



DISCUSSION PAPER PI-2004

One size fits all: How many default funds does a pension scheme need?

David Blake, Mel Duffield, Ian Tonks, Alistair Haig, Dean Blower & Laura MacPhee

February 2020

ISSN 1367-580X

The Pensions Institute
Cass Business School
City, University of London
106 Bunhill Row
London EC1Y 8TZ
UNITED KINGDOM
<http://www.pensions-institute.org/>

One size fits all: How many default funds does a pension scheme need?

By

David Blake

Mel Duffield

Ian Tonks

Alistair Haig

Dean Blower

and

Laura MacPhee*

February 2020

* Pensions Institute, Cass Business School; USS; University of Bristol; University of Edinburgh; USS; and USS, respectively. Part of this paper was written while Ian Tonks was visiting CEPAR, UNSW in Sydney, and he is grateful for the comments of Hazel Bateman and Susan Thorp.

Abstract

In this paper, we analyse the number of default investment funds appropriate for an occupational defined contribution pension scheme. Using a unique dataset of member risk attitudes and characteristics from a survey of a large UK pension scheme, we apply cluster analysis to identify two distinct groups of members in their 40s and 50s. Further analysis indicated that the risk attitudes of the two groups were not significantly different, allowing us to conclude that a single lifestyle default fund is appropriate.

Key words: investment choices, cluster analysis, risk attitudes, risk capacity, defined contribution pension schemes

JEL: G11, G41

1. Introduction

Around the world occupational pension schemes are shifting from defined benefit (DB) to defined contribution (DC). These changes have implications for scheme governance (OECD, 2009). Under a DB arrangement, employers bear the investment risks of fund performance. In contrast, a DC scheme shifts investment risks to the individual member. This means that DB scheme members receive a pension defined by the scheme rules (conditional on the solvency of the scheme sponsor), whereas DC scheme members face uncertainty of pension outcomes.

As a consequence, pension policy makers and regulators have recognised the need for a code of practice concerning the governance of DC schemes. One of the central governance requirements is the provision of a default investment fund. As recognised by the UK Pensions Commission (2005) *Second Report* (p. 378), there are two reasons for needing a default investment vehicle. First, some contributors may fail to inform the scheme of their asset allocation preferences, and second some members may “not feel well-equipped to make asset allocation decisions”.

In the UK since 2012, DC pension schemes have been required to offer a minimum of one default investment fund for members who do not wish to exercise an investment choice (under section 17(2)(b) of the Pensions Act 2008).¹ The design and suitability of these default funds have been enshrined in The Pension Regulator’s (TPR) code of practice for DC schemes which came into force in July 2016 (TPR (2016)). TPR (2019) identifies five key governance requirements (KGRs) that DC schemes should satisfy, and, in particular, KGR5 states “Trustee boards must ensure the default investment strategy is suitably designed for their members” (p. 1). The design of a suitable

¹ Section 17(2)(b) of the Pensions Act 2008 states “no provision of the scheme requires [an active member of the scheme] to express a choice in relation to any matter, or to provide any information, in order to remain an active member” (<http://www.legislation.gov.uk/ukpga/2008/30/section/17>). This has been interpreted to mean that the scheme must operate on the basis of one or more defaults, in particular, with a default minimum contribution rate and a default investment fund for those who do not express a choice. One of the reasons for expressing the legal arrangements in this way is that it allowed both the government and the scheme sponsor to avoid any legal liability for a poor pension outcome – unlike the case in which the contribution rate and investment fund were mandated by either the government or scheme sponsor. The legislation requires an employee to be auto-enrolled by the employer into a scheme selected by the employer. But it also allows the member to opt out of the scheme altogether or change the member contribution rate (so long as it does not fall below the minimum) or the investment fund. This places the legal responsibility for the pension outcome on the member.

investment strategy will depend on members' risk and returns preferences, which, in turn, will depend on the individual members' own preferences for risk and other personal attributes of the individual.² An implication is that the more heterogeneous are the membership's preferences, the more default funds are needed. The two extreme cases are one fund for each member and one-fund-for-all.³

In this paper, we consider the UK's largest private sector occupational pension scheme, the Universities Superannuation Scheme (USS), with a view to establishing the appropriate number of default funds required to reflect the pension scheme members' attitudes to risk. We make use of a survey of USS members carried out in September-October 2015 in preparation for the addition in October 2016 of a new DC section to the existing DB pension scheme. This survey had 9,755 respondents who provided information about their demographic profile, risk preferences and other characteristics. We apply cluster analysis to these survey responses to identify similar groups of members with relatively homogeneous preferences and characteristics (Everitt et al., 2011).

Our literature survey identifies member age as an important determinant of risk preferences, and we apply cluster analysis to different age cohorts: 30s, 40s and 50s (there were only small numbers of sample respondents in the other age cohorts). The cluster analysis identifies just two distinct groups of members in their 40s and 50s – the most important age cohorts in terms of the timing, size and compounding of returns on pension contributions. One group displays higher pay, longer tenure, less interest in ethical investing, lower risk capacity, a higher percentage of males, and a higher percentage of academics than the members of the other group – and significantly, *all* of its members have previously taken the active decision of making additional contributions (in the form of additional voluntary contributions (AVCs) or added years

² Not everyone agrees with this approach. For example, Bernstein (1992) calls it the 'interior decorator fallacy', namely that portfolios should reflect attitudes to risk in the same way that interior decorators attempt to reflect the personal taste of their clients. Such people would argue that if the aim is to achieve a target pension fund (or retirement income) with a specified degree of probability, then that is a technical exercise largely independent of a scheme member's risk preferences. Of course, this view is not incompatible with having a default fund investment strategy with the same aim.

³ It is, of course, arguable that there can only ever be a single default fund in any given scheme, given what the term default means. However, when scheme membership segments into mutually exclusive groupings, as for example when some members are interested in ethical investing, while the rest are not, then more than one default fund would be needed.

contributions), whereas *none* of the members of the other cluster have made additional contributions.

Further analysis indicated that the risk attitudes of the two groups were not significantly different, allowing us to conclude that a single lifestyle default fund is appropriate. Characteristics that other studies have found important determinants of risk attitudes, such as age, income and (pension) wealth, do not turn out to be as significant for USS members. Further, despite being on average more highly educated than the general population, USS members are marginally more risk averse than the general population, controlling for salary, although the difference is not significant.

The outline of the paper is as follows. Section 2 describes the USS pension scheme. Section 3 provides the theoretical background, based on two-fund separation, for the analysis of the default investment fund and reviews the existing literature on attitudes and personal characteristics. Section 4 explains cluster analysis which is the research methodology that we apply to identify similar groupings of scheme members. The empirical findings are examined in section 5, and Section 6 concludes. The survey questionnaire emailed to USS members is reproduced as an Appendix.

2. The Universities Superannuation Scheme and the survey questionnaire

The Universities Superannuation Scheme⁴ – which covers academic and professional services staff in UK universities – is one of the largest pension schemes in the UK, with 420,000 members, comprising 200,000 active, 69,000 deferred, and 151,000 pensioner members.⁵

On 31 March 2016, it closed its final salary section to future accrual and replaced this with a career average revalued earnings (CARE) section. The final salary section had already been closed to new members since 31 March 2011. This new defined benefit (DB) section, named USS Retirement Income Builder, has an annual accrual rate of

⁴ <https://www.uss.co.uk/>

⁵ USS Report and Accounts 2018.

1/75th, Consumer Price Index (CPI) uprating of the pension in payment,⁶ and a tax-free lump sum at retirement equal to three times the initial pension.

On 1 October 2016, a salary threshold was introduced, initially set at £55,000 per annum, above which member and employer contributions were paid into a new defined contribution (DC) section, named USS Investment Builder, and USS became a hybrid scheme with some members building up both DB and DC benefits. In 2016, contribution rates were set at 18% of salary for the employer and 8% of salary for members up to £55,000.⁷ Above £55,000, the employer contribution rate was 12% of the excess, while the member's contribution rate was 8% of the excess. Members at all salary levels could make AVCs into the DC section and initially this could be "matched" from the employer for additional contributions up to 1% of salary.⁸

To design the new DC section, USS undertook a programme of research in 2015 to understand member needs within the hybrid scheme. This included comparative studies of other DC schemes and prevailing pension industry best practice, demographic analysis of the USS membership to understand risk capacity, member outcome analysis based on stochastic modelling of possible investment strategies and member impacts and focus groups.⁹ As part of this programme, in October 2015, USS worked with A2Risk¹⁰ to design a risk attitudes survey of USS members. The primary purpose of the survey was to inform USS's understanding of risk attitude and investment beliefs in order to support the design of the USS Investment Builder investment fund range. USS required information on four aspects of financial planning from the survey: (1) personal circumstances,¹¹ (2) attitude to risk (ATR), (3) capacity

⁶ USS will match increases in CPI for the first 5% plus half of the difference above 5% up to a maximum increase of 10%. So, if CPI increased by 20%, the USS pension would increase would be 10%. (<https://www.uss.co.uk/members/members-home/retiring/pensions-in-payment>)

⁷ Annually uprated in line with inflation.

⁸ USS member presentation, July 2016. The 1% employer match was removed in April 2019.

⁹ A summary of this research can be found at <https://www.uss.co.uk/~media/document-libraries/uss/scheme/uss-investment-builder-a-summary-of-research.pdf?la=en>

¹⁰ <http://www.a2risk.com/>

¹¹ Information was collected on: institution, age, gender, annual salary, expected retirement age, years of membership of USS, whether the member's role was predominantly academic or professional services, whether AVCs were being made, whether the member could reasonably expect to live a long and healthy retirement, and whether the USS pension was likely to be the main household income in retirement.

to bear risk as measured by capacity for loss (CFL),¹² and (4) investment beliefs, including ethical considerations. An online questionnaire was distributed to members by participating employers in the scheme. A total of 9,755 responses were collected,¹³ making it one of the largest surveys of risk attitudes in the UK. Members were requested to answer 12 ATR questions and the results can be compared against a survey of the UK national population conducted by A2Risk via YouGov at around the same time.

As explained earlier, there was a legal requirement to put a suitable default investment fund in place for the start of the hybrid scheme in October 2016. USS reviewed the international evidence on the design of investment fund defaults. One example that it examined was the Australian pension fund, QSuper, which designed its default funds by segmenting its members according to age and size of the accumulated fund.¹⁴ QSuper has eight gender-neutral lifetime groups based on age and fund size. Members are first allocated to the group most suitable for them and then are automatically moved as they age: they are moved from higher risk to lower risk assets, consistent with an age-dependent investment strategy called lifestyling or lifecycling (Blake et al., 2014). They are also moved if their pension fund size changes sufficiently, through investment returns, contributions or transfers.¹⁵

USS's aims were to (1) design of the default lifestyle fund(s) so it (or they) were aligned with the objectives and preferences of the majority of USS members saving in USS Investment Builder, and (2) assess whether there are identifiable groups of members within the USS membership with heterogeneous objectives and preferences that may need to be actively supported towards an investment fund (default or self-select) that is better suited to meeting their long-term objectives.

¹² CFL is defined as the ability to sustain losses on an investment portfolio and this will be influenced by factors such as the number of dependants, existing financial commitments, etc.

¹³ Equal to 6.6% of active members. The sample was assessed as being broadly representative of the active membership of USS in terms of the age and salary distributions, the gender balance and the balance between academic and professional services staff.

¹⁴ <https://qsuper.qld.gov.au/our-products/investment-options/lifetime/>

¹⁵ The reason for this is that the first pillar state pension in Australia (the "age pension") is means tested, implying that a higher second pillar pension will reduce the age pension. The size-related default funds are designed to minimise reductions from the two systems.

3. Two-fund separation and factors influencing risk attitudes

3.1 Two-fund separation

Tobin (1958) demonstrated that the mean-variance model of portfolio selection framework of Markowitz (1952) leads to two-fund separation whereby all investors hold a combination of the same portfolio of risky assets and the risk-free asset. It is possible to use this framework to assess the optimality of the default fund asset allocation in a pension scheme. Figure 1 shows the location of a particular default fund on the capital market line which combines a portfolio of risky assets with the safe asset. In addition, we include the risk-return preferences of three pension scheme members (A, B, C). This default fund matches the optimal preferences of B, but the constraint of a single default means that scheme members A and C suffer a welfare loss. The risk-return combination for the default fund is too low for A and too high for C.

[Figure 1 around here]

In the extreme, the pension scheme could minimise these welfare losses by setting up a default fund for each member. However, this would be costly. Alternatively, if the scheme could segment the membership into a number of distinct groups based on similar risk-return characteristics, this would reduce the number of default funds required. If the entire scheme membership was sufficiently homogeneous, then just a single default fund would be sufficient. In reference to Figure 1, the empirical question is whether there are groups of scheme members with similar sets of preferences, and how close are these group preferences to each other. Our aim in this paper is to identify groups of sufficiently homogeneous pension scheme members to assess the minimum number of default funds required.

We explained in footnote 1 that the legislation in the UK requires DC pension funds to establish at least one default fund into which compulsory contributions are allocated when members do not record any choice. However, although members may not make explicit choices, they may have identifiable characteristics such as gender or age, which the pension scheme can utilise to make an appropriate decision on the scheme members' behalf. Individual risk-return preferences depend on a range of demographic (including gender and age), socio-economic, health and personality factors. We now summarise the evidence on these key factors.

3.2 Factors influencing risk attitudes

3.2.1 Gender

A large body of evidence suggests that, on average, women are more risk averse than men and this has consequences for financial decision making contexts, such as asset allocation, trading patterns and ethical choices (e.g., Bajtelsmit and Bernasek (1996), Powell and Ansic (1997), Jianakoplos and Bernasek (1998), Schubert et al. (1999), Finucane et al. (2000), Croson and Gneezy (2009) and Dohmen et al. (2011)). However, Nelson (2017) concludes that there was little evidence for gender differences, claiming instead that existing studies were contaminated with confirmation bias or gender stereotyping.

With respect to pension schemes, Bajtelsmit and VanDerhei (1997), Hinz et al. (1997) and Sundén and Surette (1998) report gender differences in participant-directed pension investments, with women selecting more conservative investments. Watson and McNaughton (2007) examined the impact of gender on the pension fund risk preferences of staff in the Australian university sector. They also find that women choose more conservative investment strategies than men and that, combined with lower contributions (as a result of lower salaries), explains why women have lower projected retirement benefits than men in Australian universities.

Overconfidence is another acknowledged difference between men and women. Lenny (1977), Meehan and Overton (1986), and Gervais and Odean (2001) find that men are generally more confident about their own abilities than women. Over-optimistic investors also tend to make poorer investment decisions (Hunt et al., 2015): Barber and Odean (2001) document that men transact their common stock investments 45% more frequently than women, and this excessive trading reduces men's net investment returns compared with women.

Ethical behaviour differs by gender, with women generally behaving more ethically than men (Dollar et al., 2001; Borkowski and Ugras, 1998)). Betz et al. (1989) use data from a sample of 213 business school students and find that men are more than twice

as likely as women to engage in actions regarded as unethical. Beams et al. (2003) finds the social stigma of trading on inside information was a more important deterrent for female respondents than male respondents. Often these risk aversion and trading studies are undertaken in relation to particular samples of participants, such as business school students, and the results may not be directly applicable to other distinct groups. Johnson and Powell (1994) compare the decision-making characteristics of males and females in “non-managerial” positions with those in “managerial” positions and find that for those in managerial positions of both genders display similar risk attitudes and make decisions of comparable quality. Atkinson et al. (2003) and Niessen and Ruenzi (2006) compare the performance of male and female mutual fund managers. Both studies find that male and female managed funds do not differ significantly in terms of performance, risk, and other fund characteristics.

Eckel and Grossman (2008) show that studies with contextual frames show less consistent differences in risk aversion between men and women. Perceptions are also important. Siegrist et al. (2002) show that both men and women overestimated male risk preferences, but accurately predicted female risk preferences, suggesting that predictions were influenced by knowledge about risk preferences incorporated in gender stereotypes.

3.2.2 Age

Life-style investment strategies, as frequently advocated by financial advisors, state that young people should invest in risky assets and shift gradually to safer assets as they age. This strategy has been criticised by Samuelson (1989a) on the grounds that, for a given degree of risk aversion, the optimal asset allocation should be independent of age (see also Poterba et al. (2006)). However, if it is the case that risk aversion does indeed decline with either age or the length of the financial planning horizon, then this provides a justification for life-styling (Samuelson (1989b) and Schooley and Worden (1999)).

Most studies investigating whether risk aversion changes with age show that very young people and very old people tend to be risk averse. Between these ages, risk aversion initially falls before rising again following a U-shaped pattern (e.g., Riley and

Chow (1992), Bakshi and Chen (1994), and Pålsson (1996)). Although contrary evidence is provided by Wang and Hanna (1996) who show that risk aversion falls with age controlling for other variables. Brooks et al (2018), while confirming that risk aversion falls with age (which they call the pure age effect), find evidence that falling risk aversion is associated with a reduced ability to bear losses and a declining investment horizon. There is also evidence of a specific cohort effect with different generations having different risk attitudes at the same age – possibly influenced by experience when young. Gilliam et al. (2010) find that leading baby boomers are less risk tolerant than trailing baby boomers.

Korniotis and Kumar (2011) identify two effects of age on financial investments. First, they find that older experienced investors make better investment decisions, because they follow rules of thumb that reflect greater investment knowledge. However, there is a second effect that investment skill deteriorates with age due to the adverse effects of cognitive aging. Older investors are less effective in applying their investment knowledge and exhibit worse investment skill, especially if they are less educated, earn lower income, and belong to ethnic minority groups. Overall, the adverse effects of aging appear to dominate the positive effects of experience. Dohmen et al. (2010) confirm that lower cognitive ability in otherwise healthy people is associated with greater risk aversion. Kim et al. (2016) conclude that older investors should delegate their investment decisions to experts.

3.2.3 Socio-economic, health and personality factors

Most studies find that risk aversion decreases with higher salary and wealth, controlling for other factors such as gender, age, education and financial knowledge (e.g., Riley and Chow (1992), Grable (2000), Hartog et al. (2002), Campbell (2006), Guiso and Paiella (2008), and Grinblatt et al. (2011)). However, individuals who are more likely to face salary uncertainty or to become liquidity constrained exhibit a higher degree of risk aversion (Guiso and Paiella (2008)). Similarly, individuals become more risk averse after a negative shock to wealth, such as a reduction in the value of their home (Paravisini et al. (2017)).

Individuals with higher levels of general educational attainment or higher IQs tend to be more risk tolerant (e.g., Grable (2000) and Grinblatt et al. (2011)). This is strongly reinforced if individuals also have a high degree of financial literacy (Behrman et al. (2012), and Lusardi and Mitchell (2014)). Financial literacy tends to be lower amongst the young, women, the less educated, and ethnic minorities (Lusardi and Mitchell (2011)). Individuals who score higher on the financial literacy questions are much more likely to plan for retirement. Financial planning can explain the differences in levels of retirement savings and why some people reach retirement with very little or no wealth (Lusardi and Mitchell (2007, 2011)). Bluethgen et al. (2008) find that financial advice can also help to overcome risk aversion, especially for women, and lead to more diversified portfolios that are better targeted to achieving an investor's goals.

Health is another factor that can influence risk aversion. A typical finding is that financial risk tolerance is positively associated with both health and life expectancy (Hammit et al., 2009). But particular diseases can change people's risk attitudes. For example, Tison and Hammit (2018), using data from the US Health and Retirement Study, find that people suffering from cancer and arthritis can become less risk averse, while people with diabetes can become more risk averse. Sinz et al. (2008) report that individuals with mild Alzheimer's disease (AD) gambled more often in situations with low-winning probabilities and less frequently in situations with high-winning probabilities than healthy participants in controlled experiments. Delazer et al. (2007) concluded that people with mild AD made such frequent changes between strategies that decisions were being made randomly. Smoski et al. (2008), in a controlled experiment, found that depressive participants would learn to avoid risky responses faster than control participants. They conclude that depressive individuals tend to have enhanced feedback-based decision-making abilities, but are more risk averse than non-depressive individuals.

Attitude to risk can be influenced by personality type. Psychologists distinguish between Type A personalities – who are categorised as being competitive, outgoing, ambitious, impatient and/or aggressive – and Type B personalities – who are more laid back. Type A individuals tend to take greater financial risks than Type B individuals according to a study by Carducci and Wong (1998). Another way of differentiating

between individuals is through the types of jobs they choose. Studies show that entrepreneurs are more risk tolerant than employees, private-sector employees are more risk tolerant than public-sector employees, and professionals are more risk tolerant than employees without a professional qualification (Grable (2000) and Hartog et al. (2002)).

The degree of risk aversion is also influenced by marital status. Sung and Hanna (1996) and Yao and Hanna (2005) show that single women are more risk averse than single men or married couples. Having children tends to increase risk aversion amongst both men and women according to Chaulk et al. (2003), Hallahan et al. (2004) and Gilliam et al. (2010).

None of the above factors can fully explain an individual's risk aversion. There are numerous other factors that influence risk attitude – typically given the name background risks – such as the weather, emotional factors, and the environment in which an individual lives (Hirshleifer and Shumway (2003), Kamstra et al. (2003), Guiso and Paiella (2008)). USS members typically have high levels of general educational attainment and many may also have a high degree of financial literacy. In terms of other identifiable factors, the USS questionnaire asked questions on salary, job type, and health and marital status.

4. Research methodology

We apply cluster analysis to the 9,755 responses to the USS/A2Risk member risk attitudes and investment beliefs survey conducted in September-October 2015. Cluster analysis is an exploratory data analysis technique used to identify patterns in a data sample (Everitt et al., 2011). Most cluster analysis methods use some type of distance measure, such as Euclidean distance, for determining the similarity or dissimilarity between observations. We also apply data transformations through factor analysis before applying the cluster analysis.

Cluster analysis has been used extensively in market research – to identify distinct homogeneous groups based on purchasing patterns – and we follow the best-practice approach outlined in Tuma et al. (2011) and Tuma and Decker (2013). Cluster

analysis has previously been applied to research questions in pensions. Speelman et al. (2013) undertake a cluster analysis of groups of savers in Australia, and report that gender differences dominate outcomes. Gough and Sozou (2005) identify the attitudes of UK consumers to pension savings, and analyse 540 respondents that had made inquiries about pensions, and identify 6 groups based on age, income and DB membership. Deetlefs et al. (2015) examine a sample of UniSuper members,¹⁶ and use cluster analysis to identify groups of similar members, and then use these clusters to predict the likelihood of these groups choosing default options and levels of engagement with the pension scheme.

There are two main methods of cluster analysis. Partition clustering and hierarchical clustering. A commonly used partition clustering method is “*k*-means cluster analysis”, where we specify in advance the number of clusters, *k*, and an iterative algorithm is used to determine which observation should be included in each group. Each observation in the sample is assigned to one of the *k* groups based on the closeness of the value of the observation to the mean value of the *k*th group. For each group, the group mean is computed, and an observation is reassigned to another group if it is closer to the other group’s mean. New group means are determined, and these steps continue until no observation changes groups.

Following Everitt et al. (2011, p. 114), the *k*-means partition method specifies in advance *k* groups, and then assigns observations to these groups by minimising the error sum-of-squares (*SSE*) between observations and their group mean

$$\min SSE = \sum_{m=1}^k \sum_{i=1}^{n_m} d_{mi,m}^2$$

where n_m is the number of members of the m^{th} group and $d_{mi,m}$ is the Euclidean distance between the i^{th} observation in the m^{th} group and the mean of m^{th} group.¹⁷

¹⁶ UniSuper is the one of Australia’s largest pension schemes with 460,000 members and is open to all employees in the higher education and research sectors (Dobrescu et al. (2017)).

¹⁷ In this case, the method simply minimises the sum across the *k* groups of the sum of squared differences between each observation in each group and the mean of that group.

An alternative to partition clustering is hierarchical clustering, which creates hierarchically related sets of clusters. Agglomerative hierarchical clustering methods start with each observation in the sample of N observations being in a separate group (N groups each of size 1). The closest two groups are combined (giving $N-1$ groups: one of size 2 and the rest of size 1), and this process continues until all observations belong to the same group. This process creates a hierarchy of clusters. The simplest hierarchical method is single-linkage, which computes the similarity between two groups as the similarity between the closest pair of observations in the two groups. In our analysis, to measure the closeness between groups, we apply Ward's clustering method (Ward (1963)) in which the criterion for joining groups is based on a within-cluster error sum-of-squares. Following Everitt et al (2011) let SSE be the total within-cluster error sum-of-squares, then Ward's method is to

$$\min SSE = \sum_{m=1}^k E_m$$

where $E_m = \sum_{i=1}^{n_m} \sum_{j=1}^p (x_{mi,j} - \bar{x}_{m,j})^2$ and where $\bar{x}_{m,j} = (1/n_m) \sum_{i=1}^{n_m} x_{mi,j}$, with $x_{mi,j}$ being the value of the j^{th} variable ($j = 1, \dots, p$) for the i^{th} observation ($i = 1, \dots, n_m$) in the m^{th} group ($m = 1, \dots, k$).

The objective of using cluster analysis in our case is two-fold: (1) to identify groups of individuals with similar risk attitudes and/or risk capacities; and (2) having done this, to examine whether these individuals exhibit particular demographic and personal characteristics. Based on the literature review, we will be able to answer the following questions for our data sample: (1) Does risk aversion vary by gender?, (2) Are women more likely to be interested in ethical investments?, (3) Does risk aversion vary by age?, (4) How does risk aversion vary with salary?, (5) How does risk aversion vary with job type (academic vs professional services)?, and (6) Are USS members more risk averse than members of the general public?

5. Empirical findings

5.1 Descriptive statistics

The data sample includes an anonymous code for each individual, and a series of demographic and personal characteristics self-reported by the survey respondent (Section A of the Appendix) including: age (within five-year bands), gender, marital status (including married, civil partnership, single, divorced, separated, widowed), annual salary (within £10,000 bands), expected retirement age, length of USS membership, job-type (academic or professional services), whether the member could reasonably expect to live a long and healthy retirement, and whether the USS pension is expected to be the main source of income in retirement.

[Table 1 about here]

Participants in the survey then answered a series of questions around a number of different themes establishing: (1) their previous additional contributions to USS (in terms of AVC contributions or buying additional years of service in the DB section); (2) their attitude to risk (Section B, with 14 questions¹⁸); (3) their capacity to bear risk (Section C, with 6 questions); (4) ethical beliefs in investing (Section D, with 6 questions, plus a question on attitude to Shariah-compliant funds, and a further 5 questions on the desirable properties of DC funds); and (5) intentions with respect to participating in the new DC section operated by USS (Section E, with 2 questions).

Table 1 provides some descriptive statistics on the sample of participants. Panels A and B show that the median respondent is a 47-year old married male academic who has been a member of USS for 7 years with a salary of £50,000. This person intends to retire at 65 and expects to have a long healthy retirement during which the USS pension will be the household's main source of income. Further, this person has not previously made additional contributions to USS via AVCs or additional years of service. Recall that the DB section continues to operate as the main pension scheme for individuals earning up to £55,000 (through USS Retirement Income Builder), so that many individuals completing the USS survey might not be expected to automatically participate in the new additional DC scheme (USS Investment Builder), unless they actively select the match.¹⁹

¹⁸ Twelve questions (numbered 12-23) were used to assess attitude to risk. Two additional questions (numbered 24 and 25) were used as validity checks.

¹⁹ While it was operating prior to April 2019.

5.2 Cluster analysis

Cluster analysis works most effectively when the number of observations and the number of variables is relatively small because the algorithms used compute many pairwise comparisons.²⁰ To reduce the size of the data matrix (number of participants by number of variables), we split the sample by age of the participant, and form groups of participants based on the age distribution. Panel C of Table 1 shows the age distribution of the sample of respondents, and because of the relatively small number of respondents in their 20s (426 observations) and above sixty years of age (741), we concentrate our analysis on the remaining 8,588 members. We form three cohorts of respondents in their 30s (2,412), 40s (3,041) and 50s (3,135), and apply cluster analysis to each age cohort separately.

To reduce the dimensionality of the problem even further, we note that each theme asks a range of questions, so, using factor analysis, we analyse the correlation matrix of responses to identify a smaller number of common factors. To illustrate, there are six questions in the sample inviting responses on the individual's capacity for bearing risk (CFL). The correlation matrix in this case indicated that the responses were so highly correlated across respondents that we could reduce the potential number of questions to two (Q26 and Q27). A similar analysis suggested that there were just two potential questions in the case of ethical investment beliefs (Q32 and Q33) – revealing that, at most, two factors could explain most of the responses in these two themes. Table 2 shows the results of the factor analysis for the two themes. In both cases, there is a negative value on the second factor, indicating that just a single factor can explain the responses to the two sets of questions – and indeed the full set of questions for both themes.

[Table 2 about here]

A number of questions on the themes of DC investment intentions had ambiguous responses, and so we only retained one DC investment intentions question (being Q39 on the intention to “match” the additional employer contribution). The responses to the questions on the desirable properties of DC investments were poorly answered in the

²⁰ Cluster analysis is subject to the “curse of dimensionality”.

survey, with many participants not providing any answers. Some of the responses were also ambiguous and therefore this set of questions was dropped from the subsequent analysis.

Finally, there was the set of responses around the theme of attitude to risk. Each response was aligned on a 1-to-5 point Likert-scale to represent a risk attitude and then we averaged these responses across the 12 attitude to risk questions to provide an average risk aversion score (av_ATRQ). In Figure 2, we plot the distribution of av_ATRQ by age, with higher values representing greater risk aversion. While the distributions look similar, Figure 3 plots the average value of av_ATRQ within each of the eight 5-year age groups in our data set (age range 25-29 centred on age 27, up to age range 60-64 centred on age 62). This shows the broadly U-shaped pattern previously identified in the literature. A Bartlett test for equal variances rejects the hypothesis that the distributions in Figure 2 are the same ($\chi^2(7) = 27.38$). However, pair-wise tests of the difference in means of adjacent distributions indicates that only the 35-39 and the 40-44 age groups have statistically significantly different means. We can conclude that, while there is a U-shaped distribution in Figure 3, the differences in the average ATRQ scores across age groups are economically small.

[Figure 2 and 3 around here]

How do these USS member ATRQ scores compare with the UK adult population as a whole? A2Risk conducted a YouGov survey of risk attitudes of the UK adult population at around the same time as the USS survey. Average earnings for USS members were £38,000 and over 90% of respondents reported an income above £30,000. Since at the time, average UK earnings were around £26,000, it is clear USS members have above-average salaries compared with the UK adult population. USS members are on average less risk averse than the UK adult population (an av_ATRQ of 3.41 compared with 3.56). However, when we compare the USS risk aversion scores with those of the UK adult population with an income above £30,000, USS members are marginally more risk averse (3.41 compared with 3.34).²¹ We find a higher percentage of USS members who are labelled “cautious” by A2Risk: 16% of USS members compared to 10% of UK adults earning £30,000 or more. Slightly fewer USS members are

²¹ The standard deviation of the scores of both groups is very similar (both are close to 0.7), indicating that the difference between the scores is not statistically significant.

“moderately adventurous” or “balanced”. Male respondents in the USS survey tended, on average, to be less risk averse than female respondents, and this finding is consistent with the UK population when controlled for age and salary. For both USS and UK samples, ATRQ scores tend to be correlated with income but do not vary much by age.²²

On the basis of the factor analysis and the average risk aversion scores, we now have, for each individual in the sample, one or more estimated values for the responses to each of the four sets of themed questions: (1) attitudes to risk (av_ATRQ); (2) interest in ethical investing (a single factor); (3) risk-bearing capacity (a single factor); and (4) DC investment intentions (the match). In addition, we also know from Question 9 in the Appendix whether the respondent has made additional contributions to USS in the past through AVC contributions or the purchase of additional years in the DB section. We denote these five variables “investment characteristics” to differentiate them from variables, such as gender, age and salary etc, which we denote as “personal characteristics”. We now turn to the cluster analysis results.

5.3 Findings

We wish to identify whether there are patterns or clusters in these factors across the individuals in the sample. We are particularly interested in applying cluster analysis to the three age cohorts 30s, 40s and 50s across our five standardised variables:²³ (1) av_ATRQ (with a higher value denoting greater risk aversion); (2) a single ethics factor (with a higher value denoting a greater interest in ethical investing); (3) a single risk capacity factor (with a higher value denoting lower risk-bearing capacity); (4) match intentions (with a higher value indicating a stronger intention to match the employer contribution); and (5) additional contributions (which is a dummy variable taking the value unity if previous additional USS contributions have been made, and zero otherwise).

²² We show later that these correlations are not statistically significant (see Table 10).

²³ Because cluster analysis minimises a weighted sum of error-sum-of-squares, the results will be influenced by the size of a particular variable. Therefore, each of the five variables of interest is standardised to have zero mean and unit variance.

There are two methods for judging the appropriate number of clusters in a dataset: the Calinski-Harabasz pseudo- F statistic and visual inspection of a dendrogram. Calinski-Harabasz (1974) computes a pseudo- F statistic based on the ratio of the (between-clusters sum-of-squares)/($k-1$) and the (within-cluster sum-of-squares)/($N-k$), where k is number of clusters and N is number of observations. The appropriate number of clusters is where the Calinski-Harabasz statistic is maximised. This criterion can be used for both k -means partitions and for hierarchical approaches. The second method, relevant for hierarchical approaches only, is visual inspection of a dendrogram.

Figure 4 reports the dendrogram from applying Ward's method to the 50s age-cohort, and suggests that, across the five standardised variables, there are just two clusters in this sub-sample of the dataset. The vertical axis shows how the L2squared dissimilarity measure²⁴ between groups increases as more members are added to existing groups. A large jump in the dissimilarity measure suggests a cut-off for the number of clusters – at two in this case.

[Figure 4 and Tables 3 & 4 around here]

Table 3 reports the Calinski-Harabasz pseudo- F statistics for both Ward's hierarchical and the k -means partition methods for each age cohort. The F -statistic takes its highest value for groupings of two clusters in all age cohorts (30s, 40s and 50s).

In addition, we examine the observations identified by the clusters from both the hierarchical and k -means partitions to assess whether the two methods classify the observations into the same two sets of clusters. The results of these cross-tabulations are reported in Table 4. Panel A shows the cross-tabulations of the observations in the 30s age-cohort of the two clusters (Group) formed by both Ward's hierarchical method and the k -means partition (Clusters). So, for example, there are 833 observations that are in the first cluster defined by Ward's method and also in the first cluster defined by the k -means partition. However, there are 243 observations that are in the first cluster defined by Ward's method but happen to be in the second cluster defined by the k -means partition method. The implication from Panel A, is that, for the 30s age-cohort, the two clustering methods produce different groupings, from which we conclude that clear and robust clusters do not exist for this age cohort. But this is

²⁴ The Stata name for the minimised squared Euclidean distance between groups.

not a severe problem, given the low numbers of USS members in their 30s with salaries above £55,000 and hence eligible for USS Investment Builder.

Turning to the other two panels for the 40s and 50s cohorts, the two methods produce very similar groupings for the 40s cohorts (Panel B), and identical groupings for the 50s cohort (Panel C). We can therefore be very confident that the clusters formed for the 40s and 50s cohorts are robust to the clustering method used.

We also examine the distribution of the demographic and personal characteristics of individuals allocated to each of the two groups. The results, reported in Table 5, illustrate the distribution of variables across members of the two sets of clusters for the 40s age-cohort and for the 50s age-cohort.

[Table 5 around here]

Examining the numbers for the 40s age cohort first of all, it can be seen that there are large differences between the two clusters, with Cluster 2 displaying higher pay, longer tenure, additional contributions, less interest in ethical investing, lower risk capacity, a higher percentage of males, and a higher percentage of academics than the members of Cluster 1.²⁵ The additional contributions (in the form of AVC or added years contributions) is particularly noteworthy, since *all* members of the second cluster have made these contributions, but, in contrast, *none* of the members of the first cluster have made additional contributions. A multivariate analysis-of-variance test indicates that the differences in these variables between the two groups (e.g., differences in pay) are in aggregate statistically significant, indicating that the two clusters are statistically significantly different. There are, however, only small differences between the two clusters in terms of the degree of risk aversion and the propensity to match employer contributions.

Turning to the results for the 50s age-cohort, there are similar differences between the two clusters for most of the variables, with the exception that ethical investment beliefs are now similar across the two groups. As with the 40s age cohort, *all* the members of

²⁵ We might normally expect that a cohort with higher pay would have higher risk capacity, but in this case the higher concentration of males (who might be the main source of retirement income in the family) and longer tenure (implying greater reliance on the USS pension as the main and possibly only occupational pension) mean that this higher-pay cohort has lower risk capacity. Further, the guaranteed DB pension will provide a lower percentage of the total USS pension for this cohort, reinforcing the lower risk capacity.

the second cluster have previously made additional contributions with USS, whereas *none* of the members of the first cohort have. Again, a multivariate analysis-of-variance suggests that the differences between the two groups are in aggregate statistically significant.

[Tables 6 & 7 around here]

This leads to an interesting and potentially significant conjecture. Although our data set is cross-sectional rather than longitudinal, it might be possible to treat the two 50s age-cohort groups as being the same two 40s age-cohort groups ten years on, although they have grown marginally less interested in ethical investing as they have aged. We can investigate this by combining the 40s and 50s age cohorts. Table 6 confirms that there are still two clusters in the combined age cohorts, while Table 7 confirms that the two clustering methods produce identical clusters.

Table 8 reports the distribution of the demographic and personal characteristics of individuals in their 40s and 50s allocated to each of the two clusters. We observe the same large differences between the two clusters previously observed in Table 5. But the important point is that the two clusters for the combined age cohorts are identical to the two clusters found when the two age cohorts were analysed separately, with the second of the two clusters, Cluster 2, displaying higher pay, longer tenure, higher additional contributions, less interest in ethical investing, lower risk capacity, a higher percentage of males, and a higher percentage of academics than the members of the first cluster, Cluster 1. A multivariate analysis-of-variance test indicates that the two clusters are statistically significantly different. As before, there are only small differences between the two clusters in terms of the degree of risk aversion and the propensity to match employer contributions. But the most important point to emerge from combining the two age cohorts and comparing the results with the two age cohorts separately is that again *all* the members of one cluster (Cluster 2) have previously made additional contributions with USS, whereas *none* of the members of the other cluster (Cluster 1) has previously made additional contributions.

[Tables 8 & 9 around here]

Table 9 presents estimates of a probit model of the characteristics for the combined 40s and 50s age cohort clusters. A higher or more positive value of an estimated coefficient indicates a higher probability of the member being in Cluster 2, while a

lower or more negative value indicates a higher probability of the member being in Cluster 1. So, for example, higher pay increases the probability of the member being in Cluster 2, while a higher expected retirement age increases the probability of the member being in Cluster 1.

In Tables 7 and 10, the fact that the additional contributions variable is such a striking indicator of which cluster a member belongs raises the possibility that these results are driven solely by this particular characteristic. To investigate this, we dropped the `Additional_contributions` variable and performed the cluster analysis using only the other four investment characteristics plus the personal characteristics. For the 40s cohort, the two alternative cluster methods (partition vs hierarchical) indicate two clusters as before, but each method produces a cluster that is both different from each other and different from the previous clusters. For the 50s cohort, the partition method produces two clusters, while the hierarchical method produces three clusters. So while it is impossible to say that that two clusters in Tables 7 and 10 depend only on the additional contributions variable, it would appear that the additional contributions variable has a sufficiently powerful impact that the clusters are nowhere near as strongly defined when this variable is dropped.

Finally, we conducted a cluster analysis of the average risk aversion question scores (`av_ATRQ`) alone, using both clustering methods. Figure 5 presents a histogram of the distribution of the average scores across the 9,755 individuals in the sample. Recall that each individual has `av_ATRQ` based on their responses to the 12 ATRQs. Each question is based on a Likert score between 1-5, and so the average Likert score for each individual also has this same range. Higher values indicate greater risk aversion, and the histogram clearly shows a bunching or clustering of scores. The dendrogram for the single-linkage agglomerative hierarchical clustering method suggested 18 clusters in total (with 17 in one hierarchy). The partition method also identifies 18 clusters, although the number of members in each cluster differed from the hierarchical clustering method. To assess whether these differences in cluster membership were significant, we report in Figure 6 the relationship between ATRQ scores and the intention to match across 18 clusters for both the hierarchical method (panel A) and the partition method (panel B). Both panels show very similar patterns of responses, namely the lower the risk aversion, the higher the intention of the

member to match the additional employer contribution. We conclude from this that the two clustering methods produce sufficiently close clusters.

[Figure 5 around here]

Given this, we estimated a regression model of the attitude to risk scores for each of the two clustering methods with the following potential explanatory variables: age, pay, expected retirement age, tenure, pension wealth, %female, %couple, plus the match and additional contributions factors. Table 10 shows that only %female and %couple are statistically significant for both clustering methods. For both methods, a 1% increase in females in a cluster increases av_ATRQ by 0.05, while a 1% increase in couples in a cluster reduces the av_ATRQ by a little over 0.02. The first result reconfirms one of the key findings of the study, while the second supports the idea that couples have lower risk aversion than singles because of risk sharing within the household. An examination of Table 8 shows how these findings influence the two clusters for the combined 40s and 50s age cohorts. The two clusters have av_ATRQs of 3.41 and 3.30, respectively. This difference is explained almost entirely by the higher percentage of females in the first cluster (46.4%) compared with the second (39.1%), since the percentage of couples in the two clusters is broadly similar at 73%. Other variables, such as pay, do not have a statistically significant impact on the av_ATRQ. Even the match is not statistically significant, despite Figure 6.

[Figure 6 and Table 10 here]

6. Conclusions

Our aim in this study was to determine the number of default funds appropriate for a large occupational pension scheme with a defined contribution segment where the investment risk – and hence the uncertainty concerning the pension outcome – is borne by the scheme members. We examined a survey of member characteristics and risk attitudes applying cluster analysis (both hierarchical and partitioning) to segment the members into a small number of distinct groups or clusters. We tested whether these clusters are sufficiently distinctive to justify more than one default investment fund.

For USS pension scheme members, we were able to identify two distinct clusters in the 40s and 50s age cohorts – the most important age cohorts in terms of the timing, size and compounding of returns on pension contributions:²⁶ The first identified cohort included members with lower average pay, shorter average tenure, more interest in ethical investing, higher risk capacity, a higher percentage of females, and a higher percentage of professional services staff. This cluster had not previously made additional contributions with USS (in terms of previous AVCs or added years contributions). A second identifiable cohort contained members with higher average pay, longer average tenure, less interest in ethical investing, lower risk capacity, a higher percentage of males, and a higher percentage of academics. This cluster had previously made additional contributions with USS (in terms of previous AVC or added years contributions).

There were only small (and statistically insignificant) differences between the two clusters in terms of the average degree of risk aversion and the propensity to match employer contributions: the first cluster was marginally more risk averse and less likely to match than the second cluster. Conditioning only on the attitude to risk responses, we identified 18 clusters, with similar but not identical membership, depending on which clustering method is used. The differences in risk aversion across the 18 clusters could be explained largely by differences in the percentage of females and the percentage of couples. Risk aversion increases as the percentage of females in the cluster increases because they typically are more risk averse than males, while it reduces as the percentage of couples increases because of greater risk sharing within the household.

The survey also showed that, despite being on average more highly educated than the general population, USS members as a whole are marginally more risk averse than the general population, controlling for salary, although the difference is not significant.

²⁶ Most members of the 20s and 30s age cohorts have salaries below the £55,000 threshold for participation in the DC segment of USS, while for members of USS above age 60, although they will be amongst the highest salary earners in USS and therefore paying the highest level of contributions, their shorter remaining time as active members reduces any compounded benefits from asset returns.

The DB underpin in USS gives members a much stronger capacity for taking greater risk with their DC savings than might otherwise be the case. Further, the similarity in risk aversion scores across both clusters in their 40s and 50s suggests that a single default fund might be suitable, so long as it reflects the genuine risk tolerance – which takes account of both the risk attitude and risk capacity – of the USS membership. USS members can be characterised as having an overall risk tolerance which is broadly similar to that of the national population with salaries above £30,000, since their slightly greater risk aversion is offset by greater risk-bearing capacity due to the DB underpin.

On the other hand, the survey responses indicated that in some other dimensions apart from risk attitude, the scheme members segmented into mutually exclusive groups. Two prominent examples are whether members were (1) interested in ethical investing or not, or (2) required a Shariah-compliant fund or not.

In short, there is no evidence of a requirement for multiple defaults within the current scheme structure, which simplifies matters considerably. USS decided on the basis of this research to introduce a single default lifestyle fund which would de-risk gradually in the 10 years prior to retirement.²⁷ It also introduced an ethical default lifestyle fund, although a member making no choice would be allocated to the standard lifestyle fund. USS also offered 10 other (non-lifestyle) funds for self-selectors, including a Shariah-compliant fund.²⁸

However, the low level of heterogeneity in risk tolerance across the membership suggests that it might be acceptable to offer just three funds in addition to the two default funds (rather than the current 10) to satisfy the diversity of risk attitudes: (1) a well-diversified fund with a higher level of risk than the default fund, (2) a well-diversified fund with a lower level of risk than the default fund, and (3) a Shariah-compliant fund. However, we acknowledge that self-select funds may be in place to

²⁷ This therefore has similarities with Australia's QSuper approach in having a lifestyle fund that de-risks over 10 years. However, USS decided not to have multiple age-related default funds, unlike QSuper.

²⁸ The full list of funds is available at <https://www.uss.co.uk/how-uss-invests/the-fund/investments/uss-equity-holdings>

meet the particular requirements of a small minority of members who would like more control over investment regardless of their risk preferences.

The appropriate communication and engagement strategy follows on naturally from our empirical findings. This involves informing all members at joining about the default lifestyle fund in place for those who are not interested in engaging with their scheme – as well as the lifestyle ethical fund. Self-selectors, by contrast, need to be warned about both reckless conservatism and reckless adventurism and subsequently need to be guided or nudged at key ages (e.g., 30, 40, 50 and 60) into adjusting the risk exposure of their pension fund in order to maximise their lifetime welfare.

Particular effort should therefore go into designing a suitable engagement programme for those members who have not previously made additional contributions with USS either voluntarily or as part of the match.²⁹ USS should be particularly concerned about self-selectors who had never engaged with the scheme as younger members. When it comes to the appropriate time to begin de-risking, they are unlikely to be motivated to do so without suitable USS information and guidance.

Finally, it is important to recognise that our findings might not generalise to other DC pension schemes for at least two reasons. First, most DC pension schemes do not have a DB underpin which allows certain groups of scheme members to take more investment risk than would otherwise be advisable. Second, USS members are more highly educated than most pension scheme members and this may have an influence, although we suspect, on the basis of the survey responses, that the influence might be negligible.

²⁹ In the case of USS, the sponsors ended matching in April 2019. A new replacement question in any future survey would therefore be needed.

References

- Atkinson, S.M., Baird, S.B., and Frye, M.B. (2003). Do female mutual fund managers manage differently? *Journal of Financial Research* 26(1): 1-18.
- Bajtelsmit, V. L., and Bernasek, A. (1996). Why do women invest differently than men?. *Financial Counseling and Planning* 7: 1-10
- Bajtelsmit, V. L. and VanDerhei, J. L. (1997). Risk aversion and pension investment choices. In M. S. Gordon, O. S. Mitchell, & M. M. Twinney (Eds.), *Positioning Pensions for the Twenty-First Century* (pp. 45–66). Philadelphia: University of Pennsylvania Press.
- Bakshi, G. S., and Chen, Z. (1994). Baby boom, population aging, and capital markets. *Journal of Business* 67(2): 165-202.
- Barber, B. M., and Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics* 116(1): 261-292.
- Beams, J. D., Brown, R. M., and Killough, L. N. (2003). An experiment testing the determinants of non-compliance with insider trading laws. *Journal of Business Ethics* 45(4): 309-323.
- Behrman, J. R., Mitchell, O. S., Soo, C. K. and Bravo, D. (2012). How financial literacy affects household wealth accumulation. *American Economic Review: Papers and Proceedings* 102(3): 300-304.
- Bernstein, P. (1992) *Capital Ideas*, Free Press, New York.
- Betz, M., O'Connell, L., and Shepard, J.M. (1989). Gender differences in proclivity for unethical behavior. *Journal of Business Ethics* 8(5): 321-324.
- Blake, D., Wright, D., and Zhang, Y. (2014). Age-dependent investing: Optimal funding and investment strategies in defined contribution pension plans when members are rational life cycle financial planners. *Journal of Economic Dynamics and Control* 38(1): 105-124.
- Bluethgen, R., Gintschel, A., Hackethal, A., and Müller, A. (2008). *Financial advice and individual investors' portfolios*. Working paper. Oestrich-Winkel: European Business School.
- Borkowski, S.C., and Ugras, Y.J. (1998). Business students and ethics: A meta-analysis, *Journal of Business Ethics* 7(11): 1117-1127.
- Brooks, C., Sangiorgi, I., Hillenbrand, C., and Money, K. (2018). Why are older investors less willing to take financial risks? *International Review of Financial Analysis* 56: 52-72.
- Campbell, J. Y. (2006). Household finance. *Journal of Finance* 61(4): 1553-1604.

- Carducci, B. J., and Wong, A. S. (1998). Type A and risk taking in everyday money matters. *Journal of Business and Psychology* 12(3): 355–359.
- Chaulk, B., Johnson, P. J., and Bulcroft, R. (2003). Effects of marriage and children on financial risk tolerance: A synthesis of family development and prospect theory. *Journal of Family and Economic Issues* 24(3): 257-279.
- Croson, R., and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature* 47(2): 448-474.
- Deetlefs, J., Bateman, H., Dobrescu, L., Newell, B., Ortmann A., and Thorp.S. (2015). *Suspicious minds (can be a good thing when saving for retirement)*. Working paper. Sydney: University of New South Wales.
- Delazer, M., Sinz, H., Zamarian. L., and Benke, T. (2007). Decision-making with explicit and stable rules in mild Alzheimer's disease. *Neuropsychologia* 45(8): 1632-1641.
- Dobrescu, L., Fan, X., Bateman, H., Newell, B., Ortmann, A., and Thorp. S. (2017) Retirement savings: A tale of decisions and defaults. *The Economic Journal* 128 (610): 1047-1094.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review* 100(3): 1238-1260.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* 9(3): 522-550.
- Dollar, D., Fisman, R., and Gatti, R. (2001). Are women really the 'fairer' sex? Corruption and women in government. *Journal of Economic Behavior and Organization* 46(4): 423-429.
- Eckel, C. C., and Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of Experimental Economics Results*, 1:1061-1073.
- Everitt, B. S., Landau, S., Leese, M., and Stahl, D. (2011). *Cluster Analysis*. 5th edition, Chichester, UK: Wiley.
- Finucane, M. L., Slovic, P., Mertz, C.K., Flynn, J., and Satterfield, T.A. (2000). Gender, race, and perceived risk: The 'white male' effect. *Health, Risk and Society* 2:159-172.
- Gervais, S., and Odean, T. (2001). Learning to be overconfident. *Review of Financial Studies* 14(1): 1-27.
- Gilliam, J.E., Chatterjee, S., and Zhu, D. (2010). Determinants of risk tolerance in the baby boomer cohort. *Journal of Business and Economic Research* 8(5): 79-87.

- Gough, O., and P. D. Sozou. 2005. Pensions and retirement savings: Cluster analysis of consumer behaviour and attitudes. *International Journal of Bank Marketing* 23: 558-70.
- Grable, J. E. (2000). Financial risk tolerance and additional factors that affect risk taking in everyday money matters. *Journal of Business and Psychology* 14(4): 625-630.
- Grinblatt, M., Keloharju, M., and Linnainmaa, J. (2011). IQ and stock market participation. *Journal of Finance* 66(6): 2121-2164.
- Guiso, L., and Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association* 6(6): 1109-1150.
- 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.
- Hammitt, J.K., Haninger, K., and Treich, N. (2009). Effects of health and longevity on financial risk tolerance. *The Geneva Risk and Insurance Review* 34(2): 117-139.
- Hartog, J., Ferrer-i-Carbonell, A., and Jonker, N. (2002). Linking measured risk aversion to individual characteristics. *Kyklos* 55(1): 3-26.
- Hinz, R. P., McCarthy, D. D., and Turner, J. A. (1997). Are women conservative investors? Gender differences in participant-directed pension investments. In M. S. Gordon, O. S. Mitchell, & M. M. Twinney (Eds.), *Positioning pensions for the twenty-first century* (pp. 91–103). Philadelphia: University of Pennsylvania Press.
- Hirshleifer, D., and Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *Journal of Finance* 58(3): 1009-1032.
- Hunt, S., Stewart, N., and Zaliauskas, R. (2015). *Two plus two makes five? Survey evidence that investors overvalue structured deposits*. Occasional Paper No. 9. London: Financial Conduct Authority.
- Jianakoplos, N. A., and Bernasek, A. (1998). Are women more risk averse? *Economic Inquiry* 36(4): 620-630.
- Johnson, J. E.V., and Powell, P. L. (1994). Decision making, risk and gender: Are managers different? *British Journal of Management* 5(2): 123-138.
- Kamstra, M.J., Kramer, L.A., and Levi, M.D. (2003). Winter blues: A SAD stock market cycle. *American Economic Review* 93(1): 324-43.
- Kim, H.H., Maurer, R., and Mitchell, O. S. (2016). Time is money: Rational life cycle inertia and the delegation of investment management. *Journal of Financial Economics* 121(2): 427–447.
- Korniotis, G. M., and Kumar, A. (2011). Do older investors make better investment decisions? *Review of Economics and Statistics* 93(1): 244-265.

- Lenny, E. (1977). Women's self-confidence in achievement settings. *Psychological Bulletin* 84(1):1-13.
- Lusardi, A., and Mitchell, O. S. (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics* 54(1): 205–224.
- Lusardi, A., and Mitchell, O. S. (2011). Financial literacy around the world: An overview. *Journal of Pension Economics and Finance* 10(4): 497-508.
- Lusardi, A., and Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature* 52(1): 5-44.
- Malmendier, U., and Nagel S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics* 126(1): 373-416.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance* 7: 77–91.
- Meehan, A.M., and Overton, W.F. (1986). Gender differences in expectancies for success and performance on Piagetian spatial tasks. *Merrill-Palmer Quarterly* 32(4): 427-441.
- Nelson, J. (2017). *Gender and Risk-Taking: Economics, Evidence and Why the Answer Matters*. London: Routledge.
- Niessen, A., and Ruenzi, S. (2006). *Sex matters: Gender and mutual funds*. Working paper. Cologne: University of Cologne and Centre for Financial Research Cologne.
- OECD (2009) *Guidelines for Pension Fund Governance*, Organisation for Economic Cooperation and Development, Paris, 5 June; <http://www.oecd.org/pensions/private-pensions/34799965.pdf>
- Pålsson, A.-M. (1996). Does the degree of relative risk aversion vary with household characteristics? *Journal of Economic Psychology* 17(6): 771-787.
- Paravisini, D. Rappoport, V., and Ravina, E. (2017). Risk aversion and wealth: Evidence from person-to-person lending portfolios. *Management Science* 63(2): 279-585.
- Pensions Commission (2005) *A New Pension Settlement for the Twenty-First Century: The Second Report of the Pensions Commission*, The Stationery Office; <https://webarchive.nationalarchives.gov.uk/+http://www.dwp.gov.uk/publications/dwp/2005/pensionscommreport/main-report.pdf>
- Poterba, J., Rauh, J., and Venti, S., and Wise, D. (2006). *Lifecycle asset allocation strategies and the distribution of 401(k) retirement wealth*. Working Paper No. 11974. Cambridge, MA: NBER.

Powell, M., and Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology* 18(6): 605-628.

Punj, G., and Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestions for application. *Journal of Marketing Research* 20(2): 134-148.

Riley, W.B., and Chow, K.V. (1992). Asset allocation and individual risk aversion. *Financial Analysts Journal* 48(6): 32-37.

Samuelson, P. A. (1989a). The judgment of economic science on rational portfolio management: Indexing, timing, and long-horizon effects. *Journal of Portfolio Management* 16(1): 4-12.

Samuelson, P.A. (1989b). A case at last for age-phased reduction in equity. *Proceedings of the National Academy of Sciences* 86(22), 9048–9051.

Schooley, D.K., and Worden, D. D. (1999). Investors' asset allocations versus life-cycle funds. *Financial Analysts Journal* 55(5): 37-43.

Schubert, R., Brown, M., Gysler, M., and Brachinger, H. W. (1999). Financial decision-making: Are women really more risk-averse? *American Economic Review* 89(2): 381-385.

Siegrist, M., Cvetkovich, G., and Gutscher, H. (2002). Risk preference predictions and gender stereotypes. *Organizational Behavior and Human Decision Processes* 87(1): 91-102.

Sinz, H, Zamarian, L., Benke, T., Wenning, G.K., and Delazer, M. (2008) Impact of ambiguity and risk on decision making in mild Alzheimer's disease. *Neuropsychologia* 46(7):2043-2055.

Smoski, M.J, Lynch, T.R, Zachary Rosenthal M, Cheavens, J.S, Chapman, A.L, and Krishnan R.R. (2008). Decision-making and risk aversion among depressive adults. *Journal of Behavior Therapy and Experimental Psychiatry* 39(4): 567-576.

Speelman, C. P., Clark-Murphy, M., and Gerrans, P. (2013). Decision making clusters in retirement savings: Gender differences dominate. *Journal of Family and Economic Issues* 34: 329-39.

Sundén, A. E., and Surette, B. J. (1998). Gender differences in the allocation of assets in retirement savings plans. *American Economic Review* 88(2): 207-211.

Sung, J., and Hanna, S.D. (1996) Factors related to risk tolerance. *Journal of Financial Counseling and Planning* 7: 11-20.

TPR (2016) *Code of practice no. 13: Governance and administration of occupational trust-based schemes providing money purchase benefits*, The Pensions Regulator, July.

TPR (2019) *Defined contribution trust-based pension schemes research: Report of findings on the 2019 survey*, prepared for The Pensions Regulator by OMB Research, May.

Tison, A., and Hammitt, J. K. (2018) *How do health shocks influence financial risk aversion? A theoretical and empirical investigation*. Working paper. Marseille: Aix Marseille University.

Tobin, J. (1958). Liquidity preference as behaviour towards risk. *Review of Economic Studies* 25, 65–86.

Tuma, M., and Decker, R. (2013). Finite mixture models in market segmentation: A review and suggestions for best practices. *Electronic Journal of Business Research Methods* 11: 2-15.

Tuma, M. N., Decker, R., and Scholz, S. (2011). A survey of the challenges and pitfalls of cluster analysis application in market segmentation. *International Journal of Market Research* 53.

Wang, H., and Hanna, S.D. (1996). Does risk tolerance decrease with age? *Journal of Financial Counseling and Planning* 8: 27-32.

Ward, J.H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58: 236–244.

Watson, J., and McNaughton, M. (2007). Gender differences in risk aversion and expected retirement benefits. *Financial Analysts Journal* 63(4): 52-62.

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.

Figure 1: Potential welfare losses from a single default fund with heterogeneous scheme member preferences

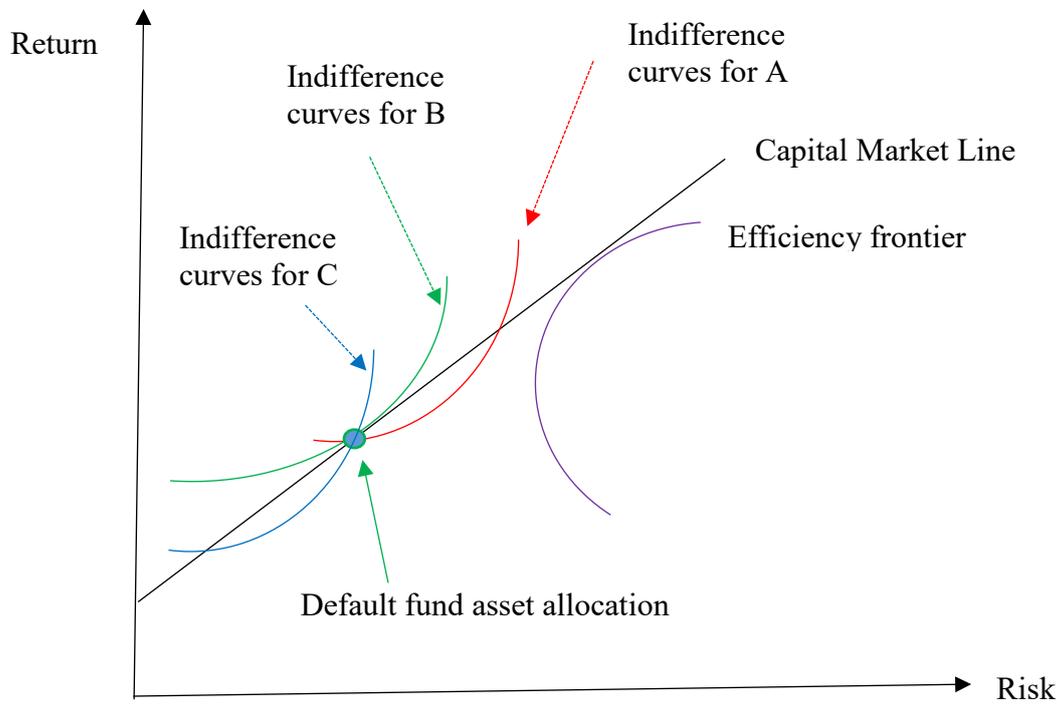
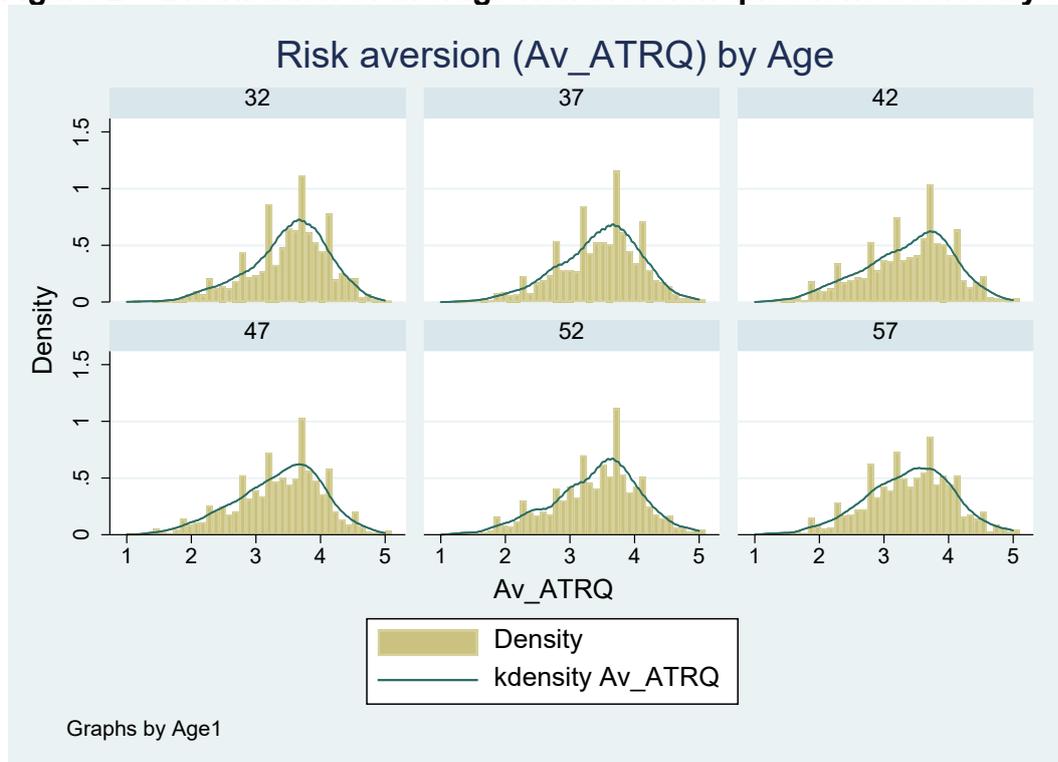


Figure 2 – Distribution of average risk aversion questions scores by age



Note: The figure shows the distribution of the attitude to risk questions score for selected ages, both in the form of a histogram and a kernel density

Figure 3 – The average risk aversion questions score by age

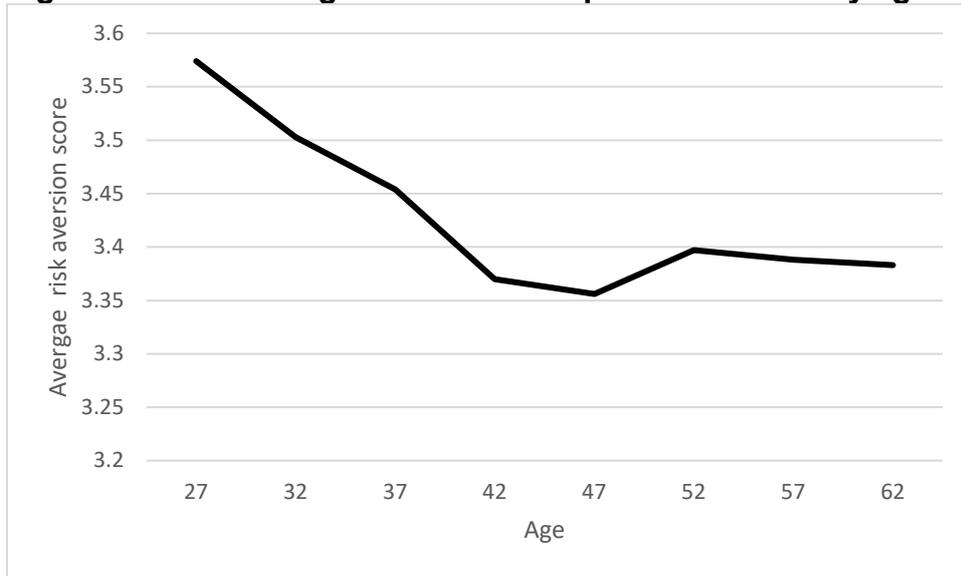
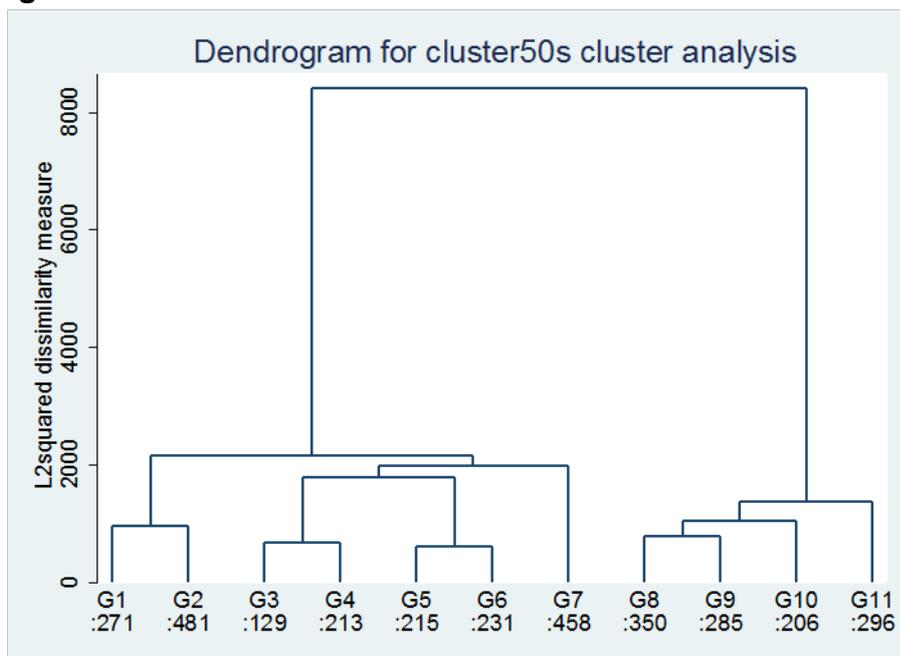


Figure 4 – Dendrogram from Ward's hierarchical clustering method for the 50s age cohort



Note: The dendrogram only reports groups with cut-off value of the L2squared dissimilarity measure > 500. There are 11 groups with cut-off > 500, and the numbers in each group are show below each group (e.g., 271 members in G1). There are many more groupings with cut-off < 500, until on the bottom row (not shown), there will be 3,135 groups with each member being in their own group, and therefore a dissimilarity measure of zero.

Figure 5 – Distribution of the average risk aversion scores across 9,755 survey respondents

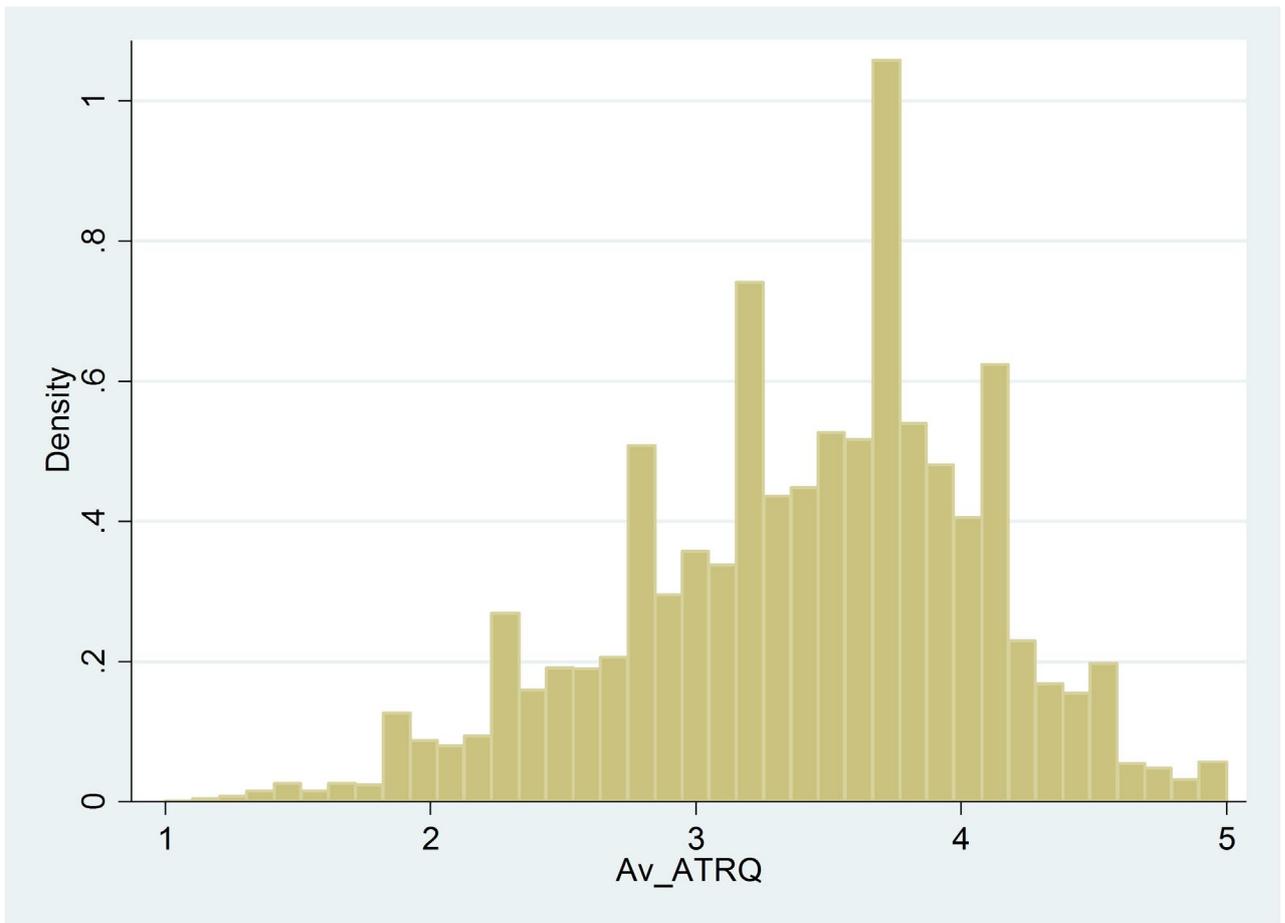
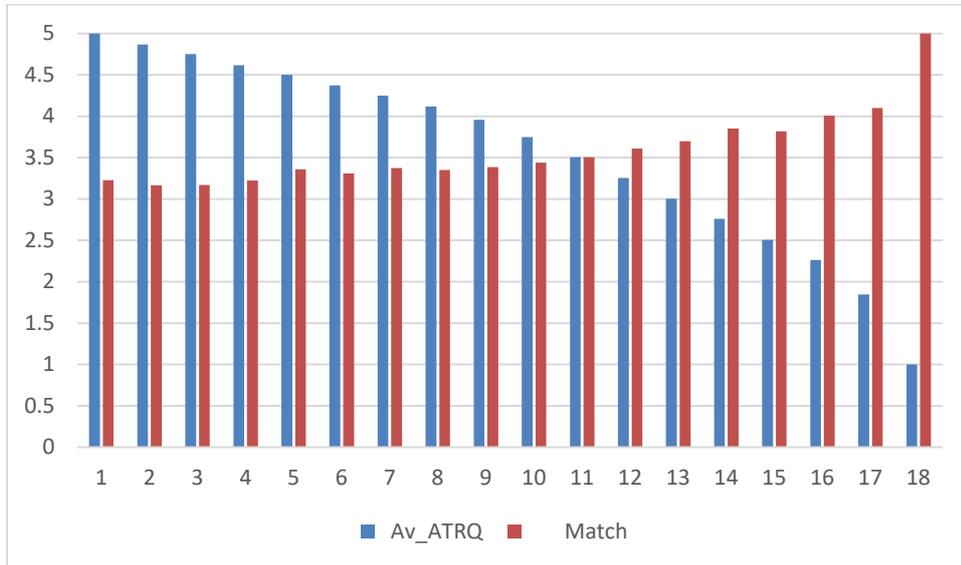


Figure 6 – Relationship between average risk aversion scores and the intention to match across 18 clusters

Panel A - Ward's hierarchical method



Panel B - k-means partition method

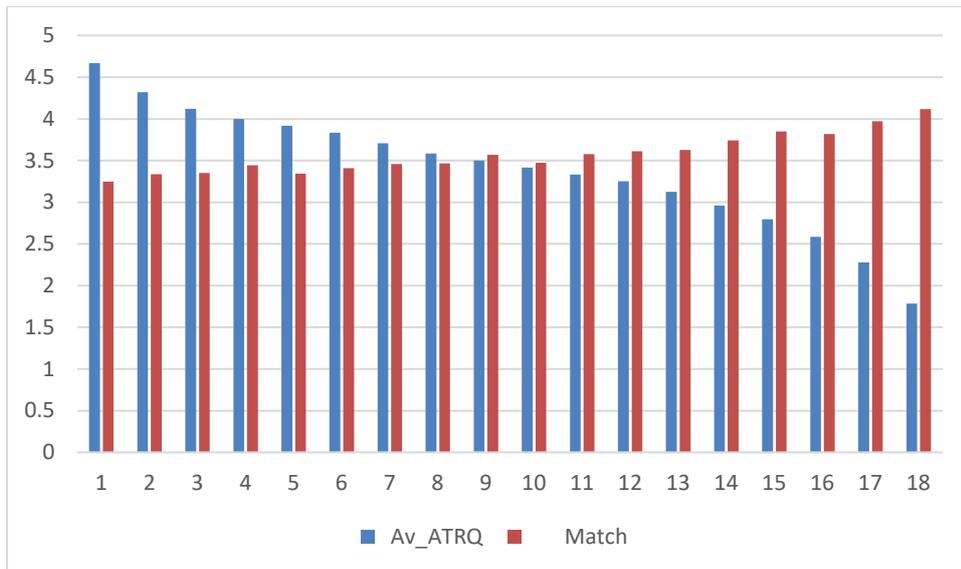


Table 1 – Summary demographics and personal characteristics of the respondents to the USS questionnaire

Panel A (Values)							
	Mean	Std. dev.	10%	25%	Median	75%	90%
Age (years)	46	0.87	32	37	47	52	57
Annual salary (£, based on bands)	£50,010	£22,830	£30,000	£40,000	£50,000	£60,000	£80,000
Expected retirement age (years)	65.01	3.39	58	65	65	67	69
USS tenure (years)	11.92	9.14	2	4	7	17	30

Panel B (Categories)			
Gender	<u>Male</u> 5,377 (55%)	<u>Female</u> 4,378 (45%)	
Marital status	<u>Married</u> (incl. civil part.) 6,360 (68%)	<u>Single</u> (incl. sep., div., wid.) 2,941 (32%)	
Job-type	<u>Academic</u> 5,768 (59%)	<u>Prof. services/Other</u> 3,987 (41%)	
Expect long, healthy retirement	<u>Agree</u> 7,177 (74%)	<u>Neither agree nor disagree</u> 2,071 (21%)	<u>Disagree</u> 507 (5%)
USS pension will be main income	<u>Agree</u> 6,745 (69%)	<u>Neither agree nor disagree</u> 1,682 (17%)	<u>Disagree</u> 1,328 (14%)
Additional contributions to USS (with AVCs/added years)	<u>No</u> 7,395 (76%)	<u>Yes</u> 2,360 (24%)	

Panel C (Age distribution)		
Age range	Number of members	Percent
<25	32	0.33
25 - 29	394	4.04
30 - 34	1,038	10.64
35 - 39	1,374	14.09
40 - 44	1,411	14.46
45 - 49	1,630	16.71
50 - 54	1,667	17.09
55 - 59	1,468	15.05
60 - 64	616	6.31
65 - 69	108	1.11
>70	17	0.17
Total	9,755	100

Note: The table presents summary information on the demographic and personal characteristics of the 9,755 members of USS who responded to the questionnaire. Numbers may not sum to 9,755 because of no-responses to some questions. In Panel A, salary information is based on mid-points of £10,000 bands. Similarly Panel C reports ages in bands, and subsequent analysis of the age variable uses mid-points of these age bands.

Table 2 – Factor analysis of responses to 2 questions on ethical investment beliefs and risk capacity

	Eigenvalues		Factor1 loading
	Factor1	Factor2	
Ethics	1.52	-0.141	0.476
Risk capacity	0.643	-0.247	0.567

Table 3 – Identifying the number of clusters for the 30s, 40s and 50s age cohorts

Number of clusters	Ward's hierarchical method			k-means partition		
	30s	40s	50s	30s	40s	50s
2	410.37	801.59	1,025.24	574.45	801.59	1,031.03
3	351.78	688.71	703.21	549.02	787.34	854.71
4	354.59	590.33	608.08	532.31	720.20	735.83
5	365.11	554.00	568.01	512.05	685.69	691.73
6	378.65	513.40	535.47	501.94	625.60	627.11
7	364.86	490.02	505.32	480.91	629.91	609.11
8	358.72	469.99	484.40	450.45	596.06	605.97
9	348.05	456.19	465.51	440.56	560.92	579.33
10	340.49	450.03	448.31	427.85	538.42	541.56

Note: Numbers in the table are values of the Calinski-Harabasz pseudo-*F* statistic for each potential cluster

Table 4 – Cross-tabulation of clusters from the *k*-means partition and Ward’s hierarchical methods for the 30s, 40s and 50s age cohorts

Panel A: 30s			
Group (Ward)	Clusters (k-means)		Total
	1	2	
1	833	243	1,076
2	589	747	1,336
Total	1,422	990	2,412

Panel B: 40s			
Group (Ward)	Cluster (k-means)		Total
	1	2	
1	2,346	0	2,346
2	0	695	695
Total	2,346	695	3,041

Panel C: 50s			
Group (Ward)	Clusters (k-means)		Total
	1	2	
1	5	1,993	1,998
2	1,137	0	1,137
Total	1,142	1,993	3,135

Note: Each panel shows the cross-tabulations of the number of observations by age-cohort of the two clusters formed by Ward’s hierarchical method (Group) and *k*-means partition (Clusters).

Table 5 – Characteristics by clusters for the 40s and 50s age cohorts (*k*-means partition method)

Variable	40s				50s			
	Cluster 1		Cluster 2		Cluster 1		Cluster 2	
	Mean	Std. dev.						
Av_ATRQ	3.39	0.68	3.26	0.70	3.43	0.66	3.33	0.70
Match	3.53	0.94	4.04	0.92	3.48	0.99	3.91	0.98
Additional_contributions	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00
Ethics_fact1	0.005	0.905	-0.058	0.950	-0.052	0.917	-0.053	0.929
Risk_capacity_fact1	-0.003	0.647	0.165	0.638	-0.062	0.698	0.173	0.635
Age	44.5	2.5	45.2	2.4	54.2	2.5	54.7	2.5
Pay	£49,856	£20,951	£55,198	£20,792	£54,041	£26,575	£61,821	£25,196
Exp_retire	65.1	3.3	64.6	3.5	64.3	3.4	63.6	3.4
Tenure	10.7	6.9	14.0	6.4	14.9	10.1	20.0	8.9
Pens_wealth	£285,334	£241,162	£399,574	£247,750	£351,053	£338,147	£504,929	£335,506
<i>F</i> (5,3013)	33.24**							
<i>F</i> (5,3103)	51.25**							
Obs	2,346		695		1,993		1,142	
%female	46.7%		38.8%		46.12%		39.25%	
%couple	72.0%		73.0%		73.03%		73.50%	
%academic	55.8%		67.8%		59.51%		68.99%	
<i>F</i> (3,2887)	13.16**							
<i>F</i> (3,3000)	10.94**							

Note: The table shows the average characteristics for the two clusters formed for the 40s and 50s age-cohorts. The *F*-statistic for a multivariate analysis-of-variance is reported to test for the joint significant differences between clusters for the five common characteristics (Age, Pay, Exp_retire, Tenure, Pens_wealth) and for the three personal characteristics (%female, %couple, %academic); ** indicates statistical significance at the 1% level. Av_ATRQ is the member's average ATRQ score. Match indicates likelihood of the member matching the available 1% employer contribution. Additional_contributions is a 0-1 dummy indicating if the member has previously made additional contributions with the scheme by making AVCs or buying added years. Ethics_fact1 is the single factor indicating the degree of member interest in making ethical investments. Risk-capacity_fact1 is the single factor indicating the member's risk capacity. Age is the member's age. Pay is the member's salary. Exp_retire is the member's expected retirement age. Tenure measures the number of years the member has been an active member of USS. Pens_wealth is the member's pension wealth. We measured this as $(1/80) \times \text{Tenure} \times \text{Pay} \times (1.051/1.022)^{(65-\text{Age})}$; this incorporates the following assumptions about USS: a capitalisation factor for the pension at retirement of 20, a lump sum of 3 x the pension at retirement, pay growth of CPI + 2%, a discount rate of gilts + 0.75% (from the USS 2017 Actuarial Valuation), with Consumer Prices Index (CPI) = 3.1% in November 2017 and the 15-year gilt yield = 1.45% on 15 December 2017. Note that this measure of pension wealth was valid at the time of the survey and does not take into account subsequent scheme rule changes from 1 April 2016. %female measures the percentage of the cluster that is female. %couple measures the percentage of the cluster that is married or in civil partnership. %academic measures the percentage of the cluster that is academic rather than professional services.

Table 6 – Identifying the number of clusters for the combined 40s and 50s age cohorts

Number of clusters	Ward's hierarchical method	<i>k</i> -means partition method
2	1868.59	1868.59
3	1443.14	1646.05
4	1210.05	1448.40
5	1074.53	1361.13
6	1008.64	1320.07
7	971.91	1258.43
8	940.97	1190.57
9	912.89	1117.80
10	876.70	1066.76

Note: Numbers in the table are values of the Calinski-Harabasz pseudo-*F* statistic for each potential cluster.

Table 7 – Cross-tabulation of the clusters from Ward's hierarchical and the *k*-means partition methods for the combined 40s and 50s age cohorts

Group (Ward)	Clusters (k-means)		Total
	1	2	
1	4,339	0	4,339
2	0	1,837	1,837
Total	4,339	1,837	6,176

Note: The table shows the cross-tabulations of the number of observations by age-cohort of the two clusters formed by Ward's hierarchical method (Group) and the *k*-means partition method (Clusters).

Table 8 – Characteristics by clusters for the combined 40s and 50s age cohorts (*k*-means partition method)

Variable	Cluster 1		Cluster 2	
	Mean	Std. dev.	Mean	Std. dev.
Av_ATRQ	3.41	0.67	3.30	0.70
Match	3.51	0.96	3.96	0.96
Additional_contributions	0.00	0.00	1.00	0.00
Ethics_fact1	-0.021	0.911	-0.055	0.937
Risk_capacity_fact1	-0.030	0.672	0.170	0.636
Age	48.9	5.4	51.1	5.2
Pay	£51,780	£23,791	£59,304	£23,831
Exp_retire	64.8	3.4	64.0	3.5
Tenure	12.6	8.7	17.7	8.6
Pens_wealth	£315,546	£291,617	£464,893	£309,331
<i>F</i> (5,6122)	113.21**			
Obs	4,339		1,837	
%female	46.4%		39.1%	
%couple	72.5%		73.3%	
%academic	57.5%		68.5%	
<i>F</i> (3,5891)	26.69**			

Note: The table shows the average characteristics for the two clusters formed for the combined 40s and 50s age-cohorts. Both clustering methods produce the same results. The *F*-statistic for a multivariate analysis-of-variance is reported to test for the joint significant differences between clusters for the five common characteristics (Age, Pay, Exp_retire, Tenure, Pens_wealth) and for the three personal characteristics (%female, %couple, %academic); ** indicates statistical significance at the 1% level. For the definition of the variables, see Table 5.

Table 9 – Probit model of the two combined 40s and 50s age cohort clusters in terms of characteristics (*k*-means partition method)

Characteristic	Coef.	z-stat
Age	0.0216	5.80
Pay	5.21e-06	3.63
Exp_retire	-0.0244	-4.79
Tenure	0.0378	7.80
Pens_wealth	-3.08e-07	-1.67
Constant	-0.7709	-2.00

Note: The table shows, for the 2 clusters formed by combining the 40s and 50s age cohorts, the results of a probit model of the five common characteristics (Age, Pay, Exp_retire, Tenure, Pens_wealth). For the definition of the variables, see Table 5. Number of obs. = 6,128, LR $\chi^2(5) = 525.74$, Prob > $\chi^2 = 0.0000$, $R^2 = 0.075$

Table 10 – Regression model of the attitude to risk scores on the characteristics for Ward’s hierarchical and *k*-means partition methods

Variable	Ward's hierarchical method		<i>k</i> -means partition method	
	Coef.	<i>t</i> -stat.	Coef.	<i>t</i> -stat.
%female	0.0474	8.21	0.0472	13.03
%couple	-0.0201	-2.06	-0.0247	-2.36
Cons.	7.4373	19.91	7.5609	14.95
Adj. R ²	0.968		0.976	
Std. err.	0.203		0.114	
No. obs.	18		18	

Note: The table shows, for the 18 clusters formed by each of the Ward’s hierarchical method and the *k*-means partition method, the results of a regression of the average risk attitude question score (av_ATRQ) in each cluster on, respectively, the percentage of females (%females) and the percentage that is married or in civil partnership (%couples) in the same cluster. For the definition of the variables, see Table 5.

