Erasmus University Rotterdam

Bachelor Thesis International Econometrics and Operations Research Proposal

Major: Financial Econometrics

Risk tolerance: the reason for gender inequality?

An analysis of risk tolerance, its differences in risk tolerance between individuals, and its impact on wealth accumulation

---

Pim van der Voet (431184)

Under supervision of Prof. dr. R.L. Lumsdaine

July 8, 2018

Abstract

This paper examines risk tolerance in the context of the allocation of Individual Retirement Account portfolios. The aim of this paper is to get a good understanding of risk tolerance for men and women, and to investigate how this helps explain the difference in wealth between the genders. To do this, the methods of Neelakantan (2009) and a maximum likelihood approach are applied to data of 2006 and 2014 from the Health and Retirement Study. This paper also investigates whether gender affects risk tolerance when controlling for other variables, and whether self-reported risk tolerance can be used to help explain risk tolerance for individuals. This paper finds inconsistent results between samples on differences in risk tolerance and wealth in a cross-sectional setting for 2006. In addition, the part of the wealth gap that can be explained by differences in risk tolerance between men and women has increased between 2006 and 2014 when gender is the only characteristic considered. Furthermore, when controlling for other variables, gender seems to have no significant effect on risk tolerance but ethnicity does. Lastly, this paper finds no significant evidence that self-reported risk tolerance can be used as a proxy for risk tolerance.

1 Introduction

In 2015, all countries in the United Nations adopted a set of Sustainable Development Goals (SDGs).[1] One of the SDGs focuses on gender equality. To achieve gender equality across the globe, it is of great importance to have a good understanding of whether differences currently exist between men and women and if so, to understand what causes these differences. One of the areas that is often studied is the equality in wealth accumulation. Several factors play a role when accumulating wealth, which can be different between the genders. One of those factors is risk tolerance. Overall, individuals with a higher tolerance of (financial) risk have a higher expected level of wealth (Finke and Huston, 2003). As a starting point, this paper uses Neelakantan (2009), in which the difference in risk tolerance between men and women is described and the effect on wealth accumulation is analyzed. As a first step this paper tries to reproduce the methods of Neelakantan (2009). This requires certain assumptions to be made, because at some points it is difficult to tell what was done exactly by Neelakantan (2009). This paper extends the research of Neelakantan (2009) by proposing a different approach for the estimation of risk tolerance distributions, by explaining risk tolerance over more characteristics than gender and by considering whether self-reported risk tolerance is useful in explaining risk tolerance. The theoretical framework in which this paper operates is similar to that of Neelakantan (2009) and is briefly presented in section 3. For the analyses in this paper, data on the Individual Retirement Accounts (IRAs) of individuals that took part in the Health and Retirement Study (HRS) in 2006 and individuals that took part in the HRS in 2014 is used.[2] Relevant descriptive statistics on the data are presented in section 4. The analyses of this paper are presented in sections 5 to 8. Concluding remarks and suggestions for future research are presented in section 9.

This paper finds that the results obtained when using the methods of Neelakantan (2009) or the maximum likelihood approach are very sensitive to the data that is used. In a cross-sectional setting for 2006 it is not possible to draw strong conclusions on differences in risk tolerance between men and women, and neither for its effect on the wealth accumulation. This paper does find indications that the part of the wealth difference that can be explained by differences in risk tolerance between men and women has increased between 2006 and 2014. This paper also finds that gender does not have a significant effect when controlling for other characteristics, but race does. Lastly, this paper finds that self-reported risk tolerance does not serve well as a proxy for risk tolerance.

2 Literature Review

It is widely agreed upon that individuals with a higher level of risk tolerance can expect higher levels of wealth. The reason for this is that they accept more financial gambles with positive expected outcomes (Rabin (2000), Pratt (1964)). This reasoning was confirmed by an empirical study conducted by Finke and Huston (2003), who found that risk tolerance is one of the strongest predictors for net worth for individuals of over 65. The idea of risk tolerance increasing expected wealth accumulation is important to take into account when comparing men and women to make inferences about gen-

[1] See: A/RES/71/313

[2] An IRA is a typical savings account in the United States that offers many tax advantages and is used to save for retirement.

der equality and equity on the matter of wealth.

The importance of considering risk tolerance becomes even more apparent when taking a close look at the differences in risk tolerance between men and women. Studies have found significant evidence that there are differences between men and women in risk tolerance (Grable (1997), Grable and Joo (2000)). Jianakoplos and Bernasek (1998) found that in the United States men tend to be more risk tolerant than women, and ties this to wealth accumulation. Yao and Hanna (2005) found that men are more risk-tolerant than women, and that marital status has a significant effect on risk tolerance. There seems to be no consensus on the effect of marital status on risk tolerance in the literature. Yao and Hanna (2005) found that single individuals tend to be more risk tolerant than their married counter parts, while Grable and Joo (2000) found that married individuals are generally more risk tolerant than single individuals. Sung and Hanna (1996) found that single women tend to be less risk-tolerant than married women.

The effect of age on risk tolerance is also not clearly consistently described in literature. Yao and Hanna (2005) and Hallahan et al. (2004) found a negative relationship between age and risk tolerance, while Wang and Hanna (1997) and Grable and Joo (2000) found that risk tolerance increases with age when controlling for other variables.

The effect of education is more clear according to literature. Higher levels of education are related to higher levels of risk tolerance (Sung and Hanna (1996), Grable and Joo (2000)).

Risk tolerance levels also seem to be different between race groups. Yao et al. (2005) found that black and Hispanic individuals are less likely to take small financial risks than whites, while they are more likely to take on substantial financial risks. Sung and Hanna (1996) found indications that whites are more risk tolerant than individuals from other ethnic groups.

Instead of considering risk tolerance as implied by the actions of individuals (in this paper, the allocation of their IRA portfolios), self-assessed risk tolerance has been considered as well in the literature, as it has been shown to be a good proxy for risk tolerance (Hallahan et al., 2004).

3 Theoretical Framework

To make inferences about risk tolerance, this paper uses data regarding the IRA accounts of individuals. This section sets a framework such that the data can be used for the analyses of this paper, following the notation of Neelakantan (2009).

It is assumed that risk tolerance plays a crucial role in the allocation choice of individuals for IRAs. All individuals face an equal allocation problem in which they aim to maximize their utility by choosing to invest part of their wealth in risky assets and part in bonds (Jagannathan et al., 1996). This utility function is assumed to be a Constant Relative Risk Aversion (CRRA) function. This means that the relative risk-aversion (and risk-tolerance) with respect to the level of wealth remains constant (Merton, 1969). The utility maximization problem is defined as:

$$\max_{\{b_t^i, s_t^i\}_{t=0}^{T-1}} \frac{w_T^{1-\gamma^i} - 1}{1 - \gamma^i} \tag{1}$$

subject to the following set of restrictions:

$$b_0 + s_0 \leq w_0$$

$$b_t + s_t \leq w_t = (1 + \tilde{r_t}^s(\theta_t))\ s^i_{t-1} + (1 + r^b)\ b^i_{t-1} \quad \forall \quad 1 \leq t \leq T-1$$

$$w_T = (1 + \tilde{r_t}^s(\theta_t) s^i_{T-1} + (1 + r^b) b^i_{T-1}$$

Here, b^i_t and s^i_t denote the amount invested in non-risky assets and risky assets respectively for individual i at time t. w_t is the wealth of the individual at time t, and γ^i is the level of risk-aversion of the individual i (the reciprocal of risk tolerance). t ranges from 1 to T and denotes the periods of one year. r^b denotes the (constant) return on risk-free assets. The return on the risky assets, $\tilde{r}^s_t$, is a stochastic return depending on the state of nature which takes on (with equal probability for each t) three states and is denoted by θ_t. [3]

When the share of wealth that is invested in the risky assets is written as $\rho^i_t = s^i_t / w^i_t$, Equation 1 can be rewritten following the same steps considered in Jagannathan et al. (1996). The obtained equation maximizes the utility and is independent on the level of wealth:

$$\mathbb{E}\{[(1+\tilde{r_t}^s)\rho^i + (1+r^b)(1-\rho^i)]^{-\gamma^i}(\tilde{r_t}^s - r^b)\} = 0 \tag{2}$$

Individuals that seek to maximize their utility will always choose to invest the same share of their wealth in risky assets, as ρ^i does not depend on wealth. Hence, the optimal share of risky assets in the IRA $\rho^{i\star}$ can be obtained by solving Equation 2 for each individual.

4 Data

This paper makes use of data from the Health and Retirement study of the waves of 2006 (wave 8) and 2014 (wave 12).[4] Specifically this paper makes use of the longitudinal RAND dataset (v.2) and the RAND Fat files of the waves of 2006 (v3.A) and 2014 (v2.A).[5] This section gives insights in the construction and the composition of the samples in this paper. This section starts off by considering the samples constructed using the data from 2006. The second part of this section describes the data used from 2014. All statistics presented, are weighted statistics using the individual probability weighting scheme of the HRS unless noted otherwise.

4.1 Wave 8 - 2006

4.1.1 Sample Construction

The first part of this paper aims to replicate the methods of Neelakantan (2009). She describes the construction of the sample in two steps. Starting by considering the full sample of 18.469 individuals, she removes all individuals without an IRA, and is left with 5.265 individuals. Next, she removes all individuals of which the amount in the IRA or the percentage in risky assets is unknown, and is left with 3.156 individuals. Even though this construction seems rather straight-forward, I was not able to replicate these steps and be left with the same number of individuals. There are multiple explanations for this, including that the HRS dataset might have been updated between 2009

[3] This paper uses the same, discrete, distribution for the return on the risky assets as Neelakantan (2009). Return $\tilde{r}^s_t$ takes on values 27.3%, 13% or -15.25% for the different states of nature θ_t, with equal probabilities of $p = \frac{1}{3}$. The return on non-risky assets is constant: $r^b = 1\%$

[4] As a condition of use I note that the HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

[5] The RAND HRS Fat Files take almost all the raw variables from the HRS survey, and collapse them into a single respondent-level dataset for each wave. These files were developed at RAND with funding from the National Institute on Aging.

and 2018, and that the used variables might differ from this paper and Neelakantan (2009). This paper considers multiple samples that are constructed in different ways, as I was not able to replicate the same sample.

The key difference between the samples of this paper is in the step that is equivalent to the first step of the sample construction of Neelakantan (2009). In an attempt to reproduce the sample used in Neelakantan (2009), this paper explores three methods to select people with an IRA. The methods deal differently with the answers on the questions that ask about the three largest IRA accounts within the household, and who owns these accounts. A complete list with the variables used in this paper is presented in appendix section A. The questions are only asked to financial respondents of households. A financial respondent is the respondent that is designated to answer questions regarding financial questions for a household. Answers given by the financial respondent on this question are copied to his or her spouse in the RAND dataset. This produces a lot of duplicate observations regarding the IRAs. This paper considers three samples that each deal differently with this issue. The construction of the samples is presented in Table 1, and do-files for STATA are presented in appendix section B.1. The Corrected-spouses(CS) sample assigns IRAs within a household according to the answers of the financial respondent. The Financial Respondents(FR) sample only considers the individuals that are designated as financial respondent, and thus disregards the IRAs of their spouses. The All-respondents (AR) sample considers all these observations as valid observations and selects everyone that has an IRA according to the RAND dataset, without imputing or mutating data. Each sample contains a considerably lower amount of individuals than was the sample of Neelakantan (2009).

4.1.2 Summary statistics - sample composition

The differences in construction result in differences in composition of the samples. This sub-section presents the composition of the samples and the next sub-section considers the differences in statistics on the IRAs of the considered individuals.

Table 2 gives an overview of the composition of each sample. It gives statistics on the distribution of male and female respondents, the mean age and the distribution of categories of marital status, race and education.

Males account for a larger part of the sample in the FR sample than in other samples. This can be explained by the tendency that men are more likely to be the financial respondent of a household. The average age of the samples that are considered in this paper is slightly lower than the average age of the sample of Neelakantan (2009), with the exception of women in the FR sample. This exception may be explained by the substantially higher share of widowed women in this sample compared to the other samples. It is also notable that the average age in the CS sample is relatively low.

Table 1: Sample construction - 2006

	Neelekantan(2009)	**CS**	**FR**	**AR**
RAND HRS 2006 data set	18.469	18.469	18.469	18.469
Is not financial respondent			-5.911	
Does not own an IRA	-13.204	-13.202	-8.997	-12.792
Amount in all owned IRA's unknown		-1.062	-671	-1.009
% risky assets in all owned IRA's unknown		-1.739	-1.195	-1.834
Amount or % risky assets in IRA unknown	-2.109			
Total size	**3.156**	**2.466**	**1.695**	**2.834**

This table shows the various steps taken to derive the different samples. Note that the first step only applies to the FR sample as this is the only sample that explicitly filters for financial respondents. Also, the last step is only applicable to the sample of Neelakantan (2009), as it is not described in the paper how this step divides over the two components amount and % risky assets. For all other samples this step is split in two components to give a clear insight in how the samples are constructed.

Table 2: Descriptive statistics - 2006

	Neelakantan (2009)	**CS**	**FR**	**AR**
Men				
Samplesize	1553(49.2%)	1291(52.4%)	1071(63.2%)	1343(47.4%))
Age	67.2	65.5	66.6	66.0
Marital status (%)				
Married/Partnered	85.1	83.0	80.8	84.8
Divorced/Seperated	7.1	8.8	9.9	7.3
Widowed	5.3	5.1	5.4	5.3
Never Married/Unknown	2.5	3.1	3.9	2.6
Race(%)				
White	94.4	83.7	81.1	84.4
Black	3.0	12.5	14.0	10.9
Other	2.6	3.8	4.9	4.8
Education(%)				
$<$ 12 years	7.6	5.4	3.7	6.1
12 years	25.8	20.3	23.2	23.0
13 - 16 years	20.5	20.4	17.6	20.5
16 years	20.3	22.3	24.9	20.2
$>$ 16 years	25.2	31.1	30.7	29.6
Unknown	0.6	0.6	0.0	0.6
Women				
Samplesize	1603(50.8%)	1175(47.6%)	624(37.86%)	1491(52.6%)
Age	65.3	62.6	65.8	63.7
Marital status (%)				
Married/Partnered	66.4	70.2	43.0	74.3
Divorced/Seperated	11.5	14.0	20.9	12.1
Widowed	19.0	12.7	28.5	10.5
Never Married/Unknown	3.1	3.1	7.7	3.1
Race(%)				
White	93.1	83.7	83.2	87.1
Black	3.8	11.3	13.2	10.0
Other	3.1	4.9	3.6	2.9
Education(%)				
$<$ 12 years	2.0	3.7	4.1	4.0
12 years	12.8	26.9	28.0	30.4
13 - 16 years	15.9	27.0	23.1	29.6
16 years	20.4	19.4	16.5	18.0
$>$ 16 years	21.8	22.4	19.3	16.7
Unknown	27.1	0.6	0.0	1.3

This table shows the distribution of men and women, the average age and distributions of marital status, race and education of the respective genders for every sample that is considered for 2006.

For marital status, race and education, a number of differences and similarities between the samples of this paper and the sample of Neelakantan (2009) is worth mentioning. First, the distribution of marital status for men is rather similar over all samples, while for women this is not the case. Here, the largest difference is seen for the FR sample, which has a low share of Married/Partnered women relative to the other samples. For the CS sample and the AR sample we observe higher rates of Married/-Partnered women. Second, the share of whites is lower for every sample, for both men and women for all samples. Lastly, men are higher-educated in the samples of this paper, and more information tends to be available on the education of women than the sample of Neelakantan (2009).

4.1.3 Summary statistics - IRA composition

The main interest of this paper is risk tolerance. The framework that is set in section 3 uses information on the content and composition of IRAs of individuals to make inferences about risk tolerance. The average content in dollars of the IRAs for men and women is presented in Table 3. Here, it is assumed that the balances presented in Neelakantan (2009) represent the balance at the moment of the interview. It is immediately clear that the differences in content of the IRAs are very large for all samples except for the AR sample. It also seems that the relative difference between the contents of the IRAs for men and women in Neelakantan (2009) ($(193.367-95.037)/95.037 = 103\%$) are quite well represented by the FR sample (101%), and in a lesser extent by the CS sample (162%). The AR sample clearly misrepresents these differences however (1%), as in this sample, women have a much larger amount of money in their account relative to the other samples. This is probably caused by the duplication of the IRA information of the financial respondents to their spouses' observations.

Table 3: Average IRA content in USD($)- 2006

	Neelakantan (2009)	**Corrected-spouses**	**Financial respondents**	**All respondents respondents**
Men	193.367	307.296	222.675	191.175
Women	95.037	117.180	110.685	189.146

This table shows the mean of the total amount of money in all IRA accounts (of which the amount and percentage invested in risky assets is known) of individuals in the samples for 2006 in US dollars($).

While the content of the IRAs is not necessarily important for the methods that are used in this paper to make inferences about risk-tolerance (but rather for investigating its effect on wealth differences), the distribution of the share of risky assets in the IRAs is important, as this is the only parameter that determines the (implied) risk-tolerance of individuals. The mean and standard deviation for this share is presented for the sample of Neelakantan (2009) and the samples of this paper in 4. The shares invested in risky assets for the samples of this paper are quite similar to each other, and are much higher than the percentages found by Neelakantan (2009) while they deviate much less. This finding is very serious, as it means that the results obtained by Neelakantan (2009) are likely different from the results obtained by using their methods on the samples of this paper.

Table 4: Share of risky assets - 2006

	Neelakantan (2009)	Corrected-spouses	Financial respondents	All respondents respondents
Men				
Mean	64.4%	85.2%	84.9%	84.9%
Standard Deviation	41.8%	24.4%	24.1%	24.6%
Women				
Mean	59.5%	85.2%	84.4%	84.1%
Standard Deviation	44.0%	24.6%	24.4%	24.8%

This table shows the mean and standard deviation of the fraction that is invested in risky assets in the IRA accounts (of which the content and the amount invested in risky assets is known) of the individuals for all samples for 2006.

4.2 Wave 12 - 2014

4.2.1 Sample construction

The HRS dataset of 2014 is used in this paper to analyze risk tolerance and its effect on wealth accumulation in a longitudinal setting and to analyze self-assessed risk tolerance. The steps taken in constructing the sample for 2014, are presented in Table 5.

Table 5: Sample construction - 2014

	CS14
RAND HRS 2014	18.747
Does not own IRA	-13.701
Amount in owned IRAs unknown	-873
% shares in owned IRAs unknown	-1.411
Total size	**2.762**

This table shows the steps taken to derive the sample used for the wave of 2014. Note that the steps to construct this sample are equivalent to the steps taken for the CS sample of the wave of 2006.

Note that the steps are equivalent to the steps described in section 4.1.1 for the CS sample. This paper refers to the sample of 2014 as the Corrected-Spouses sample for 2014 (CS214). This paper continues with this way of constructing the sample, as the CS sample seems to represent the sample of Neelakantan (2009) closest in terms of descriptive statistics on the composition of the sample. The FR sample represents the average IRA content better, but it is clearly off in the ratio male to female. The CS14 sample describes a total of 2.762 individuals.

4.2.2 Summary statistics - sample composition

When analyzing the composition of the CS14 sample, it is most interesting to compare this to the composition of the CS sample of 2006 as it was constructed following the same steps. The composition of the sample is presented in Table 6.

The main differences between the samples, is that the sample has shifted over time such that it includes a larger share of women and that the average age has increased. An increase in the life-expectancy over the years (Tuljapurkar et al., 2000) and the fact that a number of respondents of the first cohort is still alive in the wave of 2014 most likely explains the increase in average age. The shift in the share of women in the samples is possibly explained by the higher life-expectancy of women (Wingard, 1984).

Table 6: Descriptive statistics - 2014

	Men	**Women**
Samplesize	1294 (46.9%)	1468 (53.1%)
Age	68.0	66.1
Marital-status (%)		
Married Partnered	77.0	67.0
Divorced/ Seperated	9.7	12.6
Widowed	8.2	16.3
Never Unknown/ Never Married	5.1	4.0
Race(%)		
White	84.9	86.9
Black	9.7	10.1
Other	5.3	3.0
Education(%)		
$<$ 12 years	3.9	5.0
12 years	19.1	24.5
13 - 16 years	23.0	26.4
16 years	23.6	22.6
$>$ 16 years	29.1	20.7
Unknown	1.3	0.7

This table shows the distribution of men and women, the average age and distributions of marital status, race and education of the respective genders for CS14 sample.

4.2.3 Summary statistics - Individual Retirement Accounts

The content of the IRAs and the allocation decisions of the IRA owners have shifted over time. The average IRA content in USD of 2014, which is presented in Table 7, is 10.8% higher for men than for women. This difference is smaller than the difference found for the CS sample in 2006, in which we found a difference of 162%. This decrease in the difference is caused by both an increase in the average amount of USD in the IRAs of women (+81.7%), and a decrease in the average amount of USD in the IRAs of men (−23.2%).

Table 7: Average IRA content in USD($) - 2014

	CS14
Men	235.996
Women	212.923

This table shows the mean of the total amount of money in all IRA accounts (of which the amount and percentage invested in risky assets is known) of the individuals in the CS14 sample.

The average shares of risky assets in the IRA portfolios for men and women are presented in Table 8. These averages are smaller for both men and women when compared to the averages of 2006. Now, a larger difference between men and women seems to exist for this allocation decision. Men tend to have a larger share of their IRA invested in stocks. The standard deviation on the allocation decision has increased for both men and women. This increase is especially notable for women, who show an increase of almost 3 percentage-points on this standard deviation when compared to the CS sample of 2016.

Table 8: Share of stocks - All samples

	Men	**Women**
Mean	81.8%	80.1%
Standard Deviation	24.7%	27.4%

This table shows the mean and standard deviation of the fraction that is invested in risky assets in the IRA accounts (of which the content and the amount invested in risky assets is known) of the individuals for the CS14 sample.

5 Neelakantan Approach

After setting the theoretical framework and exploring the data, this paper starts with using the methods and techniques of Neelakantan

(2009). She estimated distributions for risk-tolerance for men and women and assessed the effect of differences in these distributions on the accumulation of wealth in IRAs. This paper applies her methods and techniques on the different samples described in section 4. This section gives a short overview of the way this is done, what results are obtained from using the methods and what conclusions can be drawn from the results. The analyses are done both from a cross-sectional and from a longitudinal point of view. The section builds upon the theoretical framework set in section 3.

5.1 Methodology

The methodology of Neelakantan (2009) involves two steps. First, she estimates a distribution for the risk tolerance of individuals using a simulated method of moments approach. The second step involves estimating the wealth accumulation for men and women given mean risk tolerance levels from the distributions obtained in the first step.

5.1.1 Estimating the distributions of risk tolerance

The first step of the methodology of Neelakantan (2009) involves estimating distributions for the risk tolerance of men and women separately. She assumes log-normal distributions for risk tolerance as previous literature has shown that this is reasonable to do and it fits the data well. The distributions are defined as follows, where $\frac{1}{\gamma_m^i}$ and $\frac{1}{\gamma_f^i}$ denote the risk tolerance for respectively men and women:

$$log\frac{1}{\gamma_m^i} \sim N(\mu_m, \sigma_m) \quad (3)$$

$$log\frac{1}{\gamma_f^i} \sim N(\mu_f, \sigma_f) \quad (4)$$

The parameters of these distributions are estimated in two steps. The first step starts by taking random draws from a log-normal distribution with arbitrary initial values for the parameters that represent risk-tolerance for all individuals. Using the risk-tolerance draws γ_m^i and γ_f^i, optimal shares of risky assets ${\rho^\star}_m^i$ and ${\rho^\star}_f^i$ are computed by numerically solving equation 2. Subsequently, the first two (weighted) moments (the mean and standard deviation) of ${\rho^\star}_m^i$ and ${\rho^\star}_f^i$ are computed for the second step. The moments are compared with the values observed in the sample, which are presented in Tables 4 and 8 in section 4. If the moments generated by the draws from the distribution are different (this paper allows for an absolute error of 0.01), the parameters of the distributions 3 and 4 are adjusted and the process is repeated until the moments match those of the sample.

5.1.2 The effect of risk tolerance on wealth accumulation

The second step of the methodology of Neelakantan (2009) considers the effect of the differences in distributions of risk tolerance on the wealth accumulation of men and women in their IRAs. She simulates portfolio-return paths for the average male and female using the mean risk tolerance of the obtained distributions from the first step.[6] The length of the period she considers is 38 years. The returns are simulated by randomly drawing $\tilde{r}_t^s$ and setting $r_t^b = 1\%$ for all t. The median portfolio-return path of the simulated return paths is used to obtain an estimate of the expected wealth level

[6] Neelakantan (2009) simulates 10 000 return-paths. This paper simulates 10 000 000 return-paths to decrease the randomness of these draws

Table 9: Parameter estimates and associated risk tolerance - Neelakantan (2009) approach

	Men				**Women**			
			Risk tolerance				**Risk tolerance**	
	μ	σ	Mean	Median	μ	σ	Mean	Median
Neelakantan (2009)	-1.460	0.601	0.278	0.232	-1.564	0.651	0.259	0.209
CS	-1.056	0.273	0.361	0.348	-1.054	0.277	0.362	0.349
FR	-1.058	0.276	0.361	0.347	-1.067	0.26	0.356	0.344
AR	-1.060	0.275	0.360	0.346	-1.069	0.289	0.358	0.343
CS14	-1.098	0.287	0.348	0.334	-1.100	0.285	0.347	0.333

This table shows the parameter estimates for the log-normal distributions for risk tolerance for the different samples, μ and σ using the approach of Neelakantan (2009). It also shows the associated theoretical mean and median of the risk-tolerance obtained from the log-normal distributions with these estimated coefficients. Note that the Neelakantan (2009) estimates are computed using her data, and are slightly different to the estimates in her paper due to randomness.

at the age of retirement.

To asses the effect of differences in risk tolerance between genders on wealth accumulations in IRAs, Neelakantan (2009) compares the relative difference in the estimated expected levels of wealth with the relative difference in the observed levels of wealth.

5.2 Results

5.2.1 Estimating the distributions of risk tolerance

The parameters that are estimated according to the procedure described in section 5.1.1 are presented in Table 9. The differences in average IRA portfolio allocation choices between the samples cause the parameter estimates for the samples of this paper to be smaller in magnitude than the estimates of the sample of Neelakantan (2009). Individuals in the sample of Neelakantan (2009) are on average less risk tolerant. Also in the sample of her paper, there is a clear difference in mean risk tolerance between men and women of 0.019. For the samples of this paper, the largest difference between men and women is found for the FR sample where there is a difference of 0.005. Between the samples that are considered for 2006, there are no big differences for the parameter estimates.

The estimates for the 2014 sample produce larger estimates for both μ and σ than the CS sample of 2006. This indicates that people over time have become less risk tolerant, but also that they deviate more between each other. We also see that men and women are very close in risk tolerance for both years. However, men are slightly more risk tolerant in 2014 while women are slightly more risk tolerant in 2006.

5.2.2 Estimating the effect on wealth

The effects of the differences in risk tolerance on the wealth accumulation in the IRAs of individuals are presented in Table 10. Here r_m and r_f denote the median accumulated return over the 38 years. For example, a man starting with a wealth level of w_0^i, will on average have accumulated a wealth level of $r_m\ w_0^i$ after 38 years. $\Delta\%r$ denotes the difference in expected accumulated return on the IRAs between men and women. $\Delta\%w$ denotes the difference in the wealth levels between men and women as presented in section 4. Lastly, $\Delta\%explained$ is the

Table 10: Wealth accumulation - Neelakantan (2009) approach

	r_m	r_f	$\Delta\%r$	$\Delta\%w$	$\Delta\%explained$
Neelakantan (2009)	6.468	5.929	9.1%	103%	8.8%
CS	9.223	9.260	-0.4%	162%	-0.2%
FR	9.223	9.038	2.0%	101%	2.0%
AR	9.186	9.112	0.8%	1%	81.2%
CS14	8.750	8.714	0.4%	11%	3.8%

This table shows the expected wealth accumulation for men and women and the relative differences for the expected wealth accumulation between men and women. In addition to this, it presents the relative difference in wealth in the IRAs of men and women and the percentage of this wealth difference that can be explained by the difference in expected wealth accumulation due to differences in in average risk tolerance between men and women. The results are obtained using the approach of Neelakantan (2009).

part of the difference in wealth that is explained by the difference in risk-tolerance between men and women.

While the differences in wealth accumulation between men and women are similar across all samples, the part of wealth differences that is explained by this wealth accumulation differs across samples. This is caused by the inequality in wealth difference between men and women between the samples. The wealth differences for the CS sample and the AR sample are especially unlike the differences in the other samples. The part of wealth difference that is explained by the difference in risk tolerance for the CS sample is negative. The average woman is more risk tolerant in this sample and thus has a higher expected accumulation of wealth than men, while the observed level of wealth is lower. The part of wealth difference that is explained by the difference in risk tolerance for the AR sample is relatively large. This is mainly caused by the small difference in wealth between men and women for this sample.

The results for the sample of 2014 with the CS sample of 2006 indicate that the part of the difference in wealth that is explained by the difference in risk tolerance between men and women has increased over time. This is partly caused because men have become more risk tolerant than women over time, but also by the large decrease in the wealth gap between men and women over the years.

5.3 Discussion

The replication of the methods and techniques of Neelakantan (2009) yield results that are hard to interpret. The samples of this paper give not only different results than the sample considered by Neelakantan (2009), they also differ much between each other when trying to answer the question of which part of the difference in wealth between men and women is explained by the differences in risk tolerance. The estimated distributions for risk tolerance are quite similar over the samples of this paper in that they find people to be more risk tolerant than in the sample of Neelakantan (2009), but differences between men and women are unequal between the samples. The differences in wealth-gap between men and women between the samples in combination with the differences in risk tolerance distributions lead to conclusions that are inconsistent across the samples.

It is important to note that the differences in risk tolerance between men and women is very close to zero for the samples of this paper. The method of Neelakantan (2009) has a randomness factor in the simulation. This make it possible that the randomness explains a part of this difference. All in all this makes it very hard to draw strong conclusions in a cross-sectional setting for 2006.

When considering a longitudinal setting, the results suggest that the part of the wealth-gap that is explained by differences in risk tolerance has increased. This is mainly due to the strong decrease of the wealth gap between the genders. It is hard to interpret this result as the difference in risk tolerance is so small for both 2006 and 2014.

6 Maximum Likelihood Approach

As no clear conclusions can be drawn by using the methods of Neelakantan (2009), this paper suggests a different approach. In the approach of Neelakantan (2009) there is a random factor in generating the moments from the shares of risky assets indicated by the drawn risk tolerance levels. This is possible as the observed risk tolerance levels are assumed to be generated from the log-normal distribution of which the parameters are estimated. This section considers an approach that moves the random factor on the results from the the parameter estimation to the observed sample. This is done by estimating a distribution that fits the data best.

6.1 Methodology

For the maximum-likelihood estimation approach, this paper also assumes log-normal distributions for risk-tolerance. The distributions are fitted to the *implied* risk tolerance of the individuals. Implied risk tolerance being the risk tolerance implied by Equation 2 and the chosen allocation between risky and risk-free assets in the IRA of an individual. This implied risk tolerance $\tilde{\gamma}_i$ is obtained for every individual i by numerically solving for the observed ρ_i in Equation 2. The personal sampling weights are accounted for by applying the probability weights in the regression. The maximum-likelihood parameters are estimated for the associated normal distributions for the log risk tolerance:

$$log\frac{1}{\tilde{\gamma}^i_m} \sim N(\tilde{\mu}_m, \tilde{\sigma}_m) \tag{5}$$

$$log\frac{1}{\tilde{\gamma}^i_f} \sim N(\tilde{\mu}_f, \tilde{\sigma}_f) \tag{6}$$

Using the parameter estimates for Distributions 5 and 6, the differences with the estimates of the method-of-moments simulation of section 5 are evaluated, as well as the implications on the effects on wealth accumulation.

6.2 Results

6.2.1 Estimating the distributions of risk tolerance

The parameter estimates obtained by using the maximum likelihood approach are presented in Table 11. Using this method, we observe a higher average risk tolerance for women than for men for both the CS and the FR samples. The average risk tolerance is higher for men than for women in the AR and the CS14 samples. Note that this does not necessarily contradict the observed allocation in Table 4 and Table 8. The standard deviation concerning the share of risky assets is higher for women and the relation between the optimal share of risky assets and risk tolerance from Equation 2

Table 11: Parameter estimates and associated risk tolerance - Maximum-Likelihood Approach

	Men				**Women**			
			Risk tolerance				**Risk tolerance**	
	μ	σ	Mean	Median	μ	σ	Mean	Median
CS	-1.085	0.492	0.381	0.338	-1.090	0.516	0.384	0.336
FR	-1.081	0.442	0.374	0.339	-1.103	0.544	0.385	0.332
AR	-1.089	0.484	0.378	0.337	-1.103	0.506	0.377	0.332
CS14	-1.106	0.526	0.380	0.331	-1.091	0.473	0.376	0.336

This table shows the parameter estimates for the log-normal distributions for risk tolerance for the different samples, μ and σ using the maximum-likelihood approach. It also shows the associated theoretical mean and median of the risk-tolerance obtained from the log-normal distributions with these estimated coefficients.

Table 12: Wealth accumulation - Maximum-Likelihood Approach

	r_m	r_f	$\Delta\%r$	$\Delta\%w$	$\Delta\%explained$
CS	9.985	10.103	-1.2%	162%	-0.7%
FR	9.714	10.142	-4.2%	101%	-4.2%
AR	9.868	9.830	0.4%	1%	39.4%
CS14	9.946	9.791	1.6%	11%	14.7%

This table shows the expected wealth accumulation for men and women and the relative differences for the expected wealth accumulation between men and women. In addition to this, it presents the relative difference in wealth in the IRAs of men and women and the percentage of this wealth difference that can be explained by the difference in expected wealth accumulation due to differences in in average risk tolerance between men and women. The results are obtained using maximum-likelihood approach.

is non-linear.

The estimates in Table 11 are larger in magnitude than those of Table 9, especially for σ (almost twice as high). This results in a higher mean risk tolerance for every sample for both men and women when the maximum likelihood method is considered instead of the methods-of-moments simulation. .

6.2.2 Estimating the effect on wealth

The differences in the mean risk tolerance found for men and women between using the approach of Neelakantan (2009) and using the maximum likelihood approach imply that there are different effects on wealth accumulation as well. The effects of the mean risk tolerances are presented in Table 12. As expected, the parts of the wealth differences that are explained by the differences in risk tolerance are much different from those found in section 5.2.2. According to the distributions of risk tolerance, women are expected to have a higher level of wealth than men for the CS sample and the FR sample. The result for the FR sample is especially interesting as the explained part has switched sign. For the AR sample we find that the explained part is about half of that when using the Neelakantan (2009) approach, and for the 2014 sample we find that the explained part is about four times higher between the two approaches.

6.3 Discussion

The results obtained by using the simulated method of moments and the maximum likelihood approach are different. This is caused by the differences in the estimated risk tolerance distributions, with the estimated standard deviation playing a large part in influencing the mean risk tolerance. The sensitivity to the choice of method is different per sample. The CS sample leads to similar conclusions across both methods. The results are still different from the results obtained by using the sample of Neelakantan (2009), as we find that women are expected to earn more because of a higher average risk tolerance than men. The FR sample leads to different conclusions when using both methods. Where men are on average more risk tolerant than women when the methods of Neelakantan (2009) are considered, women are more risk tolerant when the maximum-likelihood method is applied. The AR seems to have a too small difference in wealth levels between men and women to be able to draw strong conclusions on the part that of this wealth-gap that is explained by differences in risk tolerance. The small wealth-gap results in very sensitive outcomes on the part of wealth difference explained to the differences in risk tolerance. Overall, both methods do not provide consistent results in a cross sectional setting for 2006 across the different samples. When we compare the CS sample and the CS14 sample, we find for both methods that the explained part of the wealth difference by the differences in risk tolerance has changed sign and has increased in order of magnitude. This indicates that the effect of differences in risk tolerance on the inequality in wealth between men and women has increased over time.

7 Explaning Risk Tolerance

The methods of sections 5 and 6 did not prove to be very useful in providing consistent results. This section considers a method that aims to explain risk tolerance over more characteristics of individuals than their gender. This is done for the CS and the CS14 samples.

7.1 Methodology

The use of implied risk tolerance as done in section 6 allows for a close examination of the differences in said risk tolerance between individuals. To examine these differences, this paper aims to explain the implied risk tolerance by regressing the implied risk tolerance on a number of characteristics of the respondents considered in the sample according to the following model:

$$\frac{1}{\tilde{\gamma}} = X\tilde{\beta} + \epsilon \tag{7}$$

Here $\frac{1}{\tilde{\gamma}}$ denotes the risk tolerance and X denotes the characteristics of the respondents. The characteristics that are considered are gender, age, age-squared, marital status, education and race. For age, the age and the squared age are considered as Hallahan et al. (2004) has shown that using a non-linear relation might be most appropriate. Robust standard errors are used, using the individual probability weighting scheme of the HRS.

The obtained estimates for $\tilde{\beta}$ can be investigated in order to interpret the effect of the characteristics on the risk tolerance. This might not only give more insights in the effect of gender on risk-tolerance when controlling for other characteristics, but it can also shed light on the effects of different characteristics.

Table 13: Coefficient estimates and robust standard errors - explaining risk tolerance for 2006 and 2014

	2006		2014	
	Coef.	**Robust Std. Err.**	**Coef.**	**Robust Std. Err.**
Female	0.00012	0.00609	0.00409	0.00713
Age	-0.00253	0.00411	-0.00117	0.00444
Age^2	0.00002	0.00003	0.00001	0.00003
Marital status				
Unknown	0.02789(*)	0.01602	-0.01097	0.02291
Divorced/Seperated	-0.00336	0.01468	0.01446	0.01393
Married	0.01291	0.01080	0.00344	0.01016
Education				
12 years	0.01224	0.01608	-0.00076	0.01622
Between 12 and 16 years	0.00801	0.01626	-0.00120	0.01593
16 years	0.00060	0.01636	-0.00272	0.01633
>16 years	0.00489	0.01604	-0.00783	0.01633
Race				
White	-0.03353(***)	0.00967	0.02159	0.01748
Black	-0.03784(***)	0.01150	0.03842(**)	0.01850
Intercept	0.46479(***)	0.13491	0.38235(***)	0.15020
R^2	0.0103		0.0081	

This table presents the parameter estimates and the corresponding robust standard errors for the regression model presented in Equation 7, for both 2006 and 2014. The robust standard errors are used to determine the levels of significance. $* : p < 0.1$; $** : p < .05$; $*** : p < .01$. The table also reports the R^2 for both models.

7.2 Results

The coefficient estimates and the corresponding robust standard errors of the regression model of Equation 7 are presented in Table 13. Most characteristics do not have coefficients that are different from zero with a probability of more than 0.9. The Unknown Marital status seems to have a positive effect on risk tolerance with a probability of at least 0.9. Furthermore, we find that whites and blacks tend to have lower risk tolerance levels than individuals from other races, with a significance level of $p < 0.01$ for both races. For 2014, blacks tend to be more risk tolerant than people from other races, with a significance level of $p < 0.05$. For both models this paper finds low R^2 values, respectively of 0.0103 and 0.0081, meaning that the models only explain a very small part of the variation in risk tolerance (about 1%).

7.3 Discussion

By considering multiple characteristics in explaining differences in risk tolerance between individuals, this paper finds no significant evidence that gender affects risk tolerance for individuals. This paper does find significant evidence for race having a significant effect on risk tolerance. In particular it seems that for 2006 whites and blacks tend to be less risk tolerant than individuals from other races. For 2014, black individuals tend to be more risk tolerant than individuals who are not white or black. These findings might explain part of the discrepancies found in sections 5 and 6, as the composition of the samples all differ in race distribution. Most notably, the sample used by Neelakantan (2009) contains a significantly larger proportion of white people, who tend to be less risk tolerant according to the findings this section. This might explain why this paper finds higher levels of mean risk tolerance among men and women for 2006 in general. Furthermore, this paper finds that the model that it uses to explain risk tolerance explains only a small part of the variance in risk tolerance for both 2006 and 2014 (about 1%).

8 Self-reported Risk Tolerance

In section 7, this paper found that it is hard to explain the risk tolerance of individuals (as the models that the section considers explain respectively 1.03% and 0.801 % of the variation in risk tolerance). It can be very valuable in many applications to understand what affects risk tolerance, or even to predict risk tolerance for individuals. In this section, this paper analyses whether self-reported risk tolerance is accurate, and whether this subjective risk tolerance is useful in explaining risk tolerance combined with the characteristics that are considered in 7.

8.1 Methodology

In wave 12 of the HRS (2014), respondents are asked to rate themselves on a scale of 0 to 10 on their willingness to take risks. Literature has shown that self-reported risk tolerance generally accords with more objective measures (Hallahan et al., 2004). This section starts by analyzing whether this self reported risk tolerance can be used as a proxy of risk tolerance. This is done by considering the following model:

$$\frac{1}{\tilde{\gamma}} = \beta_1 + \beta_2 \frac{1}{\hat{\gamma}} + \epsilon \tag{8}$$

Here $\frac{1}{\tilde{\gamma}}$ denotes the implied-risk tolerance from section 6 and $\frac{1}{\hat{\gamma}}$ denotes the self-reported risk-tolerance. By regressing the implied risk tolerance on the self reported risk tolerance, it can be evaluated whether the self-reported risk tolerance can serve as an alternative to the implied risk-tolerance from the portfolio choice in the IRAs.

After this analysis, we will consider an extended form of model 7. The model is extended by including the self-reported risk tolerance:

$$\frac{1}{\tilde{\gamma}} = X\boldsymbol{\beta}_1 + \beta_2 \frac{1}{\hat{\gamma}} + \epsilon \tag{9}$$

Here $\frac{1}{\tilde{\gamma}}$ and X denote the implied risk-tolerance and the set of characteristics from section 7 respectively. $\frac{1}{\hat{\gamma}}$ denotes the self-reported risk-tolerance.

By analyzing the estimated coefficients and the corresponding robust standard errors for model 9, inferences can be made on whether the self-reported risk tolerance can be helpful in explaining implied risk-tolerance.

8.2 Results

When regressing the implied risk tolerance on the subjective risk tolerance for all individuals, this paper finds the following model as a result:

$$\frac{1}{\tilde{\gamma}} = \underset{(0.0064)}{0.3662^{***}} - \underset{(0.0179)}{0.0200} \frac{1}{\hat{\gamma}} \quad (10)$$

here, the robust standard errors are presented below the coefficient estimates in parentheses, and the levels of significance ($*: p < 0.1$; $**: p < .05$; $***: p < .01$) are shown for each coefficient.

We find no significant evidence that the subjective risk tolerance has an effect on the implied risk tolerance using model 8. This means that, considered alone, the subjective risk tolerance does not explain the risk tolerance of an individual.

Table 14 presents the coefficients of regression model 9. It is particularly interesting to consider whether there is an effect of this subjective risk tolerance on the implied risk tolerance when controlling for other characteristics. This paper finds no significant evidence of an effect of subjective risk tolerance on implied risk tolerance. There is an improvement in the R^2 of the model however, indicating that the subjective risk tolerance might still be useful to take into account when explaining risk tolerance. Note that there is only an improvement of 0.0015 however, so the model still does not explain the variation in risk tolerance well.

8.3 Discussion

This paper does not find the self-reported risk tolerance of the HRS of 2014 to be a good proxy for risk tolerance. When it is included in a model together with other personal characteristics, its helps explain *some* variation in risk tolerance. There is no significant effect of the self-reported risk tolerance on the implied risk tolerance however. For further research it can be useful to include different kinds of self-reported risk tolerance measures. It may also prove useful to consider different functional relations between self-reported risk tolerance and implied risk tolerance.

9 General Discussion

In this study we found that the findings of Neelakantan (2009) were difficult to replicate. This paper found no consistent results for the effect of the differences in risk tolerance on the wealth accumulation between genders across the samples it considered. We did find indications that between 2006 and 2014 the part of the wealth gap between men and women that is explained by differences in risk tolerance has increased. In addition this paper found that when controlling for other characteristics, ethnicity has a significant effect on risk tolerance, while gender does not. Lastly, this paper found that self-reported risk tolerance helps improve the explanation of risk tolerance slightly, but does not serve well as a proxy for risk tolerance. The findings of this paper apply to the U.S. elderly population (age 50 and above), as the HRS dataset was used for the analyses.

The inconsistency in the results between the samples for both the methodology of Neelakantan (2009) and the maximum likelihood approach raises questions about the conclusions of Neelakantan (2009). The conclusions seem to be very sensitive to the sample construction. For further research it is valuable to evaluate the construction of the samples in great detail as certain assumptions were made in this paper to try to replicate the sample of Neelakantan (2009), and to investigate what is the most appropriate sample to consider. The most important Stata and MATLAB code that is used for part of this paper that aims to replicate the

Table 14: Coefficient estimates and robust standard errors - explaining risk tolerance including subjective risk tolerance for 2014

	Coef.	**Robust Std. Err.**
Female	0.00368	0.00724
Age	-0.00005	0.00467
Age^2	0.00000	0.00003
Marital Status		
Unknown	-0.01063	0.02390
Divorced/Seperated	0.01497	0.01410
Married	0.00392	0.01041
Education		
12 years	-0.00217	0.01628
Between 12 and 16 years	-0.00121	0.01597
16 years	-0.00283	0.01633
$>$ 16 years	-0.00839	0.01634
Race		
White	0.02407	0.01780
Black	0.04088(**)	0.01889
Subjective Risk Tolerance	-0.01419	0.01850
Intercept	0.35048(**)	0.15646
R^2	0.0096	

This table shows the parameter estimates and the corresponding robust standard errors for the regression model presented in Equation 8 for 2014. The robust standard errors are used to determine the levels of significance. $* : p < 0.1$; $** : p < .05$; $*** : p < .01$. The table also reports the R^2 for the model.

methods of Neelakantan (2009) is included in the appendix section B.1 and B.2 respectively.

This paper did find differences in risk tolerance and wealth between 2006 and 2014, but did not try to explain these differences. For further research it can be interesting to find the causes of these differences. It might be that the banking crisis of 2008 has had an effect on the risk tolerance of individuals, especially in a retirement savings context as is considered in this paper. It would be worth investigating this issue further. Both contradictions

The lack of a significant effect of gender on risk tolerance when controlling for more characteristics, contradicts the findings of Yao and Hanna (2005) and Grable and Joo (2000). The significant effect of ethnicity on risk tolerance that we found in this paper is also not in line with the findings of the previously discussed literature (Yao and Hanna (2005) and Sung and Hanna (1996)). The difference in findings might be explained by the difference in the sample populations, as they used samples that were not constructed using the HRS dataset.

The finding that self-reported risk tolerance does not serve well as a proxy for risk tolerance is in contrast with the findings of Hallahan et al. (2004). This might be because a different measure was used in their analysis. This measure does not come from the HRS dataset. It might be worth investigating what self-assessed measures are useful as a proxy or in explaining risk tolerance and what makes them good.

In addition to the previously discussed possibilities for further research, there are a number of topics for further research that could be suggested that could provide valuable insights into the topic of risk tolerance and into the findings of this paper. First, it can be interesting to use the methods of Neelakantan (2009) on the same sample that she used, to confirm her findings. Second, it might be interesting to consider the subjective risk tolerance not as a continuous variable, but to use dummies for different categories such as 'risk tolerant', 'risk neutral' and 'risk averse'. Lastly, it may be interesting to consider more characteristics than the ones that were considered in this paper to explain risk tolerance, as almost none proved to have significant effects on risk tolerance in the model of this paper.

References

Finke, M. S. and Huston, S. J. (2003). The brighter side of financial risk: Financial risk tolerance and wealth. *Journal of Family and Economic Issues*, 24(3):233–256.

Grable, J. E. (1997). *Investor risk tolerance: Testing the efficacy of demographics as differentiating and classifying factors*. PhD thesis, Virginia Tech.

Grable, J. E. and Joo, S.-H. (2000). A cross-disciplinary examination of financial risk tolerance. *Consumer Interests Annual*, 46:151–157.

Hallahan, T. A., Faff, R. W., and McKenzie, M. D. (2004). An empirical investigation of personal financial risk tolerance. *Financial Services Review*, 13(1):57–78.

Jagannathan, R., Kocherlakota, N. R., et al. (1996). Why should older people invest less in stocks than younger people. *Federal Reserve Bank of Minneapolis Quarterly Review*, 20:11–20.

Jianakoplos, N. and Bernasek, A. (1998). Are women more risk averse? *Economic Inquiry*, 36(4):620–630.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, 51(3):247–257.

Neelakantan, U. (2009). Estimation and impact of gender differences in risk tolerance. *Economic Inquiry*, 48(1):228–233.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1):122–136.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *UC Berkeley: Departmenet of Economics, UCB.*

Sung, J. and Hanna, S. (1996). Factors related to risk tolerance. *Journal of Financial Counseling and Planning*, 7:11–19.

Tuljapurkar, S., Li, N., and Boe, C. (2000). A universal pattern of mortality decline in the G7 countries. *Nature*, 405(6788):789.

Wang, H. and Hanna, S. (1997). Does risk tolerance decrease with age? *Journal of Financial Counseling and Planning*, 8(2).

Wingard, D. L. (1984). The sex differential in morbidity, mortality, and lifestyle. *Annual review of public health*, 5(1):433–458.

Yao, R., Gutter, M., and Hanna, S. (2005). The financial risk tolerance of blacks, hispanics and whites. *Journal of Financial Counseling and Planning*, 16(1).

Yao, R. and Hanna, S. D. (2005). The effect of gender and marital status on financial risk tolerance. *Journal of Personal Finance*, 4(1):66.

10 Appendix

A. List of used RAND variables

Variable name dataset	**Dataset**	**Use**
kq165_x, x $\in \{1,2,3\}$	Fat File 2006	Who owns the largest three IRA accounts
kq166_x, x $\in \{1,2,3\}$	Fat File 2006	Content of the IRA accounts
kq514_x, x $\in \{1,2,3\}$	Fat File 2006	Part invested in risky assets
kb063	Fat File 2006	Marital status
kz216	Fat File 2006	Education
ka019	Fat File 2006	Age
gend_r	Fat File 2006	Gender
oq165_x, x $\in \{1,2,3\}$	Fat File 2014	Who owns the largest three IRA accounts
oq166_x, x $\in \{1,2,3\}$	Fat File 2014	Content of the IRA accounts
oq514_x, x $\in \{1,2,3\}$	Fat File 2014	Part invested in risky assets
ob063	Fat File 2014	Marital status
oz216	Fat File 2014	Education
oa019	Fat File 2014	Age
gend_r	Fat File 2014	Gender
RARACEM	Longitidunal RAND File	Race

B. Stata and Matlab Code

B.1 Stata do files

Construction CS Sample

```
version 14.1

set maxvar 30000
use "\\campus.eur.nl\users\home\431184pv\Documents\Scriptie\h06f3a_STATA\h06f3a.dta", clear


replace kq165_1 = abs(kq165_1 - 3) if (hhidn == hhidn[_n-1] & kq165_1 == kq165_1[_n-1] & kfin_rhp == hhidpn[_n-1]) | (hhidn == hhidn[_n+1] & kq165_1 == kq165_1[_n+1] & kfin_rhp == hhidpn[_n+1])

replace kq165_2 = abs(kq165_2 - 3) if (hhidn == hhidn[_n-1] & kq165_2 == kq165_2[_n-1] & kfin_rhp == hhidpn[_n-1]) | (hhidn == hhidn[_n+1] & kq165_2 == kq165_2[_n+1] & kfin_rhp == hhidpn[_n+1])

replace kq165_3 = abs(kq165_3 - 3) if (hhidn == hhidn[_n-1] & kq165_3 == kq165_3[_n-1] & kfin_rhp == hhidpn[_n-1]) | (hhidn == hhidn[_n+1] & kq165_3 == kq165_3[_n+1] & kfin_rhp == hhidpn[_n+1])

keep if kq165_1 == 1 | kq165_2 == 1 | kq165_3 == 1

keep if kq165_1 == 1 & (kq166_1 != 99999999 & kq166_1 != 99999998 & !missing(kq166_1)) | kq165_2 == 1 & (kq166_2 != 99999999 & kq166_2 != 99999998 & !missing(kq166_2)) | kq165_3 == 1 & (kq166_3 != 99999999 & kq166_3 != 99999998 & !missing(kq166_3))


keep if kq165_1 == 1 & (kq166_1 != 99999999 & kq166_1 != 99999998 & !missing(kq166_1) & (kq514_1 != 999 & kq514_1 != 998 & !missing(kq514_1))) |kq165_2 == 1 & (kq166_2 != 99999999 & kq166_2 != 99999998 & !missing(kq166_2)& (kq514_2 != 999 & kq514_2 != 998 & !missing(kq514_2))) | kq165_3 == 1 & (kq166_3 != 99999999 &kq166_3 != 99999998 & !missing(kq166_3)& (kq514_3 != 999 & kq514_3 != 998 & !missing(kq514_3)))
```

Construction FR Sample

```
version 14.1

set maxvar 30000
use "\\campus.eur.nl\users\home\431184pv\Documents\Scriptie\h06f3a_STATA\h06f3a.dta", clear

keep if kfin_rhp == hhidpn

keep if kq165_1 == 1 | kq165_2 == 1 | kq165_3 == 1
```

```
keep if kq165_1 == 1 & (kq166_1 != 99999999 & kq166_1 != 99999998 & !missing(kq166_1)) | kq165_2 == 1 & (kq166_2 != 99999999 & kq166_2 != 99999998 & !missing(kq166_2))
| kq165_3 == 1 & (kq166_3 != 99999999 & kq166_3 != 99999998 & !missing(kq166_3))


keep if kq165_1 == 1 & (kq166_1 != 99999999 & kq166_1 != 99999998 & !missing(kq166_1) & (kq514_1 != 999 & kq514_1 != 998 & !missing(kq514_1))) | kq165_2 == 1 & (kq166_2 != 99999999 & kq166_2 != 99999998 & !missing(kq166_2)& (kq514_2 != 999 & kq514_2 != 998 & !missing(kq514_2))) | kq165_3 == 1 & (kq166_3 != 99999999 & kq166_3 != 99999998 & !missing(kq166_3)& (kq514_3 != 999 & kq514_3 != 998 & !missing(kq514_3)))
```

Construction AR Sample

```
version 14.1

set maxvar 30000
use "\\campus.eur.nl\users\home\431184pv\Documents\Scriptie\h06f3a_STATA\h06f3a.dta", clear

keep if kq165_1 == 1 | kq165_2 == 1 | kq165_3 == 1


keep if kq165_1 == 1 & (kq166_1 != 99999999 & kq166_1 != 99999998 & !missing(kq166_1)) | kq165_2 == 1 & (kq166_2 != 99999999 & kq166_2 != 99999998 & !missing(kq166_2)) | kq165_3 == 1 & (kq166_3 != 99999999 & kq166_3 != 99999998 & !missing(kq166_3))


keep if kq165_1 == 1 & (kq166_1 != 99999999 & kq166_1 != 99999998 & !missing(kq166_1) & (kq514_1 != 999 & kq514_1 != 998 & !missing(kq514_1))) | kq165_2 == 1 & (kq166_2 != 99999999 & kq166_2 != 99999998 & !missing(kq166_2)& (kq514_2 != 999 & kq514_2 != 998 & !missing(kq514_2))) | kq165_3 == 1 & (kq166_3 != 99999999 & kq166_3 != 99999998 & !missing(kq166_3)& (kq514_3 != 999 & kq514_3 != 998 & !missing(kq514_3)))
```

Construction CS14 Sample

```
version 14.1

set maxvar 30000
use "\\campus.eur.nl\users\home\431184pv\Documents\Scriptie\h06f3a_STATA\h014f2a.dta", clear


replace oq165_1 = abs(oq165_1 - 3) if (hhidn == hhidn[_n-1] & oq165_1 == oq165_1[_n-1] & ofin_rhp == hhidpn[_n-1]) | (hhidn == hhidn[_n
```

```
+1] & oq165_1 == oq165_1[_n+1] & ofin_rhp == hhidpn[_n+1])

replace oq165_2 = abs(oq165_2 - 3) if (hhidn == hhidn[_n-1] & oq165_2 == oq165_2[_n-1] & ofin_rhp == hhidpn[_n-1]) | (hhidn == hhidn[_n+1] & oq165_2 == oq165_2[_n+1] & ofin_rhp == hhidpn[_n+1])

replace oq165_3 = abs(oq165_3 - 3) if (hhidn == hhidn[_n-1] & oq165_3 == oq165_3[_n-1] & ofin_rhp == hhidpn[_n-1]) | (hhidn == hhidn[_n+1] & oq165_3 == oq165_3[_n+1] & ofin_rhp == hhidpn[_n+1])

keep if oq165_1 == 1 | oq165_2 == 1 | oq165_3 == 1

keep if oq165_1 == 1 & (oq166_1 != 99999999 & oq166_1 != 99999998 & !missing(oq166_1)) | oq165_2 == 1 & (oq166_2 != 99999999 & oq166_2 != 99999998 & !missing(oq166_2)) | oq165_3 == 1 & (oq166_3 != 99999999 & oq166_3 != 99999998 & !missing(oq166_3))


keep if oq165_1 == 1 & (oq166_1 != 99999999 & oq166_1 != 99999998 & !missing(oq166_1) & (oq514_1 != 999 & oq514_1 != 998 & !missing(oq514_1))) |oq165_2 == 1 & (oq166_2 != 99999999 & oq166_2 != 99999998 & !missing(oq166_2)& (oq514_2 != 999 & oq514_2 != 998 & !missing(oq514_2))) | oq165_3 == 1 & (oq166_3 != 99999999 &oq166_3 != 99999998 & !missing(oq166_3)& (oq514_3 != 999 & oq514_3 != 998 & !missing(oq514_3)))
```

B.2 Matlab code

Code used for solving Equation 2

```
function ev = risky_solver(rho, gamma)
%% PURPOSE: Generate expected value of the equation used in Neelankantan(2009) to determine the optimal share of risky-assets in an IRA given a certain level of risk aversion
%
%----------------------------------------------------------
% USAGE:       ev = risky_solver(rho, gamma)
% where:       rho = the share of risky-assets in an IRA
%              gamma = level of risk-aversion of an individual
%
%----------------------------------------------------------
% RETURNS: the expected value of the marginal utility
%----------------------------------------------------------

% written by:
% Pim van der Voet, Erasmus School of Economics
% Erasmus university Rotterdam
% pimvandervoet@gmail.com
%----------------------------------------------------------

%Declare some initial values, given by Neelankantan(2009), for
```

```
    returns on
%assets
risky_good = 0.2703;
risky_med = 0.13;
risky_low = -0.1525;
bond = 0.01;
returnsvector = [risky_good, risky_med, risky_low, bond];
ev = 0;

for i = 1:3
    ev = ev + (1/3) * (( ( (1+returnsvector(i)) * rho + (1+
        returnsvector(4))*(1-rho) )^(-1*gamma)) * (returnsvector(i) -
        returnsvector(4)));
end

end
```

Code used to replicate the approach of Neelakantan (2009)

```
function output = method_moments_simulator(n_gendergroup, mu_observed, sigma_observed, mu_draw ,sigma_draw, sufficient_condition , mu_found, sigma_found)
%% PURPOSE: obtain parameter estimates for lognormal distribution on risk-tolerance by method of moments simulation for a gender subgroup
%
%------------------------------------------------------------
% USAGE: moment_estimates = method_moments_simulator(n_gendergroup, mu_observed, sigma_observed, mu ,sigma, mu_found, sigma_found)
% where:       n_gendergroup = the size of the group you want to analyze
%              mu_observed = the mean of the group you want to analyze
%              sigma_observed =  the standard deviation of the group you
%              want to analyze
%              mu_draw = the current mean used to draw the risk-tolerance
%              parameters
%              sigma_draw = the current sigma used to draw the risk-tolerance
%              parameters
%              sufficient_condition = the maximum error that is allowed to
%              give a sufficient in percentages
%              mu_found = a boolean which is false if the mean used to draw
%              the risk-tolerances is not sufficient according to the given
%              condition
%              sigma_found = a boolean which is false if the sigma used to draw
%              the risk-tolerances is not sufficient according to the given
%              condition
%
%------------------------------------------------------------
% RETURNS: mu and sigma of the lognormal distribution
%------------------------------------------------------------

% written by:
% Pim van der Voet, Erasmus School of Economics
% Erasmus university Rotterdam
% pimvandervoet@gmail.com
%-----------------------------------------------------------

%Set seed
rng('default') % For reproducibility

%Initialize
```

```
if isempty(mu_found)
    mu_found = false;
end
if isempty(sigma_found)
    sigma_found = false;
end

counter = 0;
counter_mu = 0;
counter_sigma = 0;

while (sigma_found == false || mu_found == false)
    counter = counter +1;
    disp(counter);
    %We first draw risk-tolerance parameters from a lognormal distribution
    draws = lognrnd(mu_draw, sigma_draw, n_gendergroup, 1);
    draws = 1./draws;

    %For each draw obtain optimal share risky assets
    optimal_risky = zeros(length(draws),1);
    parfor i = 1:length(draws)
        optimal_risky(i) = fsolve(@(rho)risky_solver(rho,draws(i)),0);
    end

    %Get moments of generated optimal shares
    mu_realized = mean(optimal_risky);
    sigma_realized = std(optimal_risky);

    %Compare moments and decide if necessary to get new ones
    %First check whether mu is correct already

    if mu_found == false
        %Control if mu is good enough
        counter = counter + 1;
        counter_mu = counter_mu+1;
        if abs((mu_realized/mu_observed) - 1) < sufficient_condition || counter >1000
            %if so, lock it
            mu_found = true;
        else
            %else, change it with the % that our observed sigma was deviating,
            %plus 0.0001, to avoid that we stay at 0
            mu_draw = mu_draw + sign((mu_realized/mu_observed) - 1) * -1 * 0.001;
        end
%          %Call function again with new mu_draw, to go to next iteration
```

```
%          method_moments_simulator(n_gendergroup, mu_observed, sigma_observed, mu_draw ,sigma_draw, sufficient_condition, mu_found, sigma_found);
    end

    %Next consider sigma
    if sigma_found == false && mu_found == true
        %Control if sigma is good enough
        counter = counter + 1;
        counter_sigma = counter_sigma+1;
        if abs((sigma_realized/sigma_observed) - 1) < sufficient_condition || counter > 2000
            %if so, lock it
            sigma_found = true;
        else
            %else, change it with the % that our observed sigma was deviating,
            %plus 0.0001, to avoid that we stay at 0
            sigma_draw = sigma_draw + sign((sigma_realized/sigma_observed) - 1) * -1 * 0.001;
        end
         %Call function again with new sigma_draw, to go to next iteration
%          method_moments_simulator(n_gendergroup, mu_observed, sigma_observed, mu_draw ,sigma_draw, sufficient_condition, mu_found, sigma_found);
    end

end
%If no more optimization can be done, print outcomes
display(mu_draw);
display(sigma_draw);
output = [mu_draw, sigma_draw, counter_mu, counter_sigma];

end
```

Code used to simulate return paths

```
function median_acumulated_return = return_path_simulator(number_paths, fraction_stocks)
%% PURPOSE: Simulate stock paths using the discrete outcomes used in Neelakantan 2009
%
%---------------------------------------------------
% USAGE: output = return_path_simulator(number_paths, fraction_stocks)
% where:      number_paths = the number of paths that get simulated
%             fraction_stocks = the fraction of wealth invested in stocks
%
%---------------------------------------------------
% RETURNS: the mean and median accumulated wealth over a period of 38 years of the
% simulated paths given the fraction of wealth invested in stocks
%---------------------------------------------------

% written by:
% Pim van der Voet, Erasmus School of Economics
% Erasmus university Rotterdam
% pimvandervoet@gmail.com
%---------------------------------------------------

%Set seed
rng('default') % For reproducibility

%Declare return initial values
risky_good = 0.2703;
risky_med = 0.13;
risky_low = -0.1525;
bond = 0.01;
fraction_bonds = 1-fraction_stocks;
returnsvector = [risky_good, risky_med, risky_low, bond];
realized_returns = [];

for i = 1:number_paths
    accumulated_return = 1;
    return_generator = unidrnd(3, 38);
    for years = 1:38
        drawn_stockreturn = returnsvector(return_generator(years));
        accumulated_return = (1+bond) * fraction_bonds * accumulated_return + (1+drawn_stockreturn) * fraction_stocks * accumulated_return;
    end
    realized_returns(i) = accumulated_return;
end


median_acumulated_return = median(realized_returns);
```

```
end
```