of the last four years.

2010-2013, the mean from 10.3% to 13.7% and the median increased from 7% to 10%. A weakness of the ZP-Panel is the duration. To test the homogeneity, more data than only four years is needed. In the DHS data, information about assets and income is included while data about the actual savings lack.

6.3.1. Measurement of savings over time

The pension savings are part of the total assets of a respondent. Furthermore, people can save for their retirement by other arrangements than only saving through pension fund plans. Hence, the change in assets can be seen as changes due to (inter alia) pension investment plans. The other way around, if the income adapts, the assets has to assimilate. If the income falls back, the assets can be used to supplement the income in order to hold a certain consumption level. The assets can be used as proxy for the total savings by the respondents to ensure a certain level of consumption during his total life. By testing over time if the first difference of assets significantly differs from the first difference in income, the permanent income hypothesis can be tested. Therefore a saving variable S_{it} is estimated by

$$(2) S_{it} = A - I = \frac{A_{it} - A_{it-1}}{|A_{it-1}|} - \frac{I_{it} - I_{it-1}}{|I_{it-1}|} + \varepsilon_{it}$$

where A_{it} is the total value of assets in year t , i indicate the individual and I_{it} is the annual income of individual i in year t ²⁰. If S_{it} is stable around 0, the increases and decreases in income are relative the same as the relative mutation of the assets. If S_{it} exceeds 0, the assets growth faster (or decrease less) than income does. If S_{it} is below 0, the increase in income is not reflected in the savings or the decrease in assets exceeds the decrease in income. The situation of $S_{it} < 0$ can be described as 'saving too little', or in Dutch 'ondersparen'. If the situation of $S_{it} < 0$ holds over time, the theory of income smoothing is violated.

Since the Dutch government stimulates pension savings by attractive tax arrangements, the total gross income (*btot*, per year) will be used as proxy for the income. The assets will be computed by summing the partial asset, debt and mortgage components (CentERdata, 2013). The options (call and put) are eliminated since these assets can be regarded as potential assets. If the value of options is included, the total value of the individual is deceiving since the value is twofold counted (options as well as accounts until the options are bought). The missing values in the separate data files per year are modified into zero's, so if a respondent answered 'I do not have a bank account', he has an amount of '0' on bank accounts. Nevertheless, there are still missing values if a respondent did not participate in the survey of a particular year.

²⁰ The assets as measured on January 1st.

Some respondents earned unusually high incomes in certain years (probably due to winning a lottery). Others increased or decreased their amount of assets by hundreds of percent's. Due to these outliers, the data become incredible. Therefore a range is handled; changes above +100% and below -50% are eliminated in formula (2). Table 20 shows the values of S_{it} for the self-employed and the not-self-employed.

Table 20: Values of S_{it} (1,039 self-employed, 14,578 non-self-employed)

t Year	S_{it} Self-Employed (mean)	S_{it} Non-Self-Employed (mean)	Significant difference (p -value Mann-Whitney test)
1994	-0.34	-0.46	0.048
1995	-0.23	-0.08	
1996	-0.16	-0.12	
1997	0.09	-0.09	0.031
1998	-0.09	-0.13	
1999	-0.39	-0.24	
2000	-0.03	-0.31	
2001	-0.60	-0.30	
2002	-0.13	0.00	
2003	0.11	-0.13	
2004	-0.11	-0.07	
2005	-0.13	-0.07	
2006	0.13	-0.03	
2007	0.00	-0.05	
2008	0.06	-0.05	
2009	-0.04	-0.17	0.036
2010	-0.13	-0.01	0.011
2011	0.07	-0.04	
2012	-0.1	-0.04	
2013	-0.01	-0.13	

In the column of self-employed, the number of negative values (14) compared to the number of positive values (6) is remarkable. If the amount and deviation from zero of negative values was the same as the amount and deviation of positive values, one can assume the presence of consumption smoothing behavior. Since there are twice as many negative values, the possibility of a rejection of the consumption smoothing ideas is more intuitively. The values of the non-self-employed should be carefully interpreted; e.g. the compulsory pension savings of this group do not belong to the assets of these respondents.

6.3.2. Industries

If we divide the group of self-employed in subgroups (based on the industries as used in the previous hypotheses) the dispersion between the subgroups sticks out. Table 21 shows the values of S_{it} per

subgroup. Some cells are empty due to the low number of observations. The colors are calculated per year (horizontally) and show that the extreme values can be found mainly in the case of self-employed. Remarkable is the presence of extreme values in especially the 'concrete product sectors' like agriculture, construction and repair. Using table 21 to find differences in the pension saving behavior between sub-groups is meaningless since the number of observations for a specific case can be very low. This problem can be solved by performing a regression.

Table 21: Values of S_{it} per subgroup based on industries

	Self-Employed		S per subgroup															
	Yes	No	Self-Employed									Non-Self-Employed						
			Industry: Agriculture	Industry: Manufacturing	Industry: Construction	Industry: Retail, repair and horeca	Industry: Service	Industry: Public and education	Industry: Health and care	Industry: Other industry	Industry: Agriculture	Industry: Manufacturing	Industry: Construction	Industry: Retail, repair and horeca	Industry: Service	Industry: Public and education	Industry: Health and care	Industry: Other industry
		Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	
S2013	-0,01	-0,13		,21	,81	-,17	,18	-,16	-,43	,15	-,10	-,15	-,07	-,23	-,15	-,19	-,14	-,06
S2012	-,10	-,04	-,36	,06	-,30	1,12	-,15	,02	-,22	-,05	-,17	-,07	-,07	-,05	-,01	-,01	-,05	-,03
S2011	,07	-,04	-,92	-,33	-,1,00	-,13	,11	,14	,07	,18	-,02	-,07	-,02	-,05	-,05	-,05	-,15	-,02
S2010	-,13	-,01	-,21	-,48	-,61	-,03	-,11	-,08	-,21	-,03	,00	,01	-,01	-,02	,04	,06	-,04	
S2009	-,04	-,17	,63	,08	-,1,42	-,37	,02	-,21	,13	,04	-,15	-,14	-,29	-,31	-,11	-,15	-,30	-,11
S2008	,06	-,05	-,02	-,45	-,65	,08	,00	,16	-,33	-,13	-,03	-,08	-,05	-,13	-,05	-,07	-,01	
S2007	,00	-,05		,06	-,27	,25	-,05	-,12	-,03	-,05	-,06	-,10	-,08	-,10	-,07	-,04	-,02	
S2006	,13	-,03	,72	,14	-,01	,19	,00	,07	,14	,33	,11	,03	-,11	-,09	-,06	-,04	-,02	
S2005	-,13	-,07	,24	-,17	-,1,07	-,41	-,24	-,11	-,17	,11	-,14	-,05	,00	-,10	-,07	-,18	-,04	
S2004	-,11	-,07	-,34	-,11	-,78	-,33	,02	-,37	-,14	-,14	-,02	-,07	-,01	-,16	-,01	-,05	-,17	-,06
S2003	,11	-,13		-,83	1,26	,47	-,04	-,04	-,42	-,11	-,28	-,12	-,07	-,13	-,17	-,18	-,09	-,10
S2002	-,13	,00				,12	,31	-,22	-,48		,04	,03	,06	-,05	-,13	-,05	,01	,07
S2001	-,60	-,30		-,1,05	-,1,51		-,65	-,68	-,04		,06	-,14	,33	-,36	-,13	-,26	-,34	-,42
S2000	-,03	-,31		,20					-,13	-,07		,00	-,21	-,10	-,31	-,38	-,37	-,37
S1999	-,39	-,24		-,53	-,02	-,04		-,89	-,1,06	-,30	,75	-,37	-,03	-,13	-,04	-,23	-,31	-,25
S1998	-,09	-,13		,25	,25	,14	-,1,64	,06	,04	-,12	,15	-,02	-,06	-,01	-,17	-,07	-,14	-,14
S1997	,09	-,09		,50	,26		,15	,07	,16	,08	-,79	-,14	-,06	-,14	,03	-,11	-,09	-,10
S1996	-,16	-,12		,35			,53	-,54	-,41	-,15	-,26	-,20	-,13	-,24	,06	-,17	-,16	-,13
S1995	-,23	-,08				,86	-,78	,14		-,25	-,68	-,01	,33	-,09	-,04	-,03	-,04	-,08
S1994	-,34	-,46		,14		,54			-,07	-,36	-,70	-,61	-,35	-,21	-,44	-,38	-,24	-,46

The red (lowest of S_{it}) to green (highest of S_{it}) color scale is calculated per year (horizontally).

6.3.3. Regression and subgroups based on education²¹

While these values of S_{it} provides insight in the savings of self-employed, the effect of changes in income on assets is still vague. Therefore regression (3) is performed.

$$(3) A_{it} = C + \beta_\gamma X_{it} + \varepsilon_{it} = (\beta_0 + v_i) + \beta_1(A_{it-1}) + \beta_2(\Delta I_{it}) + \varepsilon_{it} = \alpha_i + \beta_1(A_{it-1}) + \beta_2(\Delta I_{it}) + \varepsilon_{it}$$

where $\gamma = 1, \dots, N$. Since we are interested in the change of behavior of the self-employed after a change of income (within effects), a fixed model is the most apparent model (Basturk, 2013). This eliminates the various characteristics of the self-employed that do not change over time. Consequently, the Hausman test proves the use of a fixed effect model over a random effect model²². Therefore a constant α_i will be used instead of the regular $C = \beta_0$. This intercept α_i is the sum of the standard intercept β_0 and an individual fixed term v_i for all periods (so $\alpha_i = (\beta_0 + v_i)$). This variable v_i consists all fixed omitted variables for individual i . So formula (3) can be rewritten into $A_{it} = (\beta_0 + v_i) + \beta_1(A_{it-1}) + \beta_2(\Delta I_{it}) + \varepsilon_{it}$. The $\alpha_i = (\beta_0 + v_i)$ can be interpreted as average intercept since $E(v_i) = 0$ and $Var(v_i) = \sigma_v^2$. Due to the addition of the individual intercept, the abovementioned Gauss-Markov assumptions can be complemented by $E(v_i v_{i'}) = 0$ for $i \neq i'$, $E(x_{it} v_i) = 0$ for all i, t and $E(\varepsilon_{it} v_{i'}) = 0$ for all i, i', t . Overall, the econometric added value of $\bar{\alpha}_i$ in the outputs is very small due to the average effect and can be used to compare the self-employed who are not affected by a change in income.

The models 1-7 explain the total assets value, the eighth model explain the change in asset value. So the first models explain the effect of inter alia income changes on the actual value of the assets while the eighth model explain the effect of these variables on the change in asset value. Appendix 2 shows a comprehensive table of the different models, the next paragraphs will elaborate the models piece by piece.

6.3.3.1 General regression

The hypothesis assumes differences between positive and negative values of ΔI_{it} . To find coefficients which explain the relations in detail, the variables about income changes are estimated in two different ways. The first one contains only one coefficient for the difference in income. In the second the difference in income is split in a positive or negative change to check the effects of positive and negative changes separately. $\Delta INEG_{it}$ contains the negative values of ΔI_{it} and $\Delta IPOS_{it}$ contains the positive values of ΔI_{it} . Due to multicollinearity issues, the model containing ΔI_{it} , $\Delta IPOS_{it}$ and

²¹ Appendix 2 contains a comprehensive overview of the different regression models to show the differences in coefficients due to different estimation methods.

²² Hausman test statistics model 4: $X^2 = 164.346$, degrees of freedom = 2, p -value = 0.000. So the null $H_0: E(\alpha_i | x_{it}) = 0$ will be rejected. The random individual effects estimator is inconsistent, the use of a fixed effects model will be more appropriate.

$\Delta INEG_{it}$ will not be estimated. In the economic literature, there is not an explicit predilection for the use of absolute or relative values of the assets. In my opinion, values expressed in euros are more understandable than percentages, therefore I use absolute values.

Table 22: Coefficients of models for A_{it} , 1994-2013 (415 of 1,259 observations included, Least Squares Method)

Coefficient	Model 4a (€)	Model 5a (€)
Intercept ($\bar{\alpha}_t$)	133,551***	139,756***
Assets of last year (A_{it-1})	0.339***	0.339***
Change in income (ΔI_{it})	0.404	-
Negative change in income ($\Delta INEG_{it}$)	-	0.814*
Positive change in income ($\Delta IPOS_{it}$)	-	-0.143
R-squared	0.748	0.749
Adjusted R-squared	0.615	0.616
BIC	29.749	29.753

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

The data provides information about approximately twenty years. The coefficient for the negative income changes is much stronger (and of more significance) than the coefficient for the positive income changes. So if a self-employed individual is confronted with a decrease in income, he deducts his asset value by a higher proportion than he adapts his assets after an increase income. During these decades the currency changed, the internet bubble occurred and the financial crisis affected the total economic system. These economic effects on the state of the world can induce distorted interpretation. Hence, the models are also estimated for the last decade 2004-2013 to avoid a lot of disturbance due to externalities.

Table 23: Coefficients of models for A_{it} , 2004-2013 (184 of 624 observations included, Least Squares Method)

Coefficient	Model 4b (€)	Model 5b (€)
Intercept ($\bar{\alpha}_i$)	232,719***	263,761***
Assets of last year (A_{it-1})	0.164***	0.156***
Change in income (ΔI_{it})	1.225**	-
Negative change in income ($\Delta INEG_{it}$)	-	2.911***
Positive change in income ($\Delta IPOS_{it}$)	-	-2.235*
R-squared	0.842	0.846
Adjusted R-squared	0.771	0.776
BIC	29.312	29.297

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

The Bayesian Information Criteria of model 5b is the lowest. Therefore model 5b²³ provides the best insight in the effect of changes in income on the amount of assets. On average, a decrease in income leads to a 3 times larger decrease in assets. The negative sign of the increase in income can be explained by larger amount of assets in the last year (higher levels of A_{it-1}) and the individual levels of α_i . The savings as measured by the amount of assets does not show a comparable pattern over time as income does.

6.3.3.2 Education

Like the first and second hypothesis, also the third hypothesis will be tested for the subgroups based on level of education. The self-employed are divided in two subgroups, either low-educated or high-educated. To perform the regressions, a dummy is added so that $dummy_{HighEducated} = 1$ in the case of high-educated self-employed. By adding this dummy to model 4a-5b, a singularity issue arises since the education of an individual does not change over time. Therefore an individual non-fixed with time fixed effects is performed. Table 24 shows the test statistics. The asset value of high-educated self-employed exceeds those of low-educated by approximately €17.300. However, the addition of the dummy for high-educated self-employed is not

²³ Model 5b: $A_{it} = 263,760.5 + 0.155792 \cdot A_{it-1} + 2.911404 \cdot \Delta INEG_{it} - 2.234846 \cdot \Delta IPOS_{it} + \varepsilon_{it}$

significant and lowers the R-squared. Furthermore, since the models are no longer individual fixed effect models, the coefficients can be biased by unknown individual effects.

The models 6 and 7 explain the effect on the absolute value of the assets. The effect of a change in income and of education on the change in assets can be estimated by model 8²⁴. The explanatory power of model 8 falls below the explanatory power of the models 6 and 7. Furthermore, the added significant value of the different variables is low. Model 8d shows less extreme effects in the case of high educated self-employed. Likewise, low educated self-employed do not increase their assets in the case of an increased income. Higher educated increase their assets by 1,1 times the income change in times of an increase in income and lower their assets by 2.6 times the income change. So the overall effects are less negative in the case of high-educated while this less negative effect will be abated by the general dummy variable of -€7,692,-. An important remark is the low statistical significance of the partial variables and of the whole model.

Table 24: Coefficients of models for A_{it} and ΔA_{it} ; education included; 2004-2013 (184 of 624 observations included, Least Squares Method)

Coefficient	Model 6 (€)	Model 7 (€)	Model 8c (€)	Model 8d (€)
Dependent variable	Total Assets (A_{it})	Total Assets (A_{it})	Change in Assets (ΔA_{it})	Change in Assets (ΔA_{it})
$\bar{\alpha}_t$	45,636***	56,382***	3,356	28,411*
A_{it-1}	0.839***	0.848***	-	-
ΔI_{it}	1.848***	-	2.714	-
$\Delta INEG_{it}$	-	2.669***	-	4.205***
$\Delta IPOS_{it}$	-	0.312	-	-0.334
<i>dummy</i> _{HighEducated}	17,315	17,276	8,316	-7,692
$\Delta I_{it} \cdot \text{dummy}_{HighEducated}$	-	-	-0.621	-
$\Delta INEG_{it} \cdot \text{dummy}_{HighEducated}$	-	-	-	-1.588
$\Delta IPOS_{it} \cdot \text{dummy}_{HighEducated}$	-	-	-	1.416
R-squared	0.725	0.726	0.062	0.080
Adjusted R-squared	0.719	0.721	0.044	0.058
BIC	27.992	27.997	28.082	28.084

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

²⁴Model 8d: $\Delta A_{it} = \alpha_i + \beta_1 \cdot \Delta INEG_{it} + \beta_2 \cdot \Delta IPOS_{it} + \beta_3 \cdot \text{dummy}_{HighEducated} + \beta_4 \cdot \Delta INEG_{it} \cdot \text{dummy}_{HighEducated} + \beta_5 \cdot \Delta IPOS_{it} \cdot \text{dummy}_{HighEducated} + \varepsilon_{it}$

To summarize the regression findings, the models 5a and 5b suggest a larger effect of a decrease in income than of an increase in income (in model 5a to the ratio 5.7 and in model 5b to the ratio of 1.3). Model 8 underlines these findings. So these regressions provide evidence for the idea of 'saving to less': in years of income reductions the savings will shrink relatively more while in the case of growth in income the assets are unaltered. If two respondents are confronted with the same change in income, the difference in total assets will be explained by the assets of last year, by the level of education (high or low) and in the case of fixed cross section models also by the individual α_i .

6.3.4. Conclusion hypothesis 3

To conclude, the third hypothesis of the same pattern of income and assets over time will be rejected. Self-employed who are confronted with a decrease of income, deduct their asset value by 3 times the income reduction. Self-employed do not act according to the permanent income hypothesis. Consumption smoothing and income smoothing ideas will be neglected. A caveat of this conclusion is the generalization. Since fixed effect models are used, generalization to others than the respondents is impossible (Verbeek, 2012). Furthermore, the selection of one particular model is hard. The statistical significances of some of the variables are very low and economic theory does not prefer one particular model. However, by comparing the different models a general pattern can be found: high-educated act more conform the consumption smoothing and income smoothing ideas than the low-educated while both groups of self-employed diminish their asset values to a larger extent in times of a decrease in income than they increase their asset values in times of an increase in income.

Nevertheless, the multiplier of 3 in model 5b is worrisome in my opinion. The self-employed omit to compensate the decrease of assets, so the risk of a huge decrease in income after retirement increases. Since the awareness of pension issues is very low (SER, 2010), the awareness of the risks will be low. Pension consultancy firms, pension providers and governments can use this finding of 'saving too less' to develop products - especially for the low-educated self-employed - which nudge the self-employed to increase their savings when their income increased.

6.4. Hypothesis 4 – pension motives

The fourth hypothesis links characteristics of self-employed to their saving motives. Based on previous research, distinctions will be made between levels of financial literacy (myopic loss aversion et cetera), ages (hyperbolic discounting) and mutations in income (income smoothing, 'anti cycling

saving'). The saving motives are ordered in the way Mastrogiacomo and Alessie (2014) did²⁵. Before the subgroups will be discussed, some descriptive statistics are presented.

Table 25: Summary statistics of saving motives.

Means of saving motive. <i>Self-employed answered between 1 (low) to 7 (important motive)</i>	2010	2011	2012	2013
Saving motive: Retirement	4.41	5.12	4.75	4.83
Saving motive: Bequest	3.12	4.58	3.56	3.59
Saving motive: Investment	3.85	4.66	4.24	4.26
Saving motive: Precautionary	5.15	5.82	5.49	5.62

By testing on normality, only the *Saving motive: Investment* shows a normal distribution based on the Shapiro-Wilk test. Since the Mann-Whitney test can be performed in the case of normal distributions and is more accurate than the student *t*-test in the case of non-normal distributions, this non-parametric test is used. The motive to save for retirement increased significantly between 2010 and 2013²⁶. This can be a signal of an increase in the pension awareness.

6.4.1. Financial literacy

Financial literacy is measured by asking the respondents how knowledgeable they consider themselves. If we distinguish the self-employed based on their financial knowledge, the knowledgeable respondents are more motivated to save for their retirement²⁷. But this finding about financial illiteracy is insignificant at the 5% level. So this null hypothesis cannot be rejected. In some other cases, indicative but insignificant results appear. For example, if the income decreased last year, the motivation to save for retirement increases slightly²⁸. Another illustrative example is the effect of the financial situation of a household. If there are debts, the motivation to save for

²⁵ See appendix 3

²⁶ Mann Whitney Test, $p=0.001$, ratio 1.23, mean rank of 2013 exceeds mean rank of 2010.

²⁷ Mann Whitney Test, $p=0.078$, ratio 1.28, mean rank of motivation to save for retirement of financial knowledgeable self-employed exceeds mean rank of non-knowledgeable.

²⁸ Mann Whitney Test, $p=0.758$, ratio 1.06, mean rank of motivation to save for retirement of self-employed whose income decreased exceeds mean rank of self-employed whose income was unaltered or increased.

retirement is less than if there is a lot of money in the household that can be saved²⁹. A drawback of these results is the number of valid cases, while the adjusted dataset about motivations contains 415 observations; the partial variables describe around 100 self-employed. Therefore the hypothesis is also tested based on the SHARE data. 583 of the self-employed and retired respondents were asked to solve four numerical questions³⁰; the financial and numerical performance of the respondents can be estimated by ranking them on the number of corrected answers. None of the respondents answered all questions correctly; just 16.5% answered three questions correctly. Hereafter the difference in total pension income between the best and bad performing retired self-employed can be computed. Table 26 supports the insignificant finding that financial literacy is, to some degree, related to pension performance

Table 26: differences in annual pension income based on estimated financial performance

<i>p-values Mann Whitney Test ([number of correct answers with highest mean rank of pension payments]; [ratio highest/lowest])</i>	Two correct answers	Three correct answers
One correct answer (average pension income €39,775; 106 respondents)	0.323 (1.09; 2)	0.038** (1.22; 3)
Two correct answers (average pension income €32,554; 132 respondents)	-	0.184 (1.12; 3)
Three correct answers (average pension income €67,725; 96 respondents)	-	-

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

6.4.2. Age

The last years, the common awareness of pension issues increased due to more public campaigns and the changes in the regulations (e.g. accrual percentage discussions). This can be the reason why younger employees are concerned about their retirement savings. If we compare the level of motivation to save for retirement between young self-employed and older self-employed, some insignificant signs appear. The results, based on the DHS data, implicate highly motivated young self-employed while the self-employed who will retire in a few years are less motivated to save for their retirement. The lower interest by older self-employed can be caused by the maximum possible

²⁹ Mann Whitney Test, p=0.066, ratio 1.16, mean rank of motivation to save for retirement of self-employed which have a lot of money that can be saved exceeds mean rank of self-employed having debts.

³⁰ E.g. “10% of 1000 is...” and “If €6.000 is two third of the original price, what was the original price?” Since some did not answered questions, the group of ‘null correct answers’ is eliminated to overcome deceiving insights.

return: since the duration of the investment is just a few years, the investment will not be very profitable. Tables 27 till 29 provide statistics on different age-groups. These tables suggest that the youth is highly motivated to save for their retirement. If we change the subgroups a bit, this pattern will be diminished but not be broken. Since the youngest self-employed in the dataset (17 years old) do not live on their own, and do not pay the rents and other household expenses, the self-employed younger than 25 years old are eliminated of the next tests.

The second hypothesis showed a positive association between age and savings while the fourth hypothesis shows a negative association between age and motivation. This remarkable difference can be explained by, inter alia, the difference in expenses between the younger and older self-employed. Furthermore, the difference in motivation can possibly be explained by the expected yield. The relative benefits for the youth exceeds the relative benefits for the older self-employed due to the duration. Further research that discounts the difference in yields can probably shed light on the difference between savings and motivation for the young compared to the old self-employed.

Table 27: Statistics on differences in motives to save for retirement due to age

<i>p-values Mann Whitney Test ([first character of group with highest mean rank of pension saving motives]; [ratio highest/lowest])</i>	<60 years old (268 respondents)
60-65 years old (61 respondents)	0.897 (60-65; 1.01)
65+ years old (91 respondents)	0.209 (<60; 1.10)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

Table 28: Statistics on differences in motives to save for retirement due to age

<i>p-values Mann Whitney Test ([first character of group with highest mean rank of pension saving motives]; [ratio highest/lowest])</i>	<45 years old (116 respondents)
45-65 years old (213 respondents)	0.607 (<45; 1.04)
65+ years old (91 respondents)	0.125 (<45; 1.13)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

Table 29: Statistics on differences in motives to save for retirement due to age

<i>p-values Mann Whitney Test ([first character of group with highest mean rank of pension saving motives]; [ratio highest/lowest])</i>	<35 years old (50 respondents)
35-65 years old (279 respondents)	0.213 (<35; 1.11)
65+ years old (91 respondents)	0.111 (<35; 1.17)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

6.4.3. Mutations in income

In the DHS dataset, respondents were asked if there income was unusually low, regular or high compared to previous years. This question can shed light on the *perception* of mutations of income

over time. Also the comparison of the changes in real income, as mentioned before, can crystalize the motives, but the DHS dataset contains only intervals of income. So if a respondent is confronted with an increase in income within an interval, the data does not mention this increase. For example, if a respondent raised his income from 60,000 euros to 65,000 euros, he still is in the same category.

Table 30: Statistics on differences in motives to save for retirement due to income changes

<i>p-values Mann Whitney Test ([first character of group with highest mean rank of pension saving motive]; [ratio highest/lowest], 105 respondents included)</i>	Decrease (at least one year decrease in income)
Increase (at least one year increase in income)	0.290 (D; 1.18)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

Table 31: Statistics on differences in motives to save for retirement due to perception of income changes

<i>p-values Mann Whitney Test ([first character of group with highest mean rank of pension saving motive]; [ratio highest/lowest] 102 respondents included)</i>	Unusually high income
Unusually low income	0.360 (High; 1.22)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

The tables suggest that a decrease in income will result in an increase in motivation to save, but the perception of a high income will sort the same effect. It is notable that the results suggest that the effects by perception and real income are not congruent. The changes in income can be incorporated in the perception; id est if a self-employed is confronted with an income reduction, he can still think that he earns an unusually high income due to the lower income by peers. Furthermore, if the respondent is confronted with a decrease in income, he will consider his financial position critically. The experience of this income reduction is probably the reason of a higher saving motive to overcome the income reduction after retirement. The combination of table 30 and 31 is remarkable; however, deriving conclusions should be done carefully.

6.4.4. Alternative method: relative retirement saving motivation

An important weakness of the method of comparing the saving motives is the assumption that the perception of the scale between the respondents is the same. For example, individuals A and B can rank their saving motivation at level two while the individual perception differs (the motivation of individual A at level two can be the same as the motivation of individual B at level four). This caveat can be limited by introducing relative motivational levels. By checking if the retirement saving motive outperforms the overall individual saving motive (average of the four motives per individual), the individual is highly motivated to save for retirement.

$$(4) \quad R_i = \frac{Retirement_i}{(Retirement_i + Bequest_i + Precautionary_i + Investment_i)}$$

If $R_i < 1$, the motive to save for retirement falls below the average individual saving motive. If $R_i = 1$, there is no relative difference to other motives. If $R_i > 1$, the individual is motivated to save for retirement. This ratio does not say anything about differences in saving motives between subgroups, only about the individual perception. Assumed that the respondents of a certain subgroup have a certain motivation to save, the ratio R states how motivated they are to save for retirement compared to other saving motives.

By recalculation the difference in motive to save for retirement due to income change, the motives differ significantly. In the case of at least one year decrease in income $\bar{R} = 1.11$, and in the case of at least one year increase in income $\bar{R} = 1.02$. So the individual motivation to save for retirement in the case of a decrease exceeds the other saving motives by 11% on average. Also the Mann Whitney Test is performed to check the significance of the difference in ratio R_i about motives to save for retirement due to income changes. This test shows a p -value of 0.018 and shows a ratio of 1.41, the motivation to save for retirement in the case of a decrease in income exceeds the motivation in the case of an increase in income. This finding is according common sense; the respondent faced a decrease in income and can think about the expected decrease in income after retirement. Since he experienced the feeling of such a decrease, as a reaction his motivation can be increased.

By calculating the Mann Whitney Test statistics for the subgroup comparisons, in most of the cases the difference in ratio highest over lowest and p -value to the founded values in the original statistics will be strengthened. The contrary holds in the case of main wage earners. The absolute motivation of non-main wage earners exceeds the motivation of main-wage earners by a ratio of 1.05; this ratio will be lower in the case of relative retirement motivation³¹. In particular in the cases of age categories, the relative retirement saving motives differ significantly. This can mean that non-main wage earners are more motivated to save than main-wage earners but the saving goals (retirement, precautionary et cetera) differ. More intuitive results are found by comparing subgroups based on age. The comparison of the group of respondents who will retire in a few years to the younger respondents, finds proof for the hyperbolic discounting theory will be found. The relative motivation to save for retirement compared to other saving motives is higher for the respondents who will retire in a few years than for those who retire at any moment in the future. Table 32 suggests a U-shaped parabola. As mentioned before, the downward slope of motivation can be caused by the current

³¹ Mann Whitney Test absolute motives, $p=0.478$, ratio 1.05, Mann Whitney Test relative motives, $p=0.702$, ratio 1.009.

economic state of the world. Probably due to the diminishing of social securities and an increasing awareness of the necessity of pension savings, the motivation of the youth is higher.

Table 32: Statistics on differences in motives to save for retirement due to age

<i>p-values Mann Whitney Test ([first character of group with highest mean rank of pension saving motives]; [ratio highest/lowest])</i>	60-65 years old
<35 years old	0.241 (60-65; 1.15)
<45 years old	0.086* (60-65; 1.17)
<50 years old	0.063* (60-65; 1.17)
<60 years old	0.076* (60-65; 1.15)

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

This relative retirement variable depends on age, but the absolute effect of one year is still unknown. So a regression (Least Squares Method) is performed to investigate the effect of age on the relative retirement motive for self-employed who are not 65+ years old. The regression can be misrepresentative in the case of e.g. students since they can be subsidized by government or live by their parents. The costs of living of these students differ from the costs of living of ‘regular self-employed’ and hence their saving opportunities differ. So the regression is based on the self-employed between 25 and 65 years old.

$$(5) \quad R_i = \alpha + \beta \cdot age_i + \varepsilon_i = 0.960281 + 0.002113 \cdot age_i + \varepsilon_i$$

The estimation based on 299 observations found a positive relation between age and the relative retirement saving motive. The *p*-value of β is 0.022, so the marginal effect is significant. While the Mann Whitney Test found proof for the procrastination and hyperbolic discounting ideas, the regression emphasizes this finding.

The regression shows a small (but still positive) marginal effect. If the regression was performed using dummies based on age, the marginal effect became stronger. Table 33 provides the test statistics. The model shows an increase of β_x till the age of 65 (third column shows the difference to the previous age dummy), so the motivation to save for retirement increased by age. Since dummies are used, the effect of $\beta_x \cdot age_i$ is eliminated. This finding proves the hyperbolic discounting theory.

Table 33: Coefficients model R_i dummies based on age. (Ordinary Least Squared)

Coefficient	Model 5	Differences in % (regarding previous dummy)
Respondent between age of 25 and 35 (<i>dummy_{age25-35}</i>)	1.024***	
Respondent between age of 35 and 45 (<i>dummy_{age35-45}</i>)	1.037***	1.227
Respondent between age of 45 and 55 (<i>dummy_{age45-55}</i>)	1.054***	1.615
Respondent between age of 55 and 65 (<i>dummy_{age55-65}</i>)	1.075***	2.032
Respondent older than age of 65 (<i>dummy_{age65plus}</i>)	0.994***	-7.572
R-squared	0.030	
Adjusted R-squared	0.020	
BIC	-0.674	
Included observations	394	

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

6.4.5. Conclusion hypothesis 4

To summarize the findings of hypothesis 4, the conclusion will be that the hypothesis of no difference in the average pension saving motives between subgroups can be rejected. There are indications of differences between subgroups. The theories about myopic loss aversion and procrastination can be proved in the case of saving *motives* by this data. The theory of hyperbolic discounting occurs in the case of retirement saving motives compared to total saving motives. The overall high saving motive by youth can be explained by the awareness of the shrinking automatism of receiving a pension and the decrease of fiscal advantages and social securities. The performed regressions show an increasing marginal effect of age on the motivation to save for retirement.

There are effects of literacy and of income changes on the motivation. The differences in motivation between subgroups can be used as proxy to estimate the feasibility to incentivize certain self-employed to save for their retirement. Hence, an approach based on characteristics like the age of

self-employed will be more useful since the motivation depends of such characteristics³² and the significance of the effect of age is higher than the significance of literacy.

7. Conclusion and remarks

7.1. Main findings

This thesis investigated the desirability of an individual approach of self-employed instead of a collective approach for the pension plans of self-employed. While previous literature focused on the whole group of self-employed or investigated parts of their pension behavior, this thesis is –to my knowledge– the first investigation that combines the behavioral and economic insights for the retirement issues of individual self-employed.

While there are differences in pension payments between subgroups of self-employed, not all of these payments differ significantly. Self-employed in the public and education sector achieve on average higher pension payments. A caveat of this conclusion is the use of European data. Country specific factors (like differences in retirement ages and fiscal policies) and the (for some countries low) number of observations abated the findings. The lower pension payments of self-employed compared to regular employed is discussed in previous literature.

There is more variability in the savings of the self-employed and hence the rejection of the second hypothesis is stronger than the rejection of the first hypothesis. Self-employed in the service or in the 'health & care' sector save more on average. Corresponding to the common sense, the high educated and the non-main wage earners save on average more compared to their counterparts. The same holds for the self-employed who passed age of fifty compared to the self-employed younger than fifty years old.

The third and fourth hypothesis tested two of the reasons to save (more), namely the variability of income and the motivation to save. These hypotheses assumed equality between subgroups, however both hypotheses are rejected. Self-employed who are confronted with a decrease of income, deduct their savings by three times the income reduction. However, the motivation of such self-employed to save for retirement increases. These self-employed wanted to save for their pension payments but could not save due to their individual financial opportunities and constraints. Nevertheless, the missed savings will not be (fully) compensated after income increases. Simplified, self-employed cut their savings when their income falls but do not take actions to compensate the

³² If for example the thirties can be motivated to save for retirement, the savings can be altered due to a higher motivation and, consequently, the expected pension payments will increase due to duration.

deficits. This finding weakens the permanent income hypothesis. Additionally, literacy affects the savings and age affects the motivation to save. Self-employed procrastinate their savings in the case of an income decrease, furthermore, they discount (retirement) saving motivation hyperbolically. The future will demonstrate if the higher motivation of the youth to save for retirement will result in higher actual savings and pension payments.

7.2. Discussion

While this thesis does not discuss the minimum level of pension payments and does not deliberate the permission to interfere in the responsibilities of self-employed, some conclusions can be drawn. A lot of self-employed wants to save for retirement without taking actions. Income variability affects these motivation and actions. Some characteristics, like level of education and financial literacy of these self-employed can explain the differences in pension behavior.

The current system of pension providers and pension products is mainly based on the assumption of being regular employed and the arrangements are mainly based on industries (sectorial pension funds). To answer the research question, this thesis suggests that a more individual approach based on level of education, being main wage earner and also the income variability should be incorporated. An industry-approach can be efficient if there is a lot of homogeneity within an industry. However, in the normal case of heterogeneous self-employed, guidance is needed to bridge their knowledge gap and to push the self-employed into an efficient and reliable pension plan.

Pension providers can adapt to these findings by providing a product based on the individual preferences of self-employed, on their income variability and on their individual characteristics. To understand the risks and opportunities of different pension plans, some guidance should be provided. So if the pension providers can provide³³ a product like an individual defined contribution plan stored at a PPI and they present the risks and opportunities in a comprehensible way, self-employed can participate in the pension system. They can create an individual pension plan that takes income variability and individual characteristics into account. The surplus of the PPI compared to a regular DC-plan is the comprehensibility of PPI's. This comprehensibility gives the opportunity for self-employed to modify their pension plan based on individual characteristics and anticipate the income changes. Furthermore, PPI's provide information in an understandable way, so the financial knowledge gap can be bridged. If the providers commit to guide the self-employed, the PPI can be a solution to overcome the saving problems of self-employed

³³ In the Netherlands, participating by self-employed in PPI's is not allowed by law while European regulation allows PPI's in the case of self-employed. A small amendment of the Dutch law can solve this problem (Lutjens, 2011).

In short, a ‘within self-employed’ approach based on age, income variability, being main wage earner, industry characteristics, education and financial literacy instead of a ‘between self-employed’ approach based on industry and age is more appropriate to overcome the current (negative) pension issues of self-employed.

7.3.Limitations and further research

This paragraph discusses some caveats of this thesis. First of all, different datasets are used. Due to the freshness of research on self-employed, and in particular of the investigation of their retirement behavior, the number of observations is acceptable but very low and useful longitudinal data is lacking. By separating self-employed from each other, subgroups with a few self-employed hold, so the number of observations can be too low. Furthermore, deriving conclusions out of different hypothesis and datasets should be done very prudential. Additionally, the valuation of conclusions depends on the opinion about minimum levels of earnings after retirement and about the permissibility of intervening the responsibilities of self-employed.

Also variables outside the data can affect the pension behavior of self-employed. For example, the financial crisis of this decade will affect the saving behavior and perception of financial products (like the investigation by Hoffmann, Post and Pennings (2013)). Since governments economize and cut the social security expenses, inhabitants are confronted with fewer social services. The combination of this economizing and the moderation of (fiscal) pension saving facilities, have led into an increase of the retirement awareness and affects the pension behavior. This (common sense) effect is not incorporated in this research. If more recently data will be collected, an investigation of the effect of the current financial crisis on the pension behavior will be interesting.

The investigation was about self-employed, the effects of their behavior on the widows and orphans is neglected while in the case of regular employed widowers’ pensions and orphans’ pensions are regulated. This limitation strengthens the necessity of taking actions to get acceptable pension payments for self-employed. Further research can incorporate the family of the self-employed.

This thesis did not investigate the best system of pension savings and payments for self-employed but investigated the individual approach. Since using an individual approach is appropriate, further research can investigate the optimization of the composition of nudges based on the individual characteristics. The way of supplying information about such individual pension plans can affect the demand. Contextual frameworks, pricing and marketing affect the demands of people. Except the investigation of the best (individual) approach, further research can study the ideal marketing to get optimal pension behavior of self-employed.

To conclude, the 'old general industry based approach' should be replaced by an individual approach. To distinguish self-employed from each other, the use of variables like level of education, age, income variability and being main wage earner can be helpful. Further research can investigate the composition of an optimal financial product. Releasing PPI opportunities for self-employed will solve the inefficient behavior of self-employed.

8. Bibliography

- "MinSZW". (2013). *Pensioen van zelfstandigen*. The Hague: Ministry of Social Affairs and Employment.
- "Robeco". (2014, October 14). *PPI Explained*. Opgeroepen op October 26, 2014, van Robeco: <http://www.robeco.com/en/professionals/strategies-products/ppi-robeco-smart-pension/ppi-explained.jsp>
- Basturk, N. (2013). *Applied Econometrics (Lecture Notes)*. Rotterdam: Erasmus School of Economics.
- Benartzi, S., & Thaler, R. H. (1999). Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments. *Management Science*, 45(3), 364-381.
- Berk, J., & DeMarzo, P. (2011). *Corporate Finance* (Global edition of Second edition ed.). New Jersey: Pearson Education Limited.
- Bernheim, B. D. (1998). Financial Illiteracy, Education, and Retirement Saving. In O. Mitchell, & S. Schieber, *Living with Defined Contribution Plans* (pp. 38-68). Pennsylvania: Wharton School, Pension Research Council.
- Beugelsdijk, S., & Noorderhaven, N. (2005). Personality Characteristics of Self-Employed; An Empirical Study. *Small Business Economics*, 159-167.
- Borjas, G. J. (2013). *Labor Economics* (Sixth ed.). New York: McGraw-Hill Education.
- Borsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., . . . Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe. *International Journal of Epidemiology*, 42(4), 992-1001.
- Börsch-Supan, A., Brandt, M., Litwin, H., & Weber, G. (2013). *Active ageing and solidarity between generations in Europe*. Berlin: De Gruyter.
- Bosch, N., de Graaf-Zijl, M., & van Vuuren, D. (2014). *Discussiebijdrage t.b.v. 'IBO Zelfstandigen zonder personeel'*. The Hague: CPB (Netherlands Bureau for Economic Policy Analysis).
- Brown, S., Dietrich, M., Ortiz-Nuñez, A., & Taylor, K. (2011). Self-employment and attitudes towards risk: timing and unobserved heterogeneity. *Journal of Economic Psychology*, 425-433.
- Campbell, J. Y., Cocco, J. F., Gomes, F. J., & Maenhout, P. J. (2001). Investing Retirement Wealth: A Life-Cycle Model. In J. Y. Campbell, & M. Feldstein, *Risk Aspects of Investment-Based Social Security Reform* (pp. 439-482). Chicago: University of Chicago Press.
- CentERdata. (2013). *Codebook DNB Household Survey*. Tilburg: CentERdata.
- Choi, J. (2009). *Pension Schemes for the Self-Employed in OECD Countries*. Paris: OECD.
- Cole, S., Paulson, A., & Shastry, G. K. (2012). Smart Money: The Effect of Education on Financial Behavior. *Working paper*.

- Council European Union. (1996). *Council Regulation (EC) No 2223/96 of 25 June 1996 on the European system of national and regional accounts in the Community*. Brussels: Council European Union.
- De Jong, F. (2009). *Naar een flexibele pensioenregeling voor ZZP'ers*. Tilburg: Netspar.
- Dickens, W., & Katz, L. (1987). Inter-industry wage differences and theories of wage determination. *NBER, Working Paper no. 2271*.
- Eichhorst, W., Gerard, M., Kendzia, M., Mayrhuber, C., Nielsen, C., Rünstler, G., & Url, T. (2011). *Pension systems in the EU – contingent liabilities and assets in the public and private sector*. Brussels: European Union (Economic and Scientific Policy/ECON).
- European Commission. (2011, 01 04). *The Netherlands and the euro*. Opgeroepen op 10 02, 2014, van European Commission:
http://ec.europa.eu/economy_finance/euro/countries/netherlands_en.htm
- Friedman, M. (1957). The Permanent Income Hypothesis. In M. Friedman, *A Theory of the Consumption Function* (pp. 20-37). Princeton: Princeton University Press.
- Geijtenbeek, L., de Rijk, A., & Rienstra, M. (2012). *Inkomen en vermogen op de oude dag*. The Hague: Centraal Bureau voor de Statistiek.
- Genre, V., Kohn, K., & Momferatou, D. (2011). Understanding Inter-Industry Wage Structures in the Euro Area. *Applied Economics*, 43(11), 1299-1313.
- Hilgert, M., Hogarth, J., & Beverly, S. (2003). Household Financial management: The Connection between knowledge and Behavior. *Federal Reserve Bulletin*, 309-322.
- Hipple, S. (2010). Self-employment in the United States. *Monthly Labor Review*, 133(9), 17-32.
- Hirshleifer, J., Glazer, A., & Hirshleifer, D. (2005). *Price Theory and Applications: Decisions, Markets, and Information* (7 edition ed.). Cambridge: Cambridge University Press.
- Hodowanic, K., & Rubin, A. (2013, September 9). *Defined Contribution Plans*. Opgeroepen op October 26, 2014, van Towers Watson:
<http://vantage.internal.towerswatson.com/communities/services/ret-us/dc/Pages/default.aspx>
- Hoekstra, K., & van Vuuren, D. (2013). De fiscale behandeling en sociale zekerheid van zelfstandigen. *TPEdigitaal*, 7(4), 41-59.
- Hoffmann, A. O., Post, T., & Pennings, J. M. (2013). Individual investor perceptions and behavior during the financial crisis. *Journal of Banking & Finance*, 60-74.
- Jappelli, T., & Padula, M. (2011). Investment in Financial knowledge and Saving Decisions. *CSEF Working Paper University of Salerno*.
- Josten, E., Vlasblom, J. D., & Vrooman, C. (2014). *Bevrijd of beklemd?* The Hague: The Netherlands Institute for Social Research (SCP).

- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-292.
- Kalleberg, A. L. (2000). Nonstandard Employment Relations: Part-Time, Temporary and Contract Work. *Annual Review of Sociology*, 26, 341-365.
- Keynes, J. (1965 (1936)). *The General Theory of Employment, Interest and Money* (Paperback ed.). Orlando: Houghton Mifflin Harcourt.
- Kok, L., Baarsma, B., & Heyma, A. (2014). *Nieuw ontwerp sociaal beleid. Beoordeling van AWWN voorstel voor een basisregeling voor werkenden*. Amsterdam: SEO Economisch Onderzoek.
- Korniotis, G., & Kumar, A. (2011). Do Older Investors Make Better Investment Decisions? *The Review of Economics and Statistics*, 93(1), 244-265.
- Lahey, K. E., & Kim, D. (2001). Longitudinal changes in net worth by household income and demographic characteristics for the first three waves of the HRS. *Financial Services Review*, 55-73.
- Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, 112(2), 443-477.
- Lusardi, A. (2009). Household Saving Behavior: The Role of Financial Literacy, Information, and Financial Education Programs. In F. C., L. Goette, & S. Meier, *Policymaking Insights from Behavioral Economics* (pp. 109-149). Boston: Federal Reserve Bank of Boston.
- Lusardi, A., & Mitchell, O. (2013). The Economic Importance of Financial Literacy: Theory and Evidence. *NBER Working Paper*.
- Lusardi, A., & Mitchell, O. S. (2007). Financial Literacy and Retirement Preparedness: Evidence and implications for Financial Education. *Business Economics*, 35-44.
- Lusardi, A., Michaud, P., & Mitchell, O. (2011). Optimal Financial Literacy and Saving for Retirement. *RAND Corporation. Working Paper*.
- Lutjens, E. (2011). *De PPI ook voor zelfstandigen?* Amsterdam: VU University.
- Malter, F., & Börsch-Supan, A. (2013). *SHARE Wave 4: Innovations & Methodology*. Munich: Munich Center for the Economics of Aging.
- Mastrogiacomo, M., & Alessie, R. (2014). Where are the retirement savings of self-employed? An analysis of unconventional retirement accounts. *Draft paper*.
- Meager, N., & Bates, P. (2001). The Self-Employed and Lifetime Incomes: Some UK Evidence. *International Journal of Sociology*, 31(1), 27-58.
- Ministry of Finance. (2014). *Taakopdracht IBO Zelfstandigen zonder personeel*. The Hague: Ministry of Finance.

- O'Donoghue, T., & Rabin, M. (1999). Procrastination in preparing for retirement. In H. Aaron, *Behavioral Dimensions of Retirement Economics* (pp. 125-156). Washington: Brookings Institute and Russell Sage Foundation.
- Panteia/EIM. (2013). *ZZP-Panel 2013*. Zoetermeer: Panteia/EIM.
- Parker, S. C. (2004). *The Economics of Self-Employment and Entrepreneurship*. Cambridge: Cambridge University Press.
- Parker, S. C., & Rougier, J. C. (2007). The Retirement Behaviour of the self-employed in Britain. *Applied Economics*, 39(6), 697-713.
- Pensioenfederatie. (2010). *Het Nederlandse pensioensysteem. Een overzicht van de belangrijkste aspecten*. The Hague: Pensioenfederatie.
- Ross, D. (2010). Economic models of procrastination. In C. Andreou, & M. D. White, *The Thief of Time* (pp. 28-50). Oxford: Oxford University Press.
- SER. (2010). *Zzp'ers in beeld. Een integrale visie op zelfstandigen zonder personeel*. The Hague: The Social and Economic Council of the Netherlands (SER).
- Slichter, S. (1950). Notes on the Structure of Wages. *Review of Economics and Statistics*, 32(1), 80-91.
- Steenbeek, O. W., & Van der Lecq, S. (2007). *Costs and Benefits of Collective Pension Systems*. Berlin: Springer.
- Stockhammer, E. (2012). Euro-Keynesianism? The financial crisis in Europe. *Radical Philosophy*, 175, 2-10.
- U.S. Department of Commerce. (2011). *Statistical Abstract of the United States*. Washington: Government Printing Office.
- Unie van Beroepspensioenfondsen. (2012, January 2). *Pensioenfederatie*. Opgeroepen op September 5, 2014, van http://www.pensioenfederatie.nl/Document/UvB/UvB_Leden.pdf
- Van der Lecq, F., & Oerlemans, A. (2009). *Zelfstandigen zonder pensioen*. Tilburg: Netspar.
- Van der Westen, M. (2014, November 01). Nederland kan leren van andere pensioenlanden. *Het Financieele Dagblad*, p. 18.
- Van Els, P., Van Rooij, M., & Schuit, M. (2007). Why mandatory retirement saving. In O. W. Steenbeek, & S. Van der Lecq, *Costs and Benefits of Collective Pension Systems* (pp. 159-186). Berlin: Springer.
- Van Es, F., & Van Vuuren, D. (2010). *A decomposition of the growth in self-employment*. The Hague: CPB (Netherlands Bureau for Economic Policy Analysis).
- Van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial Literacy and Stock Market Participation. *Journal of Financial Economics*, 449-472.

- Van Schie, R., Donkers, B., & Dellaert, B. (2012). Saving adequacy uncertainty: Driver or obstacle to increased pension contributions? *Journal of Economic Psychology*, 882-896.
- Verbeek, M. (2012). *A Guide to Modern Econometrics* (Fourth Edition ed.). West Sussex: John Wiley & Sons Ltd.
- Von Neumann, J., & Morgenstern, O. (2007 (1953)). *Theory of Games and Economic Behavior* (60th Anniversary Commemorative Edition). Princeton: Princeton University Press.
- Vroom, V. H. (1964). *Work and motivation*. New York: John Wiley & Sons.
- Vroonhof, P., & de Vries, N. (2011). *Ondernemen voor de toekomst*. Zoetermeer: EIM.
- Wilkinson, N. (2008). *An Introduction to Behavioral Economics*. New York: Palgrave Macmillan.
- Wrasai, P., & Swank, O. (2007). Policy makers, advisers, and reputation. *Journal of Economic Behavior & Organization*, 579-590.
- Zissimopoulos, J., Maestas, N., & Karoly, L. A. (2007). The Effect of Retirement Incentives on Retirement Behavior. *Draft paper*.

9. Appendices

9.1. Appendix 1 - Subgroup composition

Table 34 shows how the subgroups are created and which variables are used per dataset

Subgroups	SHARE	DHS	ZZP-panel
Low educated	Based on variable 'years of education', low if ≤ 12 years of education.	Based on variable 'Highest level of education completed'. Low if level 1-5 or 8-9. For the years 1993-2006 low if not high level.	Based on 'opleidingsniveau'. Low if 2-6.
High educated	Based on variable 'years of education', high if ≥ 13 years of education.	Based on variable 'Highest level of education completed'. High if level 6 or 7. For the years 1993-2001 if 7-9, for the years 2002-2006 if 7.	Based on 'opleidingsniveau'. High if 1 (university).
Financial illiteracy	Based on performance mathematical financial performance	Based on 'kunde', illiteracy if 1 or 2.	
Financial literacy	Based on performance mathematical financial performance	Based on 'kunde', literacy if 3 or 4.	
Main wage earner	-	'Main wage earner'	Based on 'percentage GEZINSinkomen afkomstig van bedrijf' (aand_hhink), if >50 main wage earner.
Industry: Agriculture	Based on 'which industry active', sum of 'agriculture, hunting, forestry, fishing' and 'mining and quarrying'	Based on 'bedrtak'. Agriculture if 1	Based on 'Tot welke sector behoort uw bedrijf?'. Agriculture if 1 (landbouw).
Industry: Manufacturing	Based on 'which industry active'. No modification.	Based on 'bedrtak'. Manufacturing if 2	Based on 'Tot welke sector behoort uw bedrijf?'. Manufacturing if 2 (industrie).
Industry: Construction	Based on 'which industry active', sum of 'electricity, gas and water supply' and 'construction'	Based on 'bedrtak'. Construction if 3 or 4	Based on 'Tot welke sector behoort uw bedrijf?'. Construction if 3 (bouw).
Industry: Retail, repair and horeca	Based on 'which industry active', sum of	Based on 'bedrtak'. Retail, repair and	Based on 'Tot welke sector behoort uw

	'retail and repair' and 'horeca'	horeca if 5 or 6	bedrijf?'. Retail, repair and horeca if 4 (Reparatie, handel, horeca).
Industry: Service	Based on 'which industry active', sum of 'transport, storage and communication', 'financial intermediation' and 'real estate, renting and business activities'.	Based on 'bedrtak'. Service if 7-9	Sum of 'Tot welke sector behoort uw bedrijf?' value 5 (transport) and 'Dienstverlening' values 1-7 and 10.
Industry: Public and education	Based on 'which industry active', sum of 'education' and 'public administration and defence; compulsory social security'	Based on 'bedrtak'. Public and education if 10, 11 or 13.	Based on 'Dienstverlening van' level 9.
Industry: Health and care	Based on 'which industry active'. No modification	Based on 'bedrtak'. Health and care if 12	Based on 'Dienstverlening van' level 8.
Industry: Other industry	Based on 'which industry active', sum of 'don't know' and 'other'	Based on 'bedrtak'. Other if -14, -13, 14 or 15	Based on 'Tot welke sector behoort uw bedrijf?'. Other if 99999996 (Anders) or 99999997 (wil niet zeggen).

9.2. Appendix 2 – different regression models paragraph 6.3.3

Table 35 shows the different regression models of paragraph 6.3.3.

	Model 4a (€)	Model 4b (€)	Model 5a (€)	Model 5b (€)	Model 6 (€)	Model 7 (€)	Model 8a (€)	Model 8b (€)	Model 8c (€)	Model 8d (€)
Dependent variable	Total Assets (A_{it})	Total Assets (A_{it})	Total Assets (A_{it})	Total Assets (A_{it})	Total Assets (A_{it})	Total Assets (A_{it})	Change in Assets (ΔA_{it})	Change in Assets (ΔA_{it})	Change in Assets (ΔA_{it})	Change in Assets (ΔA_{it})
Intercept	133,551***	232,719***	139,756***	263,761***	45,636***	56,382***	4,667	25,087	3,356	28,411*
Assets of last year (A_{it-1})	0.339***	0.165***	0.339***	0.156***	0.839***	0.848***	-	-	-	-
Change in income (ΔI_{it})	0.404	1.225**	-	-	1.848***	-	2.513***	-	2.714***	-
Negative change in income ($\Delta INEG_{it}$)	-	-	0.814*	2.911***	-	2.669***	-	3.725***	-	4.205***
Positive change in income (ΔIPO_{it})	-	-	-0.143	-2.235*	-	0.312	-	0.056	-	-0.334
Dummy about education ($dummy_{HighEducated}$)	-	-	-	-	17,315	17,276	-	-	8,316	-7,692
$\Delta I_{it} \cdot dummy_{HighEducated}$	-	-	-	-	-	-	-	-	-0.621	-
$\Delta INEG_{it} \cdot dummy_{HighEducated}$	-	-	-	-	-	-	-	-	-	-1.588
$\Delta IPO_{it} \cdot dummy_{HighEducated}$	-	-	-	-	-	-	-	-	-	1.416
Periods included	20	10	20	10	10	10	10	10	10	10
Observations included	415	184	415	184	184	184	184	184	184	184
Individual fixed effect	Yes	Yes	Yes	Yes	No	No	Yes	Yes	No	No
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.748	0.842	0.749	0.846	0.725	0.726	0.178	0.184	0.062	0.080
Adjusted R-squared	0.615	0.771	0.616	0.776	0.719	0.721	-0.192	-0.185	0.044	0.058
BIC	29.749	29.312	29.753	29.297	27.992	27.997	29.817	29.820	28.082	28.084

* significant at 10% level; ** significant at 5% level; *** significant at 1% level

9.3. Appendix 3 – Classification of saving motives

The DHS data contains information about the saving motives. Twelve different motivations can be ranked on a range of 1 (low importance) to 7 (high importance). Mastrogiacomo and Alessie (2014) used these motivations and ordered them in four categories. Table 3.1 in the paper by Mastrogiacomo and Alessie (2014) provides the orders (factors) of survey questions about the motives:

Table 3.1 Classification of survey questions into saving factors (Mastrogiacomo & Alessie, 2014).

Factor	Survey question (motive): Is it to you personally of much or of little importance?
Precautionary	to have some savings to cover unforeseen expenses as a consequence of illness or accidents
Precautionary	to have some savings in case I or a member of my family get(s) unemployed
Precautionary	as a reserve to cover unforeseen expenses
Bequest	to leave money to my children (or other relatives)
Bequest	to give presents or other gifts to my (grand)children
Bequest	to pay for my children's (or other relatives') education
Investment	to buy durable goods such as furniture, electric appliances, or bicycles in the future
Investment	to generate income from interests or dividends
Investment	to set up my own business
Pension	to supplement my retirement pension, to have some extra money to spend when I am retired
Pension	to buy a house in the future
Pension	to supplement my social security benefit

10. Data disclaimers

This thesis uses data from SHARE wave 4 release 1.1.1, as of March 28th 2013 (DOI: 10.6103/SHARE.w4.111) or SHARE wave 1 and 2 release 2.6.0, as of November 29 2013 (DOI: 10.6103/SHARE.w1.260 and 10.6103/SHARE.w2.260) or SHARELIFE release 1, as of November 24th 2010 (DOI: 10.6103/SHARE.w3.100). The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5-CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

This thesis uses also data of the DNB Household Survey provided by CentERdata, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

This thesis uses data of the ZYP-panels 2010-2013, provided by Panteia/EIM Business & Policy Research, Panteia, P.O. Box 7001, 2701 AA Zoetermeer, The Netherlands.