Annex 5: Identifying metrics to aid consumer choice in the income drawdown market

Oxera and the Nuffield Centre for Experimental Social Sciences

Prepared for the
Financial Conduct Authority

March 2017

www.oxera.com

oxera
compelling economics

Contents

Boxes, figures and tables

Abstract

Following the introduction of 'pension freedoms' in April 2015, many people have chosen to shift their pension pots into income drawdown products before or at retirement, rather than buying annuities as was typical before the reforms.

Income drawdown products can be relatively complex products with multiple features and often with an array of fees. Charges and fees vary with pot size and how the product is used, so the same drawdown product could be relatively cheap for one consumer, but relatively expensive for another. In order to get the best deal, consumers need to compare products across multiple dimensions, usually with no single price or fee to focus on.

There are different ways of presenting the cost of an income drawdown pension product. In this context, the Financial Conduct Authority (FCA) wished to understand which way of summarising the cost of the product was best able to help consumers to identify, from a range of product options, which was the most cost-effective for them. The FCA therefore commissioned Oxera and the Centre for Experimental Social Sciences (CESS) to conduct a behavioural experiment to assess the effectiveness of different summary cost metrics.[1] This report presents the findings of the behavioural experiment.

In this online experiment, participants were presented with a set of drawdown products (designed on the basis of products available on the market) and a consistent decision-making environment to see how the provision of different summary cost metrics affected their product choices, with all other factors held constant. The only other variable was the scenario that the participant chose to best characterise the retirement plans they envisage, which had a limited impact on the summary cost measures (which were personalised according to these scenarios).

The experiment focused on the issue of cost—other factors, such as quality of service, were included in the comparison table, but the participants were instructed to focus on cost minimisation in this experiment. We recognise that cost is only one of a range of factors that influence the value for money offered by drawdown products. However, concerns around transparency of costs meant that a focus on this component was appropriate in this study. The findings from this study will be considered alongside evidence around other components of drawdown products as part of wider FCA work. This study focused on the cost of the drawdown 'wrapper', and not, for example, on the charges associated with the underlying investments within the wrapper (which were not included in the summary cost metrics and were held constant[2] where relevant for modelling purposes).

The results of the experiment show that two of the five personalised summary cost metrics we tested had a statistically significant positive impact on the product choices of participants, resulting in them selecting lower-cost products on average, and being better able to select the cheapest product. These two cost metrics were 'pension savings available after costs' and 'average cost per year'. The other three summary cost metrics, 'total cost', 'reduction in yield'

[1] The focus was on the costs associated with the income drawdown product directly; the underlying fund charges were not varied in the experiment.
[2] At 1% per annum.

and ‘cost rating’, did not have a statistically significant effect on improving product selection performance.

The superior performance of the ‘pension savings available after costs’ and ‘average cost’ metrics may be due to a number of reasons, including the metrics being presented in monetary terms rather than percentage figures (people often struggle to understand percentages), the salience of these metrics to participants (in terms of explaining the financial impact), and trust in their relevance and reliability (which may work against the ‘cost rating’ approach).

1 Introduction and summary

Oxera and the Centre for Experimental Social Sciences (CESS) have conducted a behavioural experiment on behalf of the Financial Conduct Authority (FCA) to identify the effectiveness of certain summary cost metrics in aiding consumer selection of income drawdown products. This report presents the findings of this experiment.

1.1 Objective

In April 2015 those saving for retirement in the UK gained new flexibility over how they used their savings. The 'pensions freedoms' provided new opportunities, including greater access to income drawdown products. Sales of income drawdown products subsequently grew sharply, from 9,500 to 18,800 per quarter (Q2 2014 to Q2 2015), rising to above the level of pension annuities being sold.[3]

Many of the people choosing income drawdown products are relying on financial advice, however, a substantial minority (32% in Q4 2015) purchase drawdown without advice.[4] This increase in people choosing drawdown products without taking financial advice is partly because, prior to April 2015, income drawdown products had been limited to those with larger pension savings and thus more able to afford financial advice.

Choosing a pension product is an infrequent event for most people—so there is limited opportunity for learning from past experience. There is also limited opportunity to learn from other people's experiences, as the market is relatively new and the consequences of the choice of an expensive pension product occur over many years. The FCA has found low levels of switching when people chose the provider of their income drawdown product. In Q4 2015, 53% of income drawdown products were purchased by existing customers.[5] For competition to work effectively, the FCA believes that non-advised income drawdown customers need to be able to shop around and compare drawdown products on cost.

However, unlike annuities, some income drawdown products can have complex charging structures, a wide range of fees and limited transparency. In particular, administration charges for income drawdown products can differ between providers in terms of the number and types of fees charged. For example, some income drawdown products have a number of different administration fees, including initial set-up fees, transfer fees, annual administration charges, and product fees, whereas some simpler products may only have an annual administration charge.

If a consumer decides to shop around for an income drawdown product, they face a decision that includes comparing products across multiple dimensions. To the extent that the consumer wishes to focus on cost, there is often no single price or fee to focus on, and certainly not charges that are comparable across products.

[3] Association of British Insurers (2015), 'UK Insurance & Long Term Savings Key Facts 2015', September, https://www.abi.org.uk/~/media/files/documents/publications/public/2015/statistics/key%20facts%202015.pdf, accessed 3 February 2017.
[4] Financial Conduct Authority (2016), 'Retirement Outcomes Review', Terms of Reference, MS16/1, July, https://www.fca.org.uk/publication/market-studies/retirement%20outcomes%20review%20tor.pdf, accessed 3 January 2017.
[5] Financial Conduct Authority (2016), 'Retirement Outcomes Review', Terms of Reference, MS16/1, July, https://www.fca.org.uk/publication/market-studies/retirement%20outcomes%20review%20tor.pdf, accessed 3 January 2017.

There are different ways of presenting the cost of an income drawdown product. The FCA wished to investigate which of these ways of summarising the cost of the product is better able to help consumers to identify which from a range of product options is the cheapest.[6]

The FCA wanted to investigate whether combining the various fees into a summary cost measure (or 'metric') could help customers better compare products according to their administration charges.[7] The objective of this study is to identify which measures of cost help consumers the most in this context.

As the summary cost metrics depend on how the consumer uses the products (because charges apply differently according to fund size, and the timing of withdrawals, for example), the cost metrics were personalised, based on scenarios for usage that the participants in the experiment chose for themselves (see section 3.2 for details). Product costs can vary significantly according to how the products are used, and this element of personalisation was therefore an important aspect of the summary cost metrics.

In theory, a summary cost measure could accomplish three goals, and enable better decision-making, by:

- bringing to light any 'hidden charges' or charges in the small print;
- reducing information overload and making it easier for consumers to compare a large number of underlying charges. Instead of computing the likely cumulative impact of numerous fees that vary by magnitude, form (percentage or fixed) and timing (one-off, annual or incurred following consumer actions), consumers will only have to compare a single figure;
- making it easier for consumers to compare charges based on likely product usage.

This study tested the effectiveness of five cost metrics to identify the measure that would be likely to help consumers the most in the experimental set-up. The experiment focused on the issue of cost—other factors, such as quality of service, were included in the comparison table, but the participants were instructed to focus on cost minimisation in this experiment. We recognise that, in real-world settings, consumers make choices along many dimensions, with cost being one of them.

1.2 The experiment

To help address these challenges in comparing the costs of drawdown products, the FCA wished to test the effectiveness of summary cost metrics in aiding the decision-making process. To this end, the experiment tested the effect of five summary cost metrics in an environment in which participants had to choose the cheapest product from a table of products—whereby the table represented a choice environment. The experiment had the following features:

- participants were all aged 40+ with a mean and median age of 58. Those in the experiment will soon face decisions over retirement income, if they have not already done so;

[6] This research is focus on income drawdown products. It is expected, however, that any results will also apply to alternatives, like uncrystallised funds pension lump sums (UFPLS). These apply similar categories of charges.
[7] Underlying fund charges or adviser charges were not included in the cost metrics. However, a flat rate for underlying fund charges was used when necessary when calculating the simulated summary cost metrics to be used in the experiment.

- participants faced a choice of realistic products. Product charges and features were similar to those of income drawdown products in the real world, albeit standardised in the table format;[8]
- participants chose product usage scenarios and the cost metrics were calculated on the basis of those scenarios;
- the language used in the experiment that explained different features or charges was based on the language mandated or recommended by the FCA;
- participants were incentivised to choose the cheapest product in the experiment.

The experiment sought to encourage participants to act as they would in real life when it came to which choice they made. In particular, the experiment was designed to reduce the risk of participants 'playing the game' to comply with what they thought was expected of them in the experiment.[9]

1.2.1 Information treatments

Our sample of 2,020 participants from the UK population nearing retirement age (i.e. over age 40) was randomly divided into seven groups. Each of the groups was shown a slightly different table of products—i.e. the table differed in which summary metric or which control version was displayed (see Figure 1.1 below), but the treatments differed in no other way.

Before facing the table of products, participants were asked to select a usage profile. This was done so as to put the participants in the mind-set of choosing an income drawdown product as they would be in the real world. Six illustrations of usage (vignettes) were presented to the participants. Three led to the participant being presented with the cost of the product if they took regular withdrawals, while the other three led to the participant being presented with the cost of the product if they used it as a long-term investment device. We can therefore compare results across usage profiles (which either essentially provide regular income immediately or act more as a long-term investment, as discussed in section 3.1 below).

To reflect the complexity of real-world choice environments, the summary cost metric was placed alongside a number of other (financial and non-financial) product characteristics. The table of products was also designed to be sufficiently complex so that the treatment summary cost metric was not overly prominent, and therefore salient in the decision-making process. If it had been too salient then it would have been difficult to observe variation between the treatments.[10] The table of products listed 18 products[11] according to four 'quality' measures and five fees (seven fees in the complex control), plus the treatment

[8] For example, if 50% of real-world products charge a fee, then half the products in the experiment also charged that fee. The minimum, maximum and mean of each fee across providers were the same in the experiment as in the real world.
[9] This is known as 'experimenter demand effects'. See Zizzo, D.J. (2010), 'Experimenter demand effects in economic experiments', *Experimental Economics*, **13**:1, pp. 75–98.
[10] If the cost metric was made to be highly prominent, or if there was little other information to draw the participant's attention, then it would have been likely that the vast majority of participants would have chosen the cheapest product. Not only is this unlike a real-world situation (where there is information about multiple product features), but the triviality of the problem presented to participants would have meant that there would have been little variation in the product that participants chose. Minimal variation would have made it difficult to identify the best summary cost metric.
[11] The products were not presented in any kind of ranking (and no indication was given to participants of any kind of ranking), and the table of products did not have any kind of 'sort by' function. A In the experiment the cheapest product was listed third in all the treatments. Therefore, if any participants reduced their decision to just the products at the top of the table, they could still choose the cheapest product.

metric. Participants were able to see definitions by hovering their mouse over question mark symbols next to the column headings, and could click on 'Explore charges', which listed the charges information as shown in the table of products.

The seven groups (five summary cost metric treatments and two 'controls', which did not include summary cost metrics) were as follows.

- **Simple control**—no summary cost metric and five core charges initial set-up fee, transfer-in payment, annual administration charge, product fee, and unscheduled withdrawal fee. This group faced a slightly less challenging task than the complex control, in order to identify the potential impact of a reduced amount of information being presented. This group represents the true 'control' for the summary cost metrics, as it is identical except for the exclusion of the summary cost metrics.

- **Complex control**—no summary cost metric and eight of the most popular types of charges for income drawdown products:[12] initial set-up fee, transfer-in payment, annual administration charge, product fee, unscheduled withdrawal fee, fee for purchasing assets online, QROPS fee[13] for transferring to a foreign pension, and drawdown review fee. This group faced the most challenging task in identifying the cheapest product. This group provides an additional 'control' group in the sense that no summary cost metric is included, although it also differs from the treatments by having additional information included.

- **Reduction in yield**—as the simple control, but with reduction in yield (%) calculated over a 20-year time period shown as a summary cost metric. Reduction in yield was described as 'The reduction in the annual returns of your pension, as a result of the charges that you will be paying'.

- **Pension savings available after costs**—as the simple control, but with the present £ value of the pension pot less the present value of costs for a product over a 20-year period (discounting based on assumed investment return) shown as the summary cost metric. Pension savings available after costs were described as 'The estimated value of your pension pot, at this time, after taking account of the impact of all charges over the next 20 years on the income drawdown product you choose'.

- **Average cost (per year)**—as the simple control, but with the average £ cost of a product over a 20-year period shown as the summary cost metric. Average cost was described as 'The estimated average £ amount per year of charges that you are likely to pay on your income drawdown product.'

- **Total cost (20 years)**—as the simple control, but with the total £ cost of a product over a 20-year period shown as the summary cost metric. Total cost (20 years) was described as 'The total charges over 20 years that you are likely to pay on your income drawdown product.'

- **Cost rating**—as the simple control, but with a rating summarising all charges from a product on a scale from £ (cheapest) to £££££ (most expensive)[14]

[12] Research undertaken by the FCA on 30 providers of income drawdown products in July 2016.
[13] A Qualifying Recognised Overseas Pension Scheme (QROPS) is an overseas pension scheme that meets requirements set by Her Majesty's Revenue and Customs (HMRC) for tax efficient transfers.
[14] Cost rating based on reduction in yield over a 20-year period, with only one product marked as being in the cheapest cost category (£), products within 10 basis points of the cheapest reduction in yield being marked in the second cheapest category (££), and further categories including subsequent groups of products in a range of 10 basis points, up to the £££££, category which included all products more than 30 basis points more expensive than the cheapest product.

shown as the summary cost metric. The cost rating was described as follows: 'The charges rating summarises all charges for the income drawdown product on a scale from £ (cheapest) to £££££ (most expensive). The scale provides a simple comparison of charges across different products on offer'.

All of the metrics assume that product fees remain constant over the lifetime of the product, and product usage remains consistent with the chosen vignette over the lifetime of the product.

Figure 1.1 Table of products with cost rating treatment shown

Please choose your income drawdown product

Pension pot £50,000 Annual drawdown 0%

Your options

Please select a product and click the 'Next' button.

Provider	Customer service rating	24/7 helpline?	Online access	Minimum monthly withdrawal	Initial set-up fee	Transfer in fee	Annual administration charge	Product fee	Cost rating	Unscheduled withdrawal fee	Your choice	Explore additional charges
Provider A	★★★½☆	✓	✓	£50	£350	-	£520	0.20%	£££	-	○	Explore
Provider B	★★★★☆		✓	£150	-	£500	£545	0.21%	££££	-	○	Explore
Provider C	★★★★½	✓	✓	£100	-	-	£125	0.43%	£	-	○	Explore
Provider D	★★½☆☆			£150	£200	-	£525	0.23%	££££	-	○	Explore
Provider E	★★★☆☆	✓	✓	£150	£150	-	£175	0.58%	££££	£250	○	Explore
Provider F	★★½☆☆			£125	-	-	£400	0.49%	£££££	-	○	Explore
Provider G	★★½☆☆	✓		£150	-	-	£300	0.51%	££££	£70	○	Explore
Provider H	★★★½☆		✓	£50	£100	-	£250	0.43%	£££	£25	○	Explore
Provider I	★★★★½			£50	£100	-	£375	0.26%	££	-	○	Explore
Provider J	★★½☆☆			£100	£350	£13	£225	0.47%	£££	£115	○	Explore
Provider K	★★☆☆☆	✓	✓	£125	£300	£20	£200	0.48%	£££	-	○	Explore
Provider L	★★★★½	✓	✓	£100	£110	£100	£475	0.31%	££££	-	○	Explore
Provider M	★★★☆☆		✓	£100	-	-	£450	0.29%	£££	-	○	Explore
Provider N	★★★★☆	✓	✓	£75	-	£210	£425	0.35%	££££	-	○	Explore
Provider O	★★★½☆	✓		£150	-	-	£540	0.39%	£££££	-	○	Explore
Provider P	★★☆☆☆	✓		£100	-	£110	£75	0.55%	££	-	○	Explore
Provider Q	★★½☆☆			£50	-	-	£150	0.60%	£££	-	○	Explore
Provider R	★★★½☆	✓	✓	£50	£100	-	£100	0.59%	£££	-	○	Explore
Provider	Customer service rating	24/7 helpline?	Online access	Minimum monthly withdrawal	Initial set-up fee	Transfer in fee	Annual administration charge	Product fee	Cost rating	Unscheduled withdrawal fee	Your choice	Explore additional charges

Note: a fund management fee of 0.76% is charged for the default investment fund in each product. Please note that the fund management fee you are charged may vary depending on your choice of investment funds.

Source: Oxera/CESS.

The best product in each table was set to be dominant across price and quality features—if participants valued quality highly and made a trade-off between quality and price, they would still choose the cheapest product.

1.3 Key findings

The key outputs of this experiment, which focused on the cost of drawdown products (explained above, the 'wrapper' costs only), were the measures of the performance of the participants in selecting the product that minimised such costs. The measures used were:

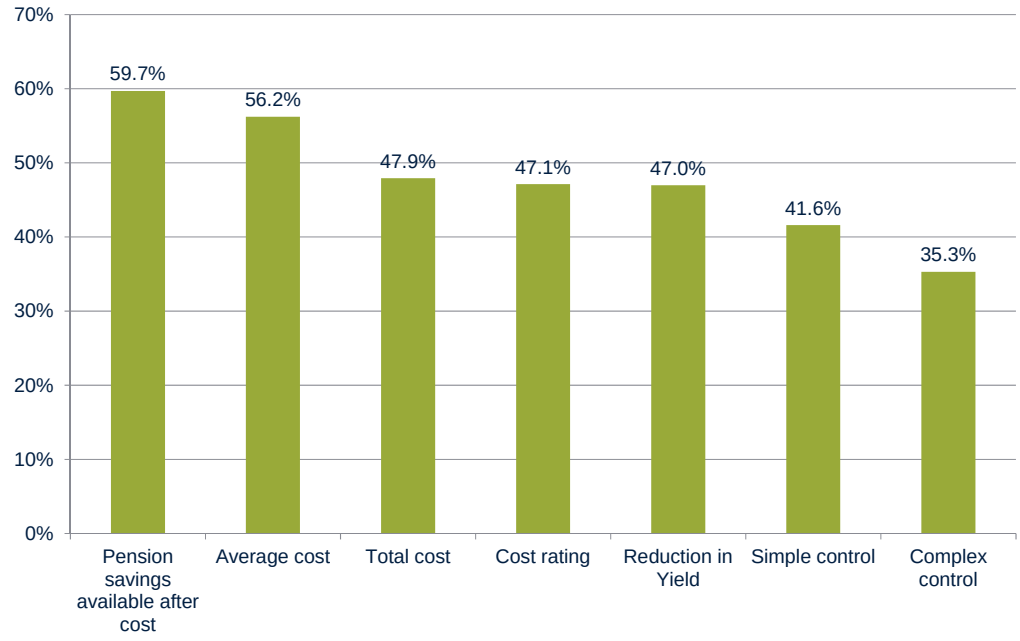
- the monetary reward paid to the participant for their choices, based on the payment schedule (see section 3.6 for details of the reward structure);
- whether the participant selected the cheapest product;[15]
- whether the participant avoided the 13 most expensive products[16] (that provided a product selection reward of less than £2).

The results of the experiment show that two of the personalised summary cost metrics, pension savings available after costs and average cost per year, had a statistically significant positive impact on the product choices of participants, resulting in them selecting lower cost products on average, and being better able to select the cheapest product (see Figure 4.2 and Figure 4.4).

When considering the monetary reward payoff measure, pension savings available after costs had a greater positive impact than the average cost per year, and the difference between the level of impact for each of these measures was statistically significant (see Figure 4.2). With the pension savings available after costs metric, 59.7% of participants chose the cheapest product, while with the average cost per year metric, 56.2% did so. However, the difference between the level of impact for each of these measures was not statistically significant (see Figure 4.4).

[15] Product C was the cheapest product for all of the profiles and was indicated to be the cheapest by all of the summary cost metrics.
[16] The bottom thirteen products were selected to provide a sufficient sample for analysis. There was no other reason for this cut-off point. Choosing 12 products, for example, does not materially change the results.

Figure 1.2 Proportion of participants who chose the cheapest product



Note: Based on whole sample of 2,020 participants.

Source: Oxera/CESS.

The outperformance of the pension savings available after costs and average cost metrics remained constant across the two usage profiles, age groupings, gender and education.

1.4 Conclusions

The research finds that two of the cost metrics were effective at encouraging participants to select lower cost products in the scenario presented to them. In reality, consumers need to make choices along many dimensions, with cost being one of them, but the research does suggest that summary cost metrics can add value through improved product selection.

The superior performance of the 'pension savings available after costs' and 'average cost' metrics may be due to a number of factors. Possible explanations, while keeping in mind the inherent uncertainties, could include that the metrics (unlike some of the other metrics):

- present the costs in £ terms rather than a percentage;
- make calculations easier by computing how the charges affect the current value of the pension pot or the annual income;
- avoid doubt over the relevance of costs, and make the salience of the charges clear.

While various potential reasons for the outperformance can be proposed, ultimately the strength of metrics in encouraging desired behaviours is an empirical question, and hence the key motivation for the behavioural experiment.

2 Motivation

One of the key insights of behavioural economics is that people can make poor decisions even when they have access to all the information necessary to make optimal decisions.

The decision about which income drawdown product to select is relatively complex. The rational, rules-based side of the brain—or 'System II'—may find it difficult to arrive at the right answer.[17] What the consumer may then rely on is what their instinct tells them, or rules of thumb (heuristics). This is the more emotive, and essentially behavioural, System I. While System I can offer useful shortcuts, it can make consumers more susceptible to making poor decisions when faced with complexity—in this case selecting a more expensive provider.

The selection of a more expensive provider may be for various reasons (all of which stem from suppression of System II reasoning, and biasing of System I), including:

- information and choice overload—this can lead to 'status quo bias', where even when offered alternative deals, consumers select their existing 'trusted' pension provider for income drawdown (e.g. due to loss aversion and regret avoidance);
- narrow framing and salience of particular charges—this will lead to customers focusing on some fees more than others (e.g. set-up fees); consumers may compare only two or three providers and select the best of these, rather than searching the whole market for the best deal;
- multiple attributes generating faulty heuristics—consumers who *do* shop around may choose the wrong product and/or provider;
- optimism bias regarding forecasting future use—consumers may, for example, underestimate the number of ad hoc withdrawals they will make. Although the uncertainty around a consumer's future consumption requirements means that they will likely find forecasting drawdown usage difficult, even if consumers were not over-optimistic;
- mistakes—consumers may simply lack the ability to compare the offerings effectively, and choose a poorer deal (e.g. by focusing on salient set-up charges only, or underestimating usage and ad hoc charges).

[17] Daniel Kahneman set out the concepts of two modes of thought, referred to as 'System 1' (fast, instinctive and emotional) and 'System 2' (slower, more deliberative and more logical), in Kahneman, D. (2011), *Thinking, Fast and Slow*, Macmillan.

Box 2.1 Kalayci and Potters (2010)

The issue of charging complexity was examined in a laboratory experiment by Kalayci and Potters (2010).[1] This involved both buyers and sellers, and considered how multiple attributes can lead to buyer confusion. The results indicated that buyers make more sub-optimal choices when the number of attributes chosen by the sellers is higher. Often consumers would simply choose the 'most popular' option.

Confusing consumers allows higher prices to be charged for the same underlying quality of the good. A further finding of the authors is that prices and profits were lower when the sellers were informed that the buyers would be replaced with perfectly rational ('robot') buyers.

The Kalayci and Potters experiment reveals how using System II in order to make a consumer purchasing decision over products with multiple attributes can be difficult. Recourse to System I heuristics can then lead to biases, with sub-optimal choices being made.

Note: [1] Kalayci, K. and Potters, J. (2010), 'Buyer confusion and market prices', *International Journal of Industrial Organization*, **29**:1, pp. 14–22.

With income drawdown products, consumers may have to compare a number of different charges. This typically includes the computation of:

- percentage fees—people often do not correctly understand percentage fees;[18]
- the addition of percentage and flat fees—this can be a challenging calculation;
- fees that vary by tier—this can also be a non-trivial calculation.

In addition, the fees depend on investment rates of return, which will have a bearing on the size of the pot in the future, and therefore the size of charges based on percentage fees. These are uncertain.

In summary, the fees applied to income drawdown products are often complex and varied, with different products charging different fees. In this case, computational requirements are likely to be quite significant, and hence consumers will face challenges in identifying the total cost of products.

2.1 The use of cost metrics

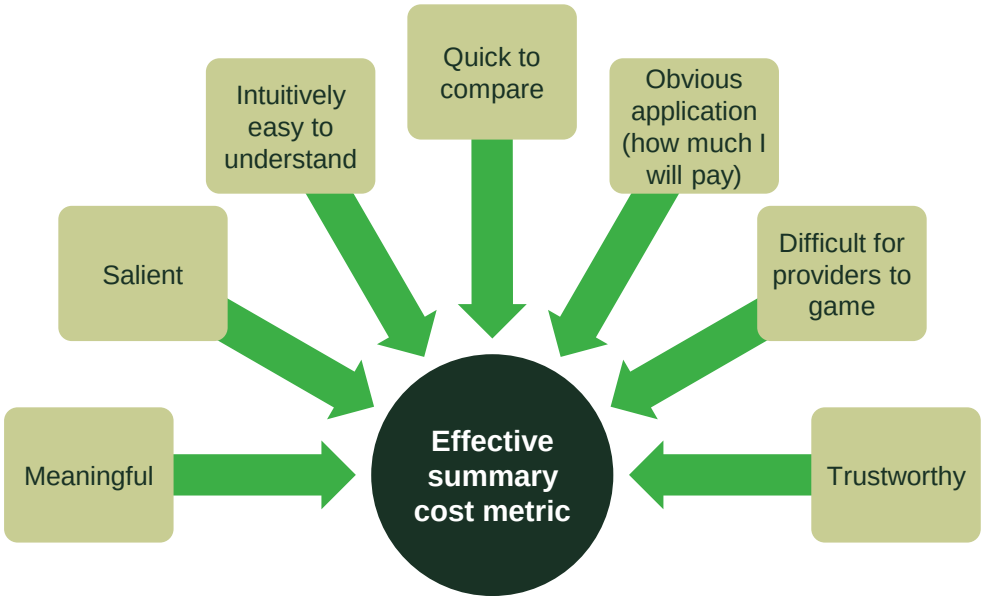
Summary cost metrics presented individually as part of a firm's drawdown product pre-sale disclosure, or incorporated into comparison tables of products may help consumers to make comparisons across a wide spectrum of providers (broad framing) and to identify the best deal for them.

Specifically, a summary cost metric will focus people's attention to reduce the impact of information overload, and if the metric is clear or familiar, it can focus attention even in an unfamiliar environment. The metric can be made more salient relative to others if it is clear that it is a total or summary cost metric.

[18] Banks, J. and Oldfield, Z. (2006), 'Understanding pensions: Cognitive function, numerical ability and retirement saving', Institute for Fiscal Studies Working Paper WP06/05, Final version February 2006, available at: http://discovery.ucl.ac.uk/2690/1/2690.pdf [Accessed 3 January 2017].

However, not all summary cost metrics are equally effective in this regard. Figure 2.1 shows the factors that make summary cost metrics effective.

Figure 2.1 Effective summary cost metric



Source: Oxera.

Ideally, people should find it intuitively easy to understand what the cost metric is telling them and whether a high or low value is optimal. They should be able to quickly compare the metric across different products. In short, the summary cost metric should appeal to, and be easily understood by, System I rather than requiring excess attention from System II. It should also be trustworthy—people should not doubt whether the metric captures all of the relevant available fees and charges.

3 The experiment

The prime objective of the experiment was to explore how summary cost metrics affect people's performance when selecting income drawdown 'wrapper' products according to their cost. We assumed that other charges such as underlying fund costs did not vary by product. The experiment aimed to identify differences in the impacts of different summary cost metrics. The main elements of the experiment, described in this section, were as follows.

1. Experimental environment
2. Usage profile selection
3. Information treatments
4. Group assignment
5. Data collection
6. Outcomes

3.1 Experimental environment

Developing an appropriate online environment in which participants would select income drawdown products was a key element of the experiment. This design process involved two stages of pre-testing: a laboratory test followed by a focus group to see how well participants understood the tasks; and a small online experiment using Amazon Mechanical-Turk[19] with 311 participants to see whether the experiment generated meaningful results. Both were successful and resulted in only minor changes to the experiment.

We conducted the final online experiment with a sample of 2,020 participants aged 40+ who were registered panel members with Respondi, over the period 6–19 January 2017. Respondi is a large online panel in the UK with 45,000 registered subjects, 5,000 of whom are in the 40–65 age bracket.[20]

Respondi uses many types of invitation to bring in people with diverse motivations to take part in research. These include email invitations, text messages, telephone alerts, and banners and messaging on websites and in online communities. The messages themselves are also varied, and include invitations to 'give your opinion' or 'let your voice be heard'. This diversity of motivations is likely to have contributed to a high-quality sample. To avoid self-selection bias, specific project details were not included in the invitation. Rather, participants were invited to 'take a survey', with the details disclosed later.

Participants were told that the whole experiment would take around 25 minutes to complete, with two-thirds of participants completing it in 25 minutes or less. The five key stages of the experiment are described below.

- **A: introduction and instruction**: participants were first asked a series of questions about themselves, and were provided with a set of instructions.
- **B: explanation and context**: relevant terms and concepts were explained, such as what an income drawdown product is and how it differs from an annuity. Participants were tested on their understanding of income drawdown

[19] Amazon Mechanical-Turk is an online platform that matches self-employed workers with online tasks. See https://www.mturk.com.
[20] Respondi is an ISO 26362-certified survey company.

products, and had to demonstrate some understanding before they could continue.

- **C: usage profile selection**: participants selected a usage profile of how they would expect to use an income drawdown product: either to provide regular income or as a long-term investment. These usage profiles were described using vignettes to encourage participants to think about how they would use an income drawdown product. This informed the price comparison website task by helping to generate representative summary cost metric figures using this information.

- **D: comparison table task**: participants were asked to choose the cheapest product from a table of products and they were instructed that they would be paid based on their decision. The information they saw varied by treatment. This was the core part of the experiment.

- **E: other questions**: the participants answered a series of questions around time and risk preferences, as well other factors. This data allowed us to see whether there were any underlying drivers behind the results other than the treatment.

As the focus of this study was to assist customers who had already decided to shop around for an income drawdown product, and this behaviour will typically take place online, the experimental environment was designed to represent a real-world environment for selecting an income drawdown product.

In order to reflect the complexity of real-world choice environments, the summary cost metric was placed alongside a number of other (financial and non-financial) product characteristics. This was done in the price comparison website setting, familiar to consumers. The table of products screen in the experiment was designed to be sufficiently complex that the treatment summary cost metric was not overly salient.

3.2 Usage profile selection

Participants were asked to select a usage profile regarding how they would use an income drawdown product. This encouraged them to put themselves into the mind-set of purchasing an income drawdown product. There were two usage profiles in the experiment.

Participants selected their usage profile from a list of six vignettes that described how someone who is considering an income drawdown product would use the product.[21] Each vignette provided a description of a person with different motivations and a different life situation.[22] Three of the vignettes described a usage pattern that was consistent with drawing regular income from the income drawdown product, and three described a usage pattern that was consistent with using an income drawdown product as a vehicle for long-term investment.

The two usage profiles were chosen to reflect how people may use income drawdown products in the real world.[23] As the cost of an income drawdown

[21] See Figures A2.9 to A2.20 in Appendix A2.
[22] The vignettes used male names for male participants and female names for female participants to avoid the name influencing choice.
[23] A third usage 'instant cash' profile was considered where the pot is totally withdrawn (exhausted) within a very short period after purchasing the product. However, this was excluded from the experiment because (a) differences in charges have less impact over a single-period product lifetime; (b) immediately withdrawing the whole pot is not representative with the optimal service that many flexi-access income drawdown products are designed to offer; and (c) adding another usage profile reduces experiment cell size.

product depends on the usage pattern,[24] each usage profile resulted in a different set of costs presented in the price comparison website. We could compare results between these two usage profiles because we tested whether or not people chose the cheapest product.

Two of the six vignettes are shown in Table 3.1, one from each of the usage profiles.

Table 3.1 Example usage profile vignettes

Regular income usage profile	Long-term investment usage profile
Barbara is looking forward to retiring in the near future. She owns her house outright, but has limited funds invested or saved outside of her pension pot. She budgets well on a monthly basis, and may expect a few larger expenditures every now and again—perhaps a trip to Spain in the summer or a new boiler in a few years' time. Barbara has considered using her pension pot to purchase an annuity, but this would not leave an inheritance for her children. She therefore prefers an income drawdown product. Barbara would like to use her pension to receive a regular income during retirement, but with the flexibility to make additional withdrawals.	Charles loves his job and plans to work for several more years to come. As he is nearing retirement age, Charles is also looking at options for using his pension. He might like a little extra income to supplement his pay, but generally wants to keep his pension pot invested for years to come. He would also like to have this as a rainy-day fund: in case he needs a larger sum of money for an unforeseen circumstance. Charles would like to keep his pension pot invested with an option to withdraw additional funds if necessary.

Source: Oxera/CESS.

The pre-testing found roughly similar proportions of individuals who stated an interest in each usage pattern,[25] and this was confirmed in our main experiment. The pre-testing focus group also found that the vignettes helped participants put themselves in the mind-set of someone choosing an income drawdown product.

3.3 Information treatments

The experiment had 14 cells, as per Table 3.2 below, which resulted in an average of 144 participants per cell.

There were five summary cost metrics and two control treatments. The simple control treatment was the same as the summary cost metric treatments except for the lack of a summary cost metric. The simple control treatment was therefore the comparison that should be made when comparing the effectiveness of the treatments against having no treatment.

The complex control was the same as the simple control but with three more charges shown (three more columns). Thus the difference between the simple and complex controls was the impact of reducing the amount of information. This was tested to see whether the effect of simply reducing the information shown to consumers had the same effect as introducing a summary cost metric. Specifically, the effect of reducing the information along three dimensions, which were: the fee for purchasing assets online; the QROPS[26] fee for transferring to a foreign pension; and the drawdown review fee.

The purpose of having two controls is summarised below.

[24] For example, a consumer who makes use of ad hoc withdrawals will face different charges to a consumer who does not.
[25] In pre-testing, 54% of participants chose 'regular income' while 46% chose 'long term investment'.
[26] QROPs refers to a Qualifying Recognised Overseas Pension Scheme.

- The simple control means that the effect of having any summary cost metric can be measured against the counterfactual where there is no such metric.
- The complex control means that the effect of having a comparison that simplifies the decision by limiting the displayed product dimensions can be estimated. This is done by comparing the simple and complex controls.

Table 3.2 Cells in experiment

Usage profile	Information treatment	Number of participants
Regular income (ad hoc withdrawals)	Simple control	150
Regular income (ad hoc withdrawals)	Complex control	149
Regular income (ad hoc withdrawals)	Reduction in yield	136
Regular income (ad hoc withdrawals)	Pension savings available after costs	148
Regular income (ad hoc withdrawals)	Average cost	162
Regular income (ad hoc withdrawals)	Total cost (20 years)	134
Regular income (ad hoc withdrawals)	Cost rating	141
Long-term investment	Simple control	148
Long-term investment	Complex control	154
Long-term investment	Reduction in yield	171
Long-term investment	Pension savings available after costs	135
Long-term investment	Average cost	119
Long-term investment	Total cost (20 years)	133
Long-term investment	Cost rating	140

Note: As usage profiles were chosen by participants, it was not possible to allocate the same number of each usage profile to a treatment, since there were slightly more 'regular income' participants (50.5%) than 'long-term investment' participants (49.5%). We did, however, balance the sample within each usage profile such that a similar number of participants of each usage type were allocated to each information treatment.

Source: Oxera/CESS.

Each price-comparison screen included the following characteristics, in addition to fee information, which varied by treatment: provider name, customer service rating, 24/7 helpline, online access, and minimum monthly withdrawal. In each information treatment, there was a button-activated pop-up charge sheet including all charges (but no summary metrics) for a product. The fee information shown in each of the information treatments was as follows.

- **Complex control**—no summary cost metric, and eight of the most popular types of charges for income drawdown products:[27] initial set-up fee, transfer-in payment, annual administration charge, product fee, unscheduled withdrawal fee, fee for purchasing assets online, QROPS fee for transferring to a foreign pension, and drawdown review fee.
- **Simple control**—no summary cost metric, and five of the most common charges: initial set-up fee, transfer-in payment, annual administration charge, product fee, and unscheduled withdrawal fee.
- **Reduction in yield**—reduction in yield (%) calculated over a 20-year time period shown as the summary cost metric, as well as the individual charges from the simple control. Reduction in yield was described as 'The reduction in

[27] Research undertaken by the FCA on 30 providers of income drawdown products in July 2016.

the annual returns of your pension, as a result of the charges that you will be paying'.

- **Pension savings available after costs**—the £ present value of the pension pot less the present value of the costs for a product over a 20-year period (calculated and discounted on assumed investment return) shown as the summary cost metric, as well as the individual charges from the simple control. Pension savings available after costs was described as 'The estimated value of your pension pot, at this time, after taking account of the impact of all charges over the next 20 years on the income drawdown product you choose'.

- **Average cost**—the average annual £ cost of a product over a 20-year period shown as the summary cost metric, as well as the individual charges from the simple control. Average cost was described as 'The estimated average £ amount per year of charges that you are likely to pay on your income drawdown product.'

- **Total cost (20 years)**—the total £ cost of a product over a 20-year period shown as the summary cost metric, as well as the individual charges from the simple control. Total cost (20 years) was described as 'The total charges over 20 years that you are likely to pay on your income drawdown product.'

- **Cost rating**—a rating summarising all charges from a product on a scale from £ (cheapest) to £££££ (most expensive)[28] shown as the summary cost metric, as well as the individual charges from the simple control. The cost rating was described as follows: 'The charges rating summarises all charges for the income drawdown product on a scale from £ (cheapest) to £££££ (most expensive). The scale provides a simple comparison of charges across different products on offer'.[29]

Figure 3.1 shows the price comparison website screen that participants saw, with the cost rating treatment shown.

[28] Cost rating based on reduction in yield over a 20-year period, with only one product marked as being in the cheapest cost category (£), products within 10 basis points of the cheapest reduction in yield marked as being in the second cheapest category (££), and further categories including subsequent groups of products in a range of 10 basis points, until the £££££ category, which includes all products more than 30 basis points more expensive than the cheapest product.

[29] There is evidence that simple rating metrics can be effective in both enabling consumer choice and encouraging better competitive outcomes. For example, the introduction of hygiene ratings in California found that consumers switched to higher rated restaurants and restaurants improved their hygiene. These findings have been replicated elsewhere. See Jin, G.Z. and Leslie, P. (2003), 'The effect of information on product quality: Evidence from restaurant hygiene grade cards', *The Quarterly Journal of Economics*, pp. 409–51. See also City of New York (2012), 'Restaurant Grading in New York City at 18 Months'; and da Cunha, D.T., de Freitas Saccol, A.L., Tondo, E.C., de Oliveira, A.B.A., Ginani, V.C., Araujo, C.V., Lima, T.A.S., de Castro, A.K.F. and Stedefeldt, E. (2016), 'Inspection Score and Grading System for Food Services in Brazil: The Results of a Food Safety Strategy to Reduce the Risk of Foodborne Diseases during the 2014 FIFA World Cup', *Frontiers in Microbiology*, **7**:614.

Figure 3.1 Choice screen, cost rating treatment

Please choose your income drawdown product

Pension pot £50,000 Annual drawdown 0%

Your options

Please select a product and click the 'Next' button.

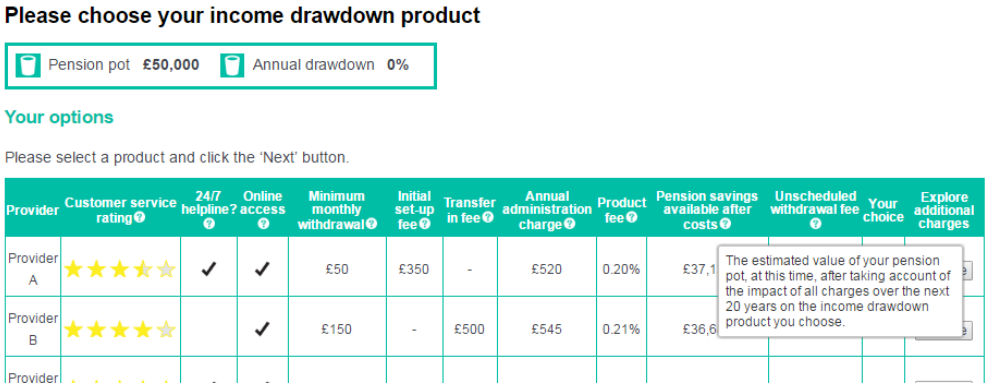
Provider	Customer service rating	24/7 helpline?	Online access	Minimum monthly withdrawal	Initial set-up fee	Transfer in fee	Annual administration charge	Product fee	Cost rating	Unscheduled withdrawal fee	Your choice	Explore additional charges
Provider A	★★★★½	✓	✓	£50	£350	-	£520	0.20%	£££	-	○	Explore
Provider B	★★★★☆		✓	£150	-	£500	£545	0.21%	££££	-	○	Explore
Provider C	★★★★½	✓	✓	£100	-	-	£125	0.43%	£	-	○	Explore
Provider D	★★½☆☆			£150	£200	-	£525	0.23%	££££	-	○	Explore
Provider E	★★★☆☆	✓	✓	£150	£150	-	£175	0.58%	££££	£250	○	Explore
Provider F	★★½☆☆			£125	-	-	£400	0.49%	£££££	-	○	Explore
Provider G	★★½☆☆	✓		£150	-	-	£300	0.51%	££££	£70	○	Explore
Provider H	★★★½☆		✓	£50	£100	-	£250	0.43%	£££	£25	○	Explore
Provider I	★★★★½			£50	£100	-	£375	0.26%	££	-	○	Explore
Provider J	★★½☆☆			£100	£350	£13	£225	0.47%	£££	£115	○	Explore
Provider K	★★☆☆☆	✓	✓	£125	£300	£20	£200	0.48%	£££	-	○	Explore
Provider L	★★★★½	✓	✓	£100	£110	£100	£475	0.31%	££££	-	○	Explore
Provider M	★★★☆☆		✓	£100	-	-	£450	0.29%	£££	-	○	Explore
Provider N	★★★★☆	✓	✓	£75	-	£210	£425	0.35%	££££	-	○	Explore
Provider O	★★★★☆	✓		£150	-	-	£540	0.39%	£££££	-	○	Explore
Provider P	★★☆☆☆	✓		£100	-	£110	£75	0.55%	££	-	○	Explore
Provider Q	★★½☆☆			£50	-	-	£150	0.60%	£££	-	○	Explore
Provider R	★★★★☆	✓	✓	£50	£100	-	£100	0.59%	£££	-	○	Explore
Provider	Customer service rating	24/7 helpline?	Online access	Minimum monthly withdrawal	Initial set-up fee	Transfer in fee	Annual administration charge	Product fee	Cost rating	Unscheduled withdrawal fee	Your choice	Explore additional charges

Note: a fund management fee of 0.76% is charged for the default investment fund in each product. Please note that the fund management fee you are charged may vary depending on your choice of investment funds.

Source: Oxera/CESS.

The explanatory text for each of the columns in the price comparison website appeared when the participant 'hovered' the mouse over the question mark symbols. Figure 3.2 shows the explanatory text for the pension savings available after costs metric.

Figure 3.2 Price comparison website screen with pension savings available after costs explanatory text shown



Source: Oxera/CESS.

3.4 Group assignment

Participants were randomly assigned to the control group or one of the other treatments using the block randomisation method (described in Appendix A1). This method first divides participants into sub-groups ('blocks') based on observable characteristics (e.g. age or gender). Then, within each block, participants are randomly assigned to each group.

This method ensures that any characteristics of the participants that might influence the outcome are accounted for. For example, if highly educated people are more likely to choose the cheapest product, block randomisation prevents this from skewing results because it prevents any one of the treatment groups containing too many highly educated people.

The analysis verified that the socio-demographic characteristics of each of the seven groups were similar. The results are reported in Appendix A1.

3.5 Data collected

Data was collected on a range of participant characteristics, as listed below. For statistics on data collected, see Appendix A1.

- **Demographics**—data was collected on each participant's age, gender, household income and education. Only participants aged 40+ were permitted to complete the experiment, in order to match the participant population as closely as possible to the population that will be considering retirement income in the near future.[30]

- **Understanding of income drawdown products**—at the beginning of the experiment, we asked participants about their background familiarity with, and understanding of, income drawdown products, before any information about these was given. We then explained income drawdown products and options for retirement income, and asked several comprehension questions to capture whether participants had understood the information that we had

[30] We did not impose an upper bound on participants' ages, to allow for a sufficiently large panel population from which to obtain participants. In addition, as people make decisions on their retirement income at different times (or at multiple points in time), an upper bound to participants' ages might have been inappropriate for other reasons. Finally, only 66 participants (3% of the sample) were aged above 75, mitigating the effect that including individuals who have already been retired for some time might have had on the results.

given them. We then collected data on most likely usage for an income drawdown product through the participant's choice of usage profile.

- **Selection of income drawdown product**—we recorded the income drawdown product that participants selected from the price comparison screen, which products participants viewed charge sheets for, and how long participants spent on this screen.

- **Product ownership and financial advice**—participants reported their experience with, and planned usage of, financial advisers, ownership of defined contribution (DC) and defined benefit (DB) pension pots, and plans for retirement. This involved asking the following questions:

  - have you previously used a financial adviser for pension or retirement advice?

  - do you plan to use a financial adviser when planning for retirement?

  Participants were then asked whether they had a DC pension pot and/or a DB pension pot, with both terms explained.[31]

- **Financial literacy**—information was collected on income and financial literacy (through stated responses on ability and understanding, as well as questions that tested participants' ability to perform simple financial calculations). This involved asking participants:

  - about their self-reported maths ability aged 10;

  - whether they agreed with the statement 'Financial services are complicated and confusing to me';

  - two questions that involved understanding compound interest rates;

  - a cognitive reflection test:[32] 'you buy a bat and a ball for £1.10. The bat costs £1 more than the ball. How much does the ball cost?'.[33]

- **Preferences and behavioural biases**—we collected responses on impulsivity, risk preference, time preference (and consistency of time preference), and life behaviours, which could be correlated with behavioural biases.[34]

3.6 Outcomes

It was made clear to participants at the start of the experiment that it was incentivised—there was a participation payment (£4.00 for each participant), and the potential for more payments depending on their responses.

Participants also received payment based on which income drawdown product they selected, up to a maximum of £3.00, with no payment made for this activity if a participant selected one of the six most expensive products, and other payment following the schedule in Table 3.3. Full payment for this activity was provided only to participants who selected the cheapest product.

[31] See Questions 13 to 16 in Appendix A2.
[32] This tests the extent to which participants are engaging System II or relying on System I.
[33] See Questions 17 to 21 in Appendix A2.
[34] See, for example, Barlow, P., McKee, M., Reeves, A., Galea, G. and Stuckler, D. (2016), 'Time-discounting and tobacco smoking: a systematic review and network analysis', *International Journal of Epidemiology*', November, https://doi.org/10.1093/ije/dyw233.

Table 3.3 Payment schedule for selection of income drawdown product

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13–18
Task earning (£)	3.00	2.75	2.50	2.25	2.00	1.75	1.50	1.25	1.00	0.75	0.50	0.25	0.00

Note: The ranking of the products was broadly the same for all treatments (Provider C was always the cheapest, etc.), but there were small differences in the ranking between reduction in yield and the other metrics (average cost, total cost, and pension savings available after costs). This is due to the way in which reduction in yield is calculated (compounding effects). Each treatment had a payoff function that corresponded to the relevant ranking. Both controls and the cost rating used the reduction in yield ranking.

Source: Oxera/CESS.

A linear payment schedule was chosen because it was simple and transparent for participants to understand.[35]

This schedule was constructed so as to preserve the incentives for all participants to select a cheaper product. The cheapest product was also the dominant product across all quality dimensions (no product had a better quality on any of the quality measures). Therefore, if a participant was considering a trade-off of price versus quality, this payment schedule created the incentive to choose the product that was just as good as another, but also cheaper.

Participants also received payment based on their response to the risk-preference question (Q22): where participants selected a risky gamble, payoffs were based on a realisation of a random variable based on the distribution of the participant's response. The payoff for the risk-preference question could be between zero and £2.16, with a median payoff of £0.72.

The whole experiment had a median payoff of £7.29, with a range of £4.00 to £9.16.

[35] We also considered a payment schedule that reflected the magnitude of the difference in cost between products. However, such a scheme would have (a) been complex for participants to understand; (b) been less transparent, in that the full payment schedule would not have been shown before the experiment (or it would have given information about the products); and (c) allowed a situation to arise where two products were very similarly rewarded, and thus a choice between those two products might not have been adequately incentivised.

4 Experiment results

4.1 Overview of product selection performance

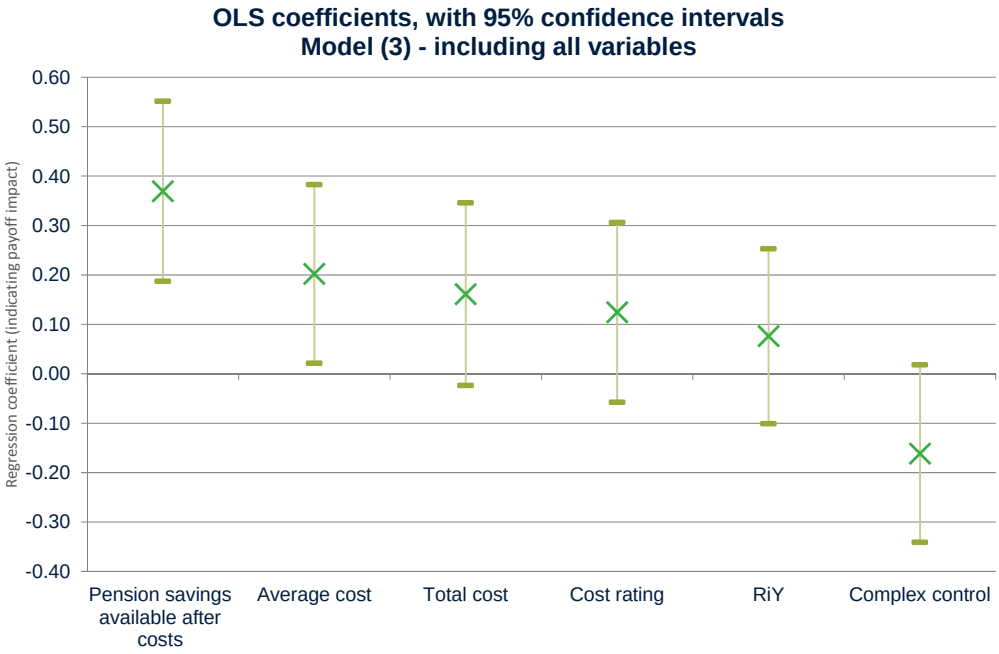
The key outputs from the experiment were the measures of the performance of the participants in selecting the income drawdown 'wrapper' product that best minimises costs. The measures considered were:

- the reward to the participant for their selection, based on the payment schedule (see section 3.6);
- whether or not the participant selected the cheapest product;
- whether or not the participant avoided the 13 more expensive products (which provided a product selection reward of less than £2).

The different measures of performance were assessed using regressions that controlled for other factors that could influence product selection performance. While the experimental approach involved allocating individuals to treatments in order to ensure a fair distribution in terms of key observable factors such as gender and education, there remained a risk that, by chance, the allocation of individuals was not as fair in terms of unobservable factors, such as motivation and financial ability. To ensure that these factors, which were measured by other questions (such as the 'bat and ball' question), did not influence the results, they were included in the regressions. Full details of the regression analysis can be found in Appendix A3.

The performance of each metric in terms of the product selection reward is presented in Figure 4.1. The ordinary least squares (OLS) coefficients of the regression represent the difference in payoff (in £) between the treatments relative to the complex control. The higher the coefficient, the higher the payoff, and therefore the more effective the treatment. For example, the effect of the pension savings available after costs treatment is plus £0.37 over the simple control (with 95% confidence intervals of plus £0.19 to plus £0.55).

Figure 4.1 OLS coefficients, with 95% confidence intervals



Note: The simple control is the baseline treatment. OLS regression coefficients from the linear model that includes all variables—Model (3). As shown, the pension savings available after costs and the average cost treatments are significantly better than the simple control at the 5% level. Full regression results are provided in Appendix A3. RiY, reduction in yield.

Source: Oxera.

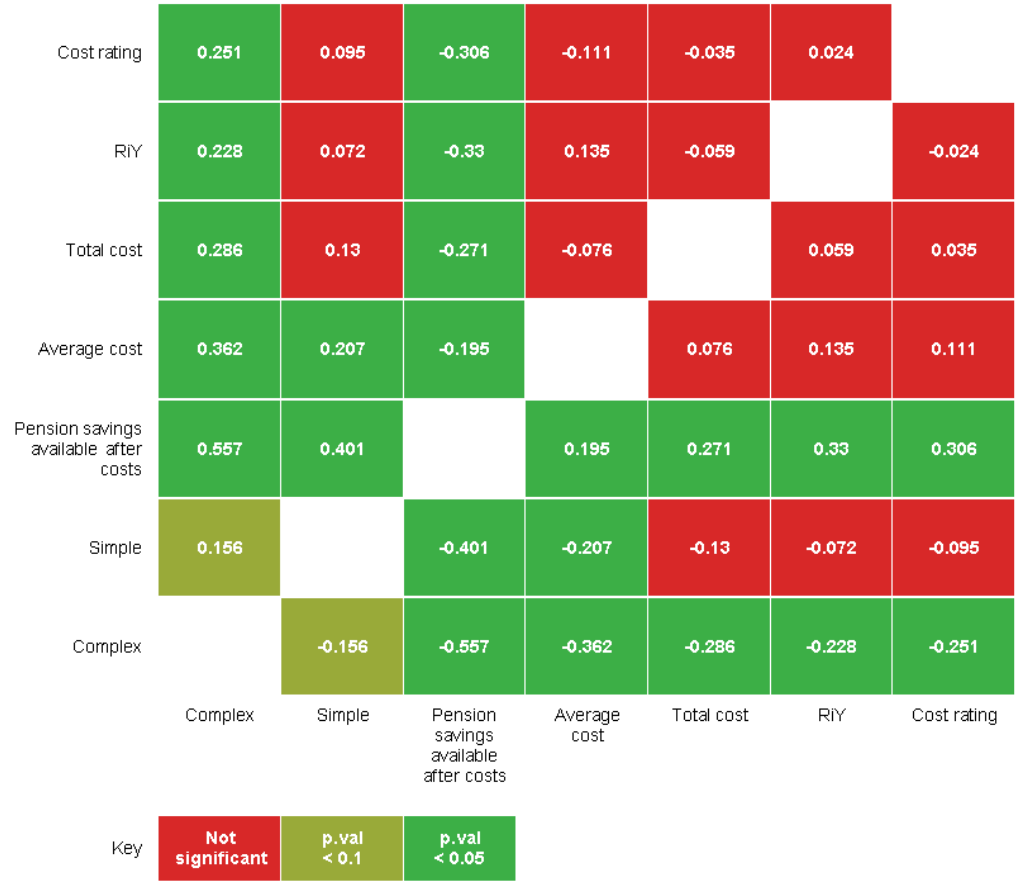
In terms of the *product selection reward*, based on the full sample, the pension savings available after costs metric performed significantly better (at the 95% significance level) than all of the other treatments. The average cost metric performed significantly better than the two control groups. The total cost, reduction in yield and cost rating metrics did not perform significantly better than the simple control.[36]

Figure 4.2 shows the differences in the mean payoff between treatments, and whether the differences are statistically significant. For example, the top left cell of the table shows that participants in the cost rating treatment received a mean payoff that was 25p higher than those in the complex control treatment. The colour of the top left cell indicates that it is statistically significant at the 5% level using the T-Test (this test on the OLS regression results takes into account all control variables and therefore isolates for the effect of the treatment).

As shown, the mean payoff for the participants who were in the pension savings available after costs treatment was higher than for all the other treatments (and this is significant at the 5% level for all six comparisons).

[36] Although they were statistically significantly better than the complex control, unlike the simple control.

Figure 4.2 Differences in mean payoff (£), and T-Test results

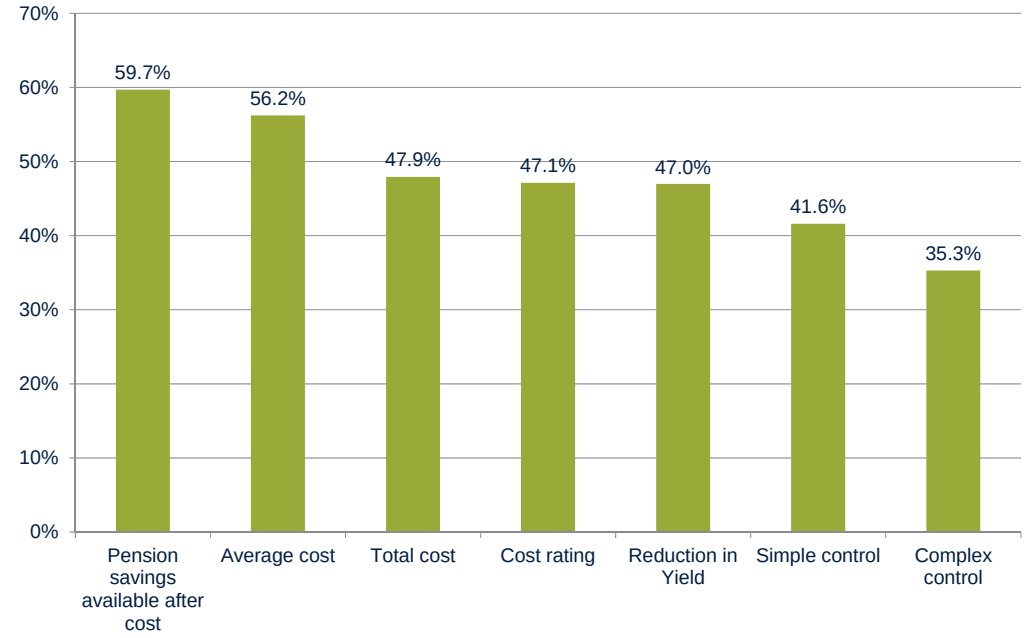


Note: Based on whole sample of 2,020. RiY, reduction in yield.

Source: Oxera/CESS.

Similar results were found in terms of the proportion of participants who chose the cheapest product. On average across the entire sample, 47% chose product C, the cheapest product. With the pension savings available after costs metric, 59.7% of participants chose the cheapest product, and 56.2% with average cost, and both of these result were statistically significant versus the simple control. However, pension savings available after costs' small outperformance was not statistically significant versus average cost on this measure. At the other end of the scale, 35.3% of participants chose the cheapest product when presented with the complex control treatment. This is shown in Figure 4.3.

Figure 4.3 Proportion of participants who chose the cheapest product

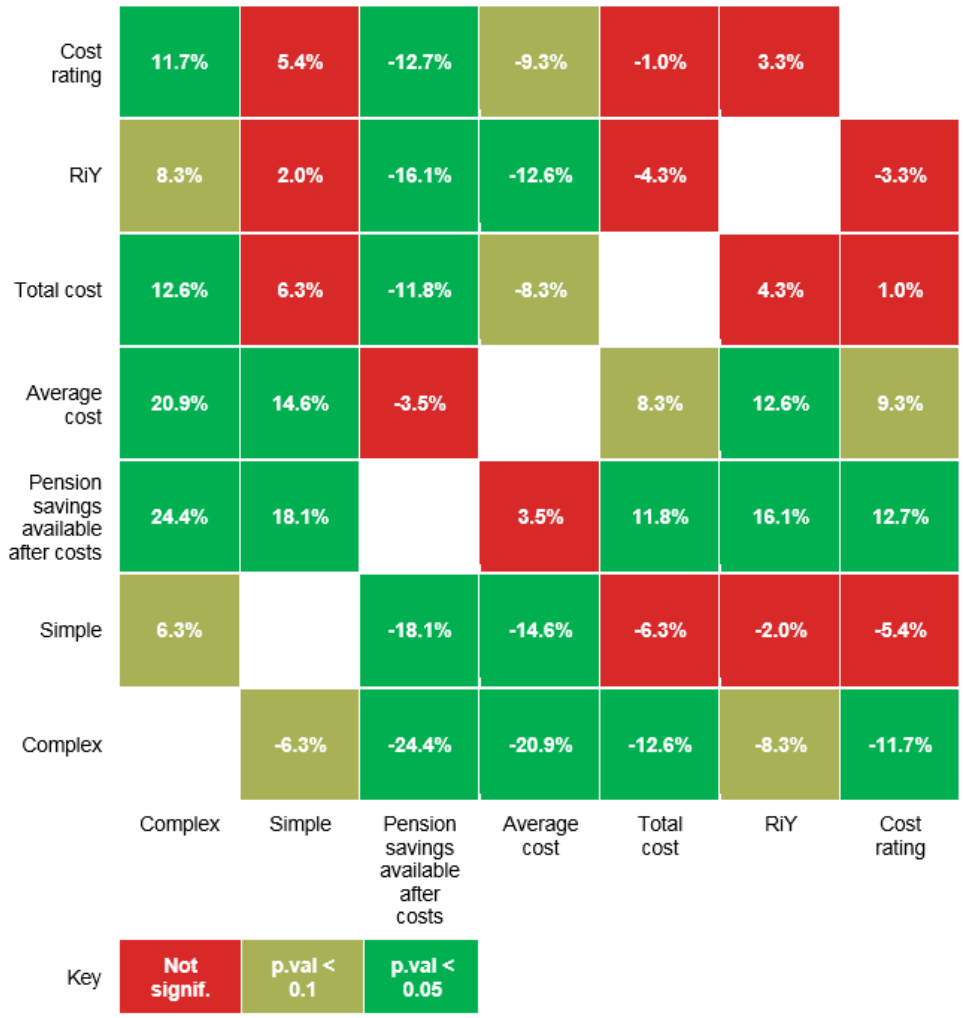


Notes: Based on whole sample of 2,020 participants.

Source: Oxera/CESS.

Figure 4.4 gives the differences between treatments in terms of the percentage of participants who chose the cheapest product. For example, the top left cell of the table shows that the percentage of participants who chose the cheapest product was 11.6% higher in the cost rating treatment than in the complex control. The colour of the top left cell indicates that this difference is significant at the 5% significance level using the Z Test (this test on the logistic regression results takes into account the control variables and therefore isolates for the effect of the treatment).

Figure 4.4 Differences in percentage of participants who chose the cheapest product, and Z-Test results

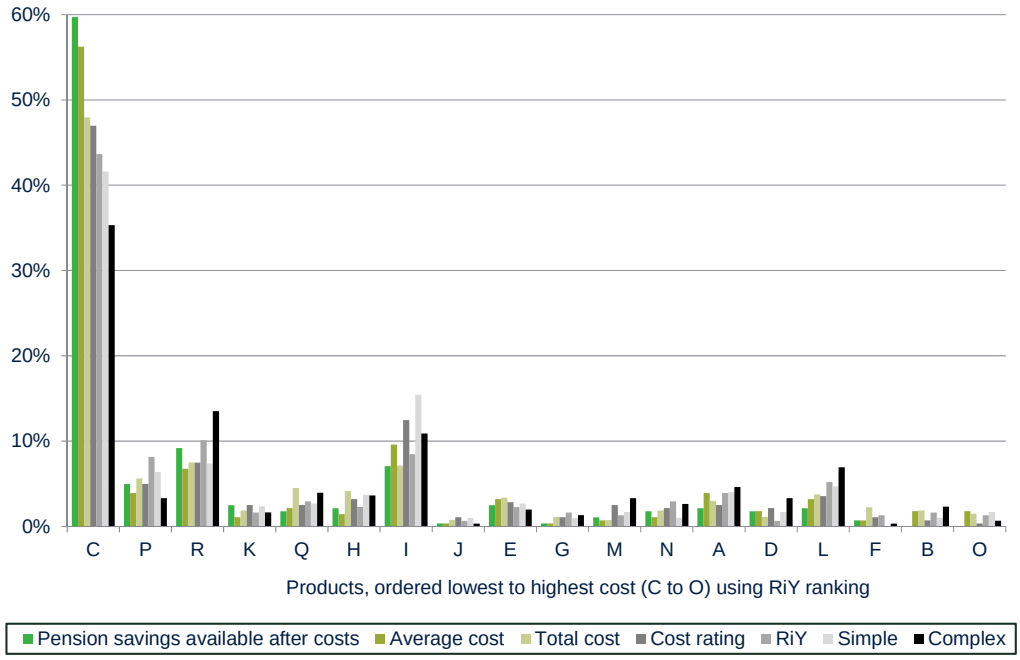


Note: Based on whole sample of 2,020. RiY, reduction in yield.

Source: Oxera/CESS.

There were similar trends across the treatments in terms of the products selected. Figure 4.5 shows the proportions of participants who selected each product, for each treatment. The figure presents the products in order of cost, with the cheapest on the left (product C) and the most expensive on the right (product O). While there is a general trend for participants to choose the cheapest products, it is also clear that some products were more popular than their cost might suggest (e.g. products I and L), and some products were less popular than their cost might suggest (e.g. products K and J).

Figure 4.5 Product choice by treatment



Note: Whole sample of 2,020 participants. Ordering of products done using reduction in yield cost ranking. The cost ranking is slightly different to this for three treatments (pension savings available after costs, average cost, and total cost (20 years)) due to compounding, but that does not affect the overall pattern shown in the chart. RiY, reduction in yield.

Source: Oxera/CESS.

Similar results were also found in terms of participants avoiding the more expensive products, which is also apparent from Figure 4.5.

4.2 Results by profile and demographics

Profile 1 was chosen by 50.5% of participants, while 49.5% chose profile 2. The choice of profile was not found to have a large impact on the results, as shown in Table 4.1. In both profiles the proportion of participants who chose the cheapest product was 47%, and the ordering of the two best treatments stayed the same.

Table 4.1 Participants who chose the cheapest product, by treatment and profile

	Profile 1	Profile 2	Overall
Pension savings available after costs	61.5%	57.8%	59.7%
Average cost	58.0%	53.8%	56.2%
Total cost (20 years)	46.3%	49.6%	47.9%
Cost rating	44.0%	50.0%	47.0%
Reduction in yield	44.9%	42.7%	43.6%
Simple control	40.0%	43.2%	41.6%
Complex control	34.9%	35.7%	35.3%
Overall	47.3%	47.0%	47.1%

Source: Oxera/CESS.

There was no clear trend in terms of age or gender, as shown in the following two tables. The performance of younger versus older people, or women versus men, was not found to be statistically significantly different.

Table 4.2 Mean payoff by age (£)

	Aged <58	Aged 58+	Overall
Pension savings available after costs	2.34	2.46	2.40
Average cost	2.25	2.17	2.21
Total cost (20 years)	2.14	2.12	2.13
Cost rating	2.09	2.10	2.10
Reduction in yield	1.94	2.19	2.07
Simple control	2.03	1.98	2.00
Complex control	1.72	1.95	1.85
Overall	2.07	2.13	2.10

Note: Median age was 58 years old. Mean age was 58.5 years old.

Source: Oxera/CESS.

Table 4.3 Mean payoff by gender (£)

	Male	Female	Overall
Pension savings available after costs	2.44	2.36	2.40
Average cost	2.26	2.15	2.21
Total cost (20 years)	2.23	2.03	2.13
Cost rating	2.04	2.17	2.10
Reduction in yield	1.99	2.15	2.07
Simple control	1.83	2.19	2.00
Complex control	1.86	1.83	1.85
Overall	2.09	2.12	2.10

Source: Oxera/CESS.

Education was, however, found to be statistically correlated with performance in the experiment. Those with a higher level of education were more likely to earn a higher reward. Pension savings available after costs remained the treatment that produced the higher payoff in each education bracket, as shown in the following table.

Table 4.4 Mean payoff by education (£)

	No qualification or secondary qualification	Post-secondary qualification	University qualification	Overall
Pension savings available after costs	2.24	2.49	2.50	2.40
Average cost	2.00	2.23	2.43	2.21
Total cost (20 years)	1.87	2.03	2.48	2.13
Cost rating	1.91	2.13	2.33	2.10
Reduction in yield	1.88	2.28	2.16	2.07
Simple control	2.07	1.97	1.93	2.00
Complex control	1.82	1.89	1.83	1.85
Overall	1.97	2.14	2.22	2.10

Note: No qualification and secondary qualification were combined, due to the small sample size of those with no qualification.

Source: Oxera/CESS.

4.3 Results by financial understanding, risk preferences and behaviour metrics

The econometric analysis found that the following metrics were statistically significantly correlated with performance in product selection:

- those who correctly answered the 'bat and ball' question (Q21), which was used as a measure of cognitive reflection;
- those who correctly calculated the compound interest (Q18 and Q19), which tested financial literacy;
- those who revealed a lower implicit discount rate (Q25–Q30);
- those who stated a greater degree of confidence in their ability to understand financial services (Q17).

In addition, the time spent by the participant completing the experiment was found to be significantly correlated with performance measures. The longer they spent on the experiment, the better they did on average. Time was therefore included as part of the robustness checks, as described in section 4.4.1 below.

While these metrics were positively correlated with product selection performance, the performance of all participants was improved by the summary cost metrics.

However, the following metrics were not statistically significantly correlated with performance in product selection:

- those who were planning for retirement (Q23);
- those who were in retirement (Q23);
- those who had used a financial adviser (Q13).

Again, the performance of all participants was improved by the summary cost metrics, irrespective of their answers to these questions.

Full details of the econometric results can be found in Appendix A3. Screenshots of each question can be found in Appendix A2.

4.4 Robustness checks

The robustness of the results was tested through a wide range of econometric analyses, including different model specifications for the OLS regressions on reward and the Logit regressions for the likelihood of picking the cheapest product, or avoiding the more expensive products. The results are presented in Appendix A3.

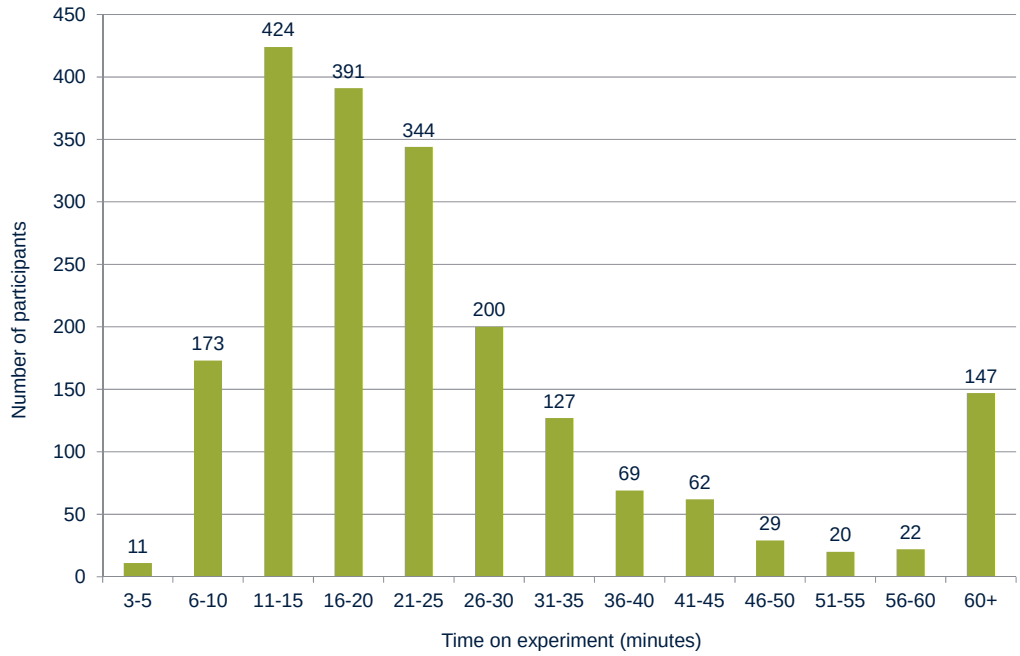
The ordering of the treatments was robust to these different model specifications. Pension savings available after costs was the most successful treatment in all cases, followed by average cost. The other treatments were less successful in all cases.

4.4.1 Excluding those who took less time on the experiment

Based on testing by Oxera team members, the view was taken that it was not plausible that those who spent ten minutes or less on the experiment had genuinely thought about it and made a considered decision on the price comparison website task. Excluding those who spent ten minutes or less on the

experiment cuts the sample by 9.1% (184 people). The distribution of participants by time spent in the experiment is shown in Figure 4.6.

Figure 4.6 Time spent on experiment



Source: Oxera/CESS.

This time cut-off of ten minutes does not have a material impact on the econometric results, as shown in Table 4.5.

Table 4.5 OLS regression results for sample excluding those who spent ten minutes or less on the experiment

	Coefficient	95% confidence intervals
Pension savings available after costs	0.522***	0.335–0.710
Average cost	0.445***	0.257–0.633
Total cost (20 years)	0.366***	0.175–0.557
Cost rating	0.329***	0.140–0.518
Reduction in yield	0.275***	0.092–0.458
Simple control	0.199**	0.012–0.387

Note: * Indicates significance at the 10% level. ** Indicates significance at the 5% level. *** Indicates significance at the 1% level.

Source: Oxera.

However, there is some variability in the proportion of participants who chose the best product. For example, while pension savings available after costs is still the best metric, it has a very similar proportion getting the best choice to average cost when the sample excludes those spending ten minutes or less on the experiment. This is shown in Figure 4.7 below.

In interpreting this result, we note that:

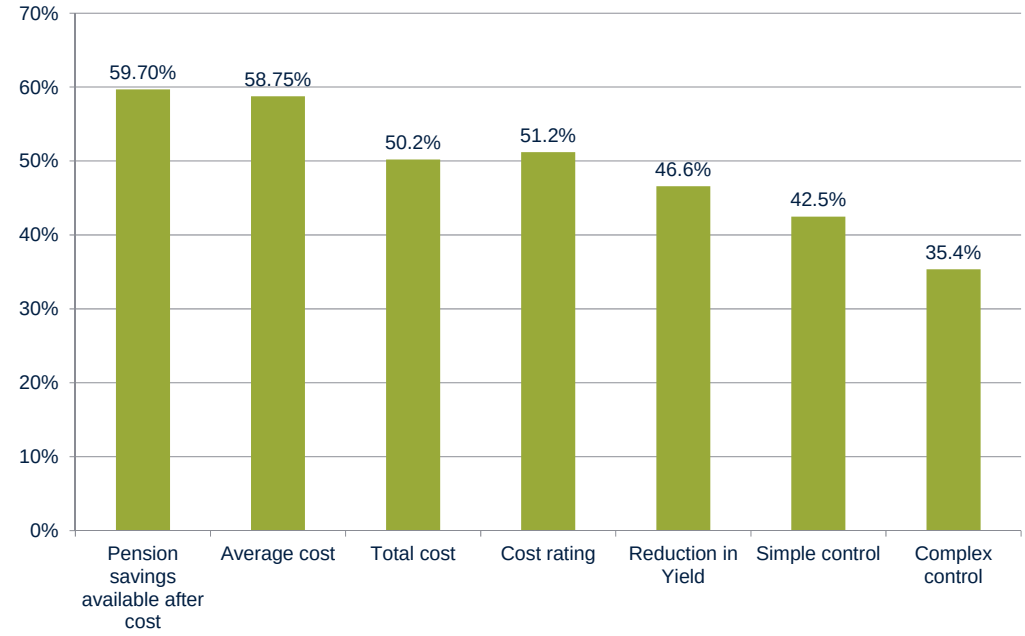
- pension savings available after costs is still the best metric;
- different time cut-off points (for example, excluding those who spent five minutes or less on the experiment) result in different gaps between the

metrics. This shows that the results are sensitive to the choice of time cut-off point, a choice that could be seen as arbitrary;

- the econometric analysis clearly shows that pension savings available after costs is the best metric. The econometric analysis is the most robust basis on which to form policy because it takes account of all other factors (e.g. age, education, financial literacy) and therefore isolates the impact of the treatment more clearly. The econometrics takes account of all the choices over all the products, rather than just whether the best product was chosen.

We therefore conclude that, while the magnitude of the difference between pension savings available after costs is smaller in this chart, it does not undermine the result that pension savings available after costs is the best metric.

Figure 4.7 Proportion of participants who chose the cheapest product sample: those who spent over 10 minutes in the experiment



Note: Sample limited to those who spent over ten minutes on the experiment (1,836 participants).

Source: Oxera/CESS.

5 Interpretation and conclusions

This study tested five summary cost metrics for income drawdown products through an online experiment. The experimental results are clear: two of the summary metrics, pension savings available after costs and the average cost metric, caused a statistically significant improvement in the selection of the cheaper drawdown product versus the simple control. The total cost, reduction in yield and cost rating metrics did not have a statistically significant effect.

When considering the monetary reward payoff measure, pension savings available after costs was more effective than average cost per year, and this difference was statistically significant (see Figure 4.2). With the pension savings available after costs metric, 59.7% of participants chose the cheapest product, while with the average cost per year metric 56.2% did so. However, the difference in performance between the metrics was not statistically significant (see Figure 4.4).

5.1 Why might pension savings available after costs and average cost be the best metrics?

Why might the pension savings available after costs and average cost metrics be better than reduction in yield?

- **They are presented as a £ rather than a % figure**. Many people find it hard to understand percentages. A cost metric that uses % figures will increase the chances of the consumer making a mistake or being discouraged from engaging with the metric.[37] In addition, the levels of % figures are often lower and therefore less salient than £ figures (i.e. reduction in yield is less likely to be as salient with figures that vary from 1.37% to 2.01%).

Why might these metrics be better than cost rating?

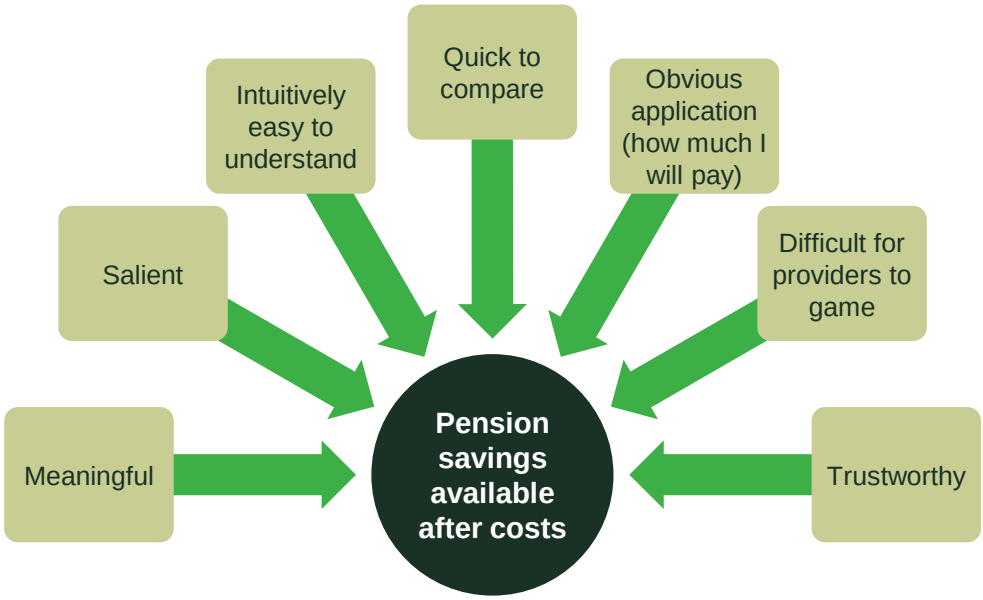
- **They are trustworthy**. People may not trust cost rating to truly reflect and incorporate all the fees relating to a product, but the other treatments, including pension savings available after cost, do not face this problem. The simplicity of cost rating may reduce trust in its comprehensiveness or accuracy. An alternative explanation is that consumers have learned from other contexts that simple cost ratings are not to be trusted, or are easily manipulated.

- **They are specific**. Although the usage of the product is approximated (e.g. there is a certain number of ad hoc withdrawals in each profile), the metric is otherwise directly relevant and specific for the consumer viewing it.

These reasons are summarised in Figure 5.1.

[37] Banks, J. and Oldfield, Z. (2006), 'Understanding pensions: Cognitive function, numerical ability and retirement saving', Institute for Fiscal Studies Working Paper WP06/05, Final version February 2006, http://discovery.ucl.ac.uk/2690/1/2690.pdf, accessed 3 January 2017.

Figure 5.1 Why pension savings available after costs and average cost might be the best metrics



Source: Oxera.

5.2 Why did some products perform better than others, apart from due to differences in cost?

As noted in Figure 4.4 above, some products were more popular than their cost might suggest (e.g. products I and L), and some products were less popular than their cost might suggest (e.g. products K and J). These trends were common across treatments, except that the more effective treatments reduced the extent of these observed deviations from the cost ordering.

The reasons for these deviations were not directly explored in the experiment, but again hypotheses can be put forward. For example, Provider I is consistently a popular choice. This might be because Provider I is in the group of products with lower product fees (below 0.3%), and has the lowest annual administration fee (at £375) of those products; but product I is not actually very cheap compared with other products that have product fees above 0.3%. Participants may be using a shortcut to estimate total costs by placing greater emphasis on the provider fee rate, which leads them to a sub-optimal choice.

5.3 Conclusions

The research finds that two of the cost metrics were effective at encouraging participants to select lower-cost products in the simplified scenario presented to them. In reality, consumers need to make choices along many dimensions, with cost being one of them, but the research does suggest that these summary cost metrics can add value through improved product selection.

While various potential reasons for the outperformance can be proposed, ultimately the strength of metrics in encouraging desired behaviours is an empirical question, and hence the key motivation for the behavioural experiment.

The introduction of one of these summary cost metrics may make it easier for consumers to choose the cheapest income drawdown product. It may also facilitate the creation and functionality of income drawdown price comparison websites—a single cost figure would be available for each product, and products could therefore be ranked by cost.

A1 Experimental methods

A1.1 Recruitment methodology

The experiment recruited 2,020 participants from the UK population of people aged 40+. This age bracket, along with the residency requirement, were the only variables that qualified potential subjects for participation in the experiment. The experimental subjects were recruited by Respondi, a large online panel in the UK with 45,000 registered subjects, 5,000 of whom are in the 40–65 age bracket.[38]

Respondi uses many types of invitation to bring in people with diverse motivations to take part in research. These include email invitations, text messages, telephone alerts, and banners and messaging on websites and in online communities. The messages themselves are also varied, and include invitations to 'give your opinion' or 'let your voice be heard'. This diversity of motivations is likely to have contributed to a high-quality sample. To avoid self-selection bias, specific project details were not included in the invitation. Rather, participants were invited to 'take a survey', with the details disclosed later.

A1.2 Group assignment

Participants were assigned to treatment and control groups using block randomisation. This method first divides participants into sub-groups ('blocks') based on observable characteristics (e.g. their age or gender). Then, within each block, participants are randomly assigned to treatment and control groups.

This method ensured that any characteristics of the participants that might influence the outcome were accounted for. For example, if highly educated people are more likely to choose the cheapest product, the objective was to avoid any one of the treatment groups containing too many highly educated people.

The blocking variables used for the experiment were gender, education (high and low) and income (high and low). There were two categories in each group for a total of eight blocks (e.g. block 1 would be male, high income, high education). Information on each of these variables was collected before the participant started the experiment.

Once the participant had submitted this information, they were assigned to the appropriate block and then randomly assigned to one of the treatment or control groups. The probability of assignment into each treatment group varied according to the number of participants within each block who had been assigned to that group. This was done so as to maintain a balance across all treatment groups within each block. For example, if there were too many men in the personalised quote comparison group relative to the other groups, the probability of being assigned to that group was reduced for the next male participant. The likelihood of this occurring was typically 1/3 or 1/4. This method is called 'biased coin'.

To test whether the groups were balanced, Table A1.1 shows the breakdown in the three blocking variables across the six treatment groups.

[38] Respondi is an ISO 26362-certified survey company.

A1.3 Sample characteristics

Participants were recruited so as to be representative of the UK population of people over 40 in terms of gender, education and income. The socio-demographic characteristics of the sample are shown in Table A1.1. As can be seen, the characteristics of those in each treatment were broadly the same (due to the group assignment process, as detailed above).

The age distribution and socio-demographic characteristics of the sample are also shown in Figure A1.1 and Figure A1.2 below.

Table A1.1 Socio-demographic characteristics of participants across treatment groups

	Simple control	Complex control	Reduction in yield	Pension savings available after costs	Average cost	Total cost (20 years)	Cost rating
Gender*	1.483	1.482	1.515	1.452	1.495	1.502	1.445
	(0.501)	(0.5)	(0.501)	(0.499)	(0.501)	(0.501)	(0.498)
Income**	2.513	2.601	2.508	2.572	2.448	2.536	2.466
	(0.953)	(0.978)	(0.934)	(0.909)	(0.952)	(0.943)	(0.902)
Education***	2.839	2.993	2.879	2.961	2.893	2.929	2.811
	(0.954)	(0.959)	(0.971)	(0.935)	(0.931)	(0.957)	(0.896)
Observations	298	303	307	283	281	267	281

Note: Standard deviations in parentheses. *1 = Male; 2 = Female. **1 = <£12,000; 2 = £12,000–£24,999; 3 = £25,000–£49,999; 4 = >£50,000. ***1 = No qualification; 2 = Secondary education; 3 = Post-secondary education; 4 = University degree.

Source: Oxera/CESS.

According to the 2011 UK census, out of those aged 40–89, 52% were female.[39] The experiment sample was 48% female. Also according to the 2011 UK census, 53% of those aged 65+ did not have any qualifications (25% for those aged 50–64).[40] However, the experiment sample had only 6% without any qualifications: the sample is over-educated relative to the population (although we note that those choosing income drawdown products may also be over-educated relative to the general population).

Table A1.2 shows the income distribution of the age groups in question in the UK according to the 2016 Annual Survey of Hours and Earnings.[41] As can be seen, while the sample is not unrepresentative of the population, it has a slightly higher income distribution than the population.

[39] Office for National Statistics (2012), '2011 Census, Population and Household Estimates for the United Kingdom', December, http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-270247.
[40] Office for National Statistics (2014), 'Over 4 in 10 People Aged 25 to 34 had a Degree Level or Above Qualification', 7 March, http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/rel/census/2011-census-analysis/local-area-analysis-of-qualifications-across-england-and-wales/sty-qualification-levels.html.
[41] Office for National Statistics (2016), 'Dataset: Age Group - ASHE: Table 6, provisional', 26 October, https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/agegroupashetable6.

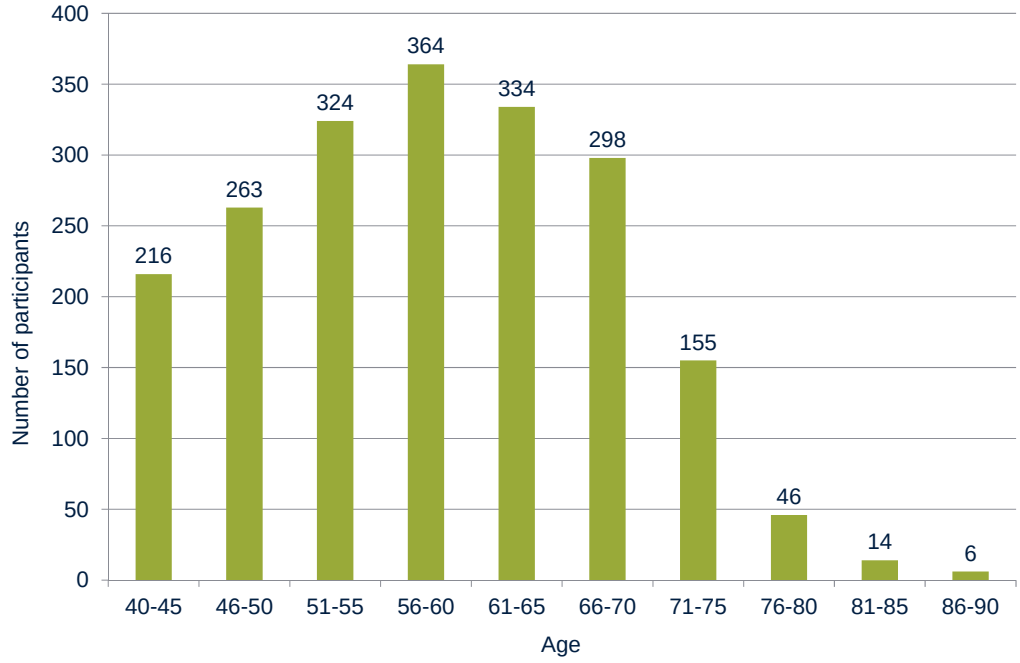
Table A1.2 Distribution of earnings by age group

	<£12,000	£12,000–£24,999	£24,99–£50,000	>£50,000
Population aged 40–49 and 50–59	10–20%	25–35%	30–40%	10–20%
Population aged 60+	30–40%	30–40%	25–35%	<10%
The experiment sample	16%	32%	37%	16%

Source: Office for National Statistics, 'The 2016 Annual Survey of Hours and Earnings', Oxera/CESS.

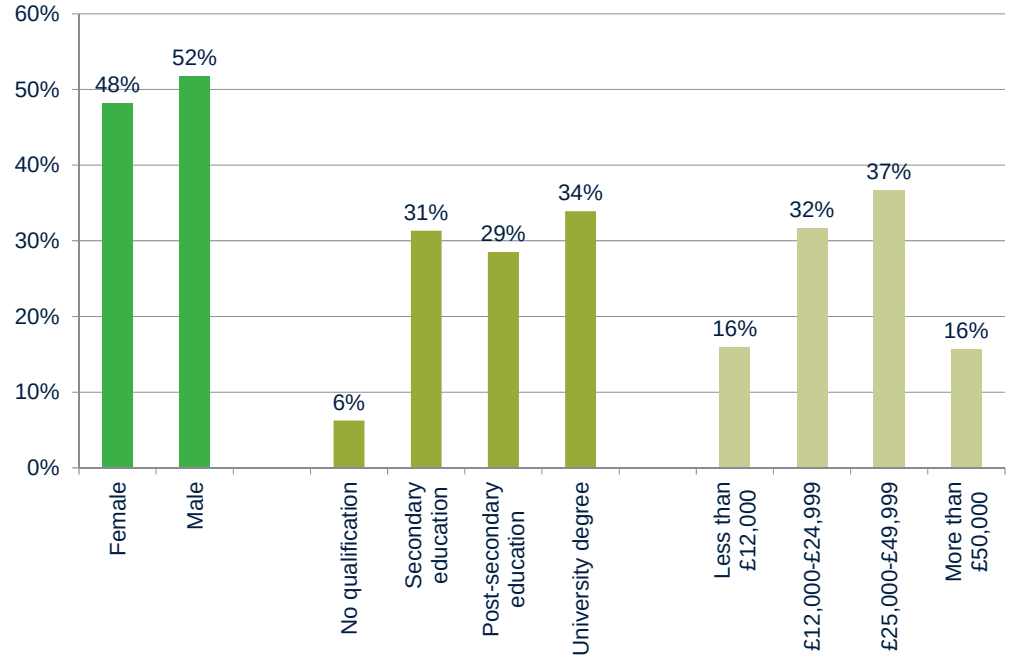
Bearing this in mind, the analysis employed methods to control for socio-demographic characteristics: multivariate regressions controlled for a number of demographic characteristics to test the robustness of the results.

Figure A1.1 Age distribution of the sample



Source: Oxera/CESS.

Figure A1.2 Socio-demographic characteristics of the sample

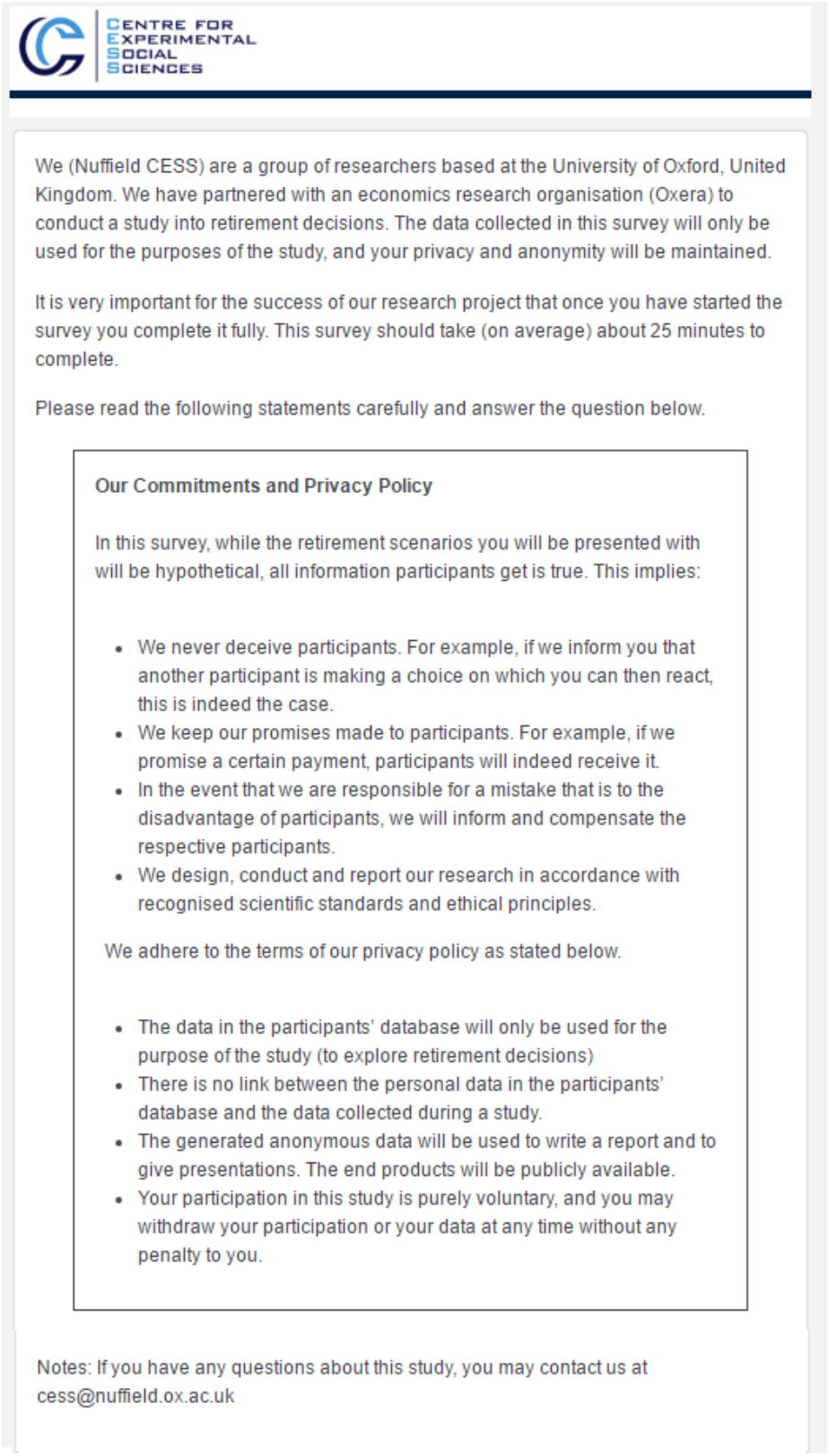


Source: Oxera/CESS.

A2 Experiment materials

Below are screenshots of the whole experiment. We have assigned question numbers to each question for ease of reporting, but these were not seen by participants. Each figure is a different screen, and all the content within each figure was within a single screen.

Figure A2.1 Introduction screen and Question 1



CENTRE FOR EXPERIMENTAL SOCIAL SCIENCES

We (Nuffield CESS) are a group of researchers based at the University of Oxford, United Kingdom. We have partnered with an economics research organisation (Oxera) to conduct a study into retirement decisions. The data collected in this survey will only be used for the purposes of the study, and your privacy and anonymity will be maintained.

It is very important for the success of our research project that once you have started the survey you complete it fully. This survey should take (on average) about 25 minutes to complete.

Please read the following statements carefully and answer the question below.

Our Commitments and Privacy Policy

In this survey, while the retirement scenarios you will be presented with will be hypothetical, all information participants get is true. This implies:

- We never deceive participants. For example, if we inform you that another participant is making a choice on which you can then react, this is indeed the case.
- We keep our promises made to participants. For example, if we promise a certain payment, participants will indeed receive it.
- In the event that we are responsible for a mistake that is to the disadvantage of participants, we will inform and compensate the respective participants.
- We design, conduct and report our research in accordance with recognised scientific standards and ethical principles.

We adhere to the terms of our privacy policy as stated below.

- The data in the participants' database will only be used for the purpose of the study (to explore retirement decisions)
- There is no link between the personal data in the participants' database and the data collected during a study.
- The generated anonymous data will be used to write a report and to give presentations. The end products will be publicly available.
- Your participation in this study is purely voluntary, and you may withdraw your participation or your data at any time without any penalty to you.

Notes: If you have any questions about this study, you may contact us at cess@nuffield.ox.ac.uk

Notes: If you have any questions about this study, you may contact us at cess@nuffield.ox.ac.uk

Please confirm that you have read and understood the rules stated above.

- ○ Yes, I have read and understood the rules
- ○ No

Figure A2.2 Welcome and Your earnings screen

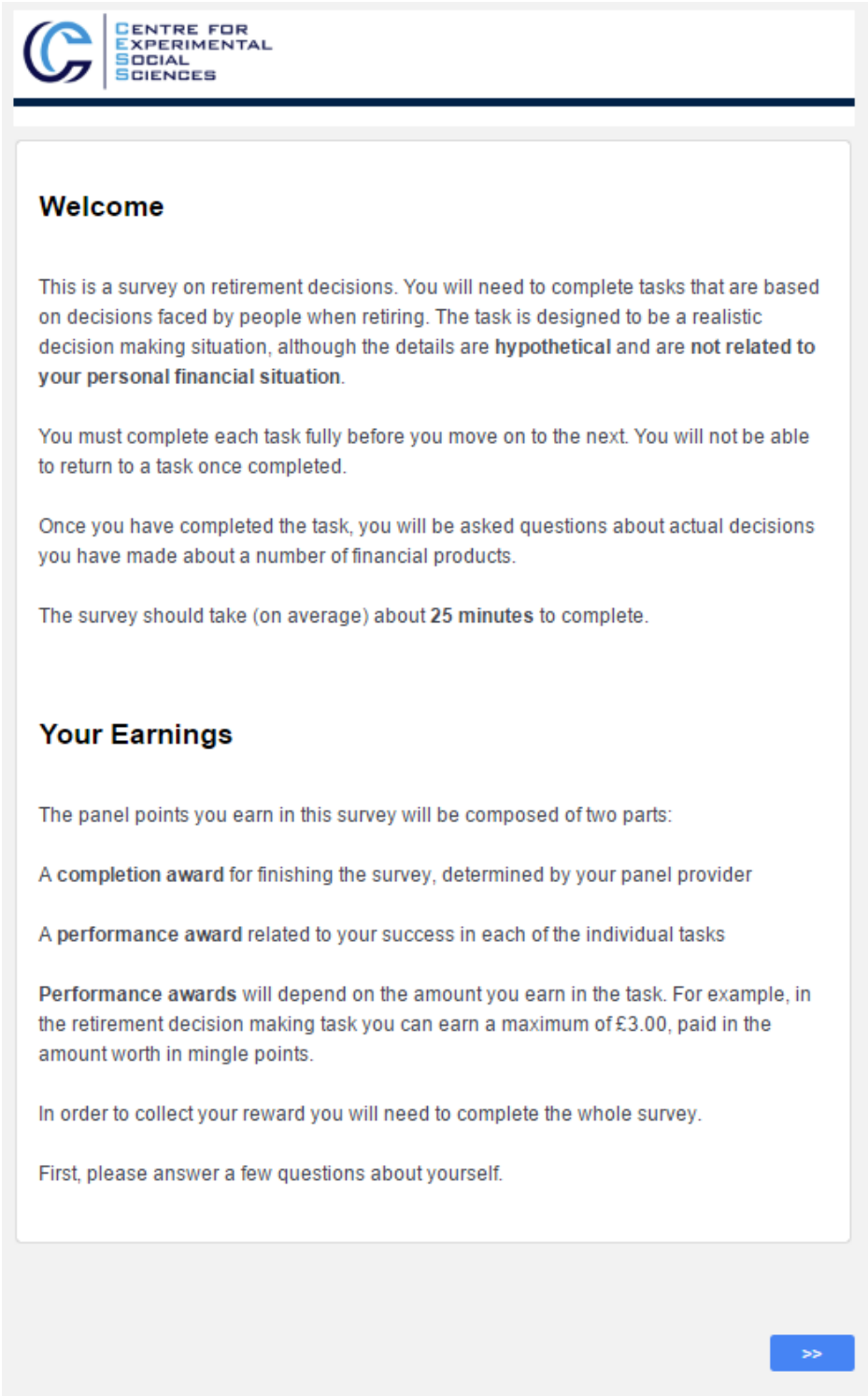


Figure A2.3 Questions 1, 2, 3 and 4

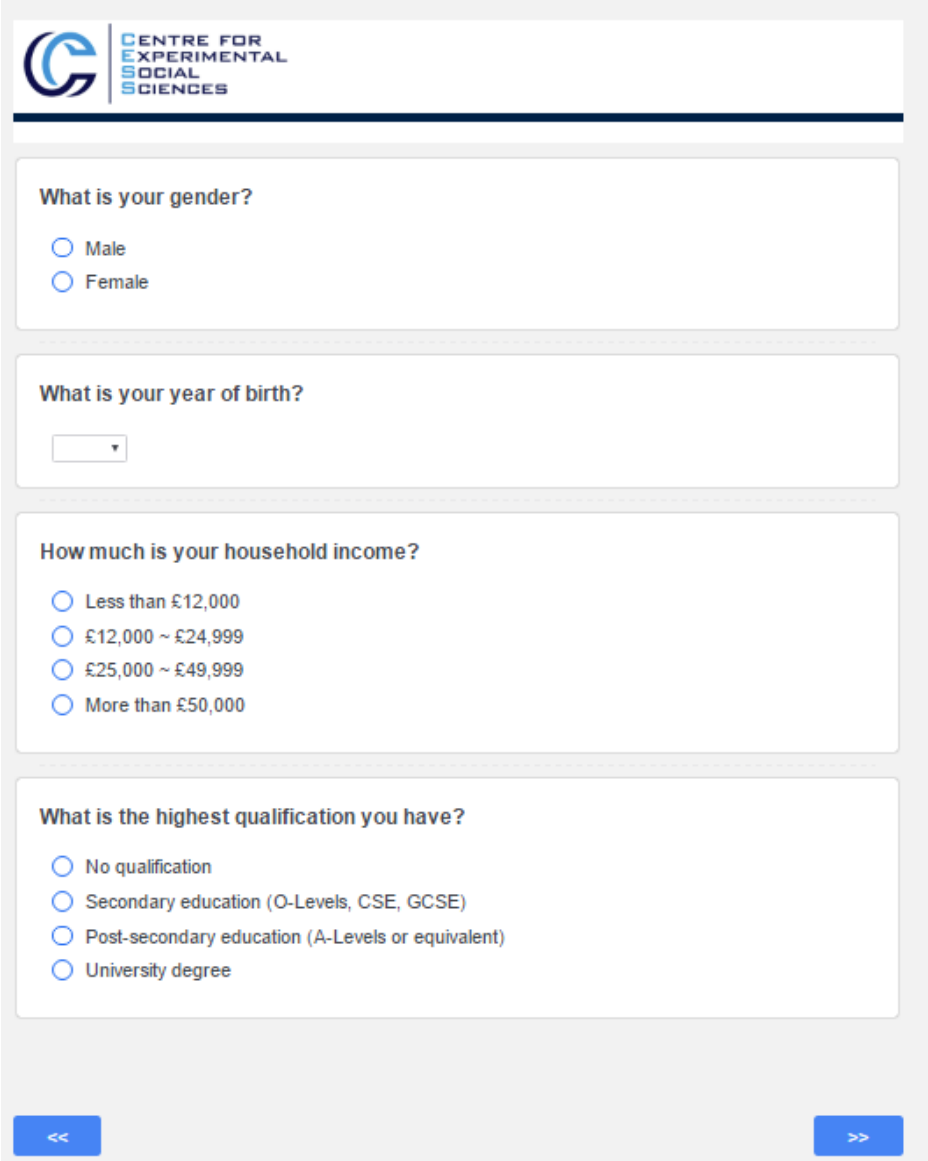


Figure A2.4 Questions 5, 6 and 7

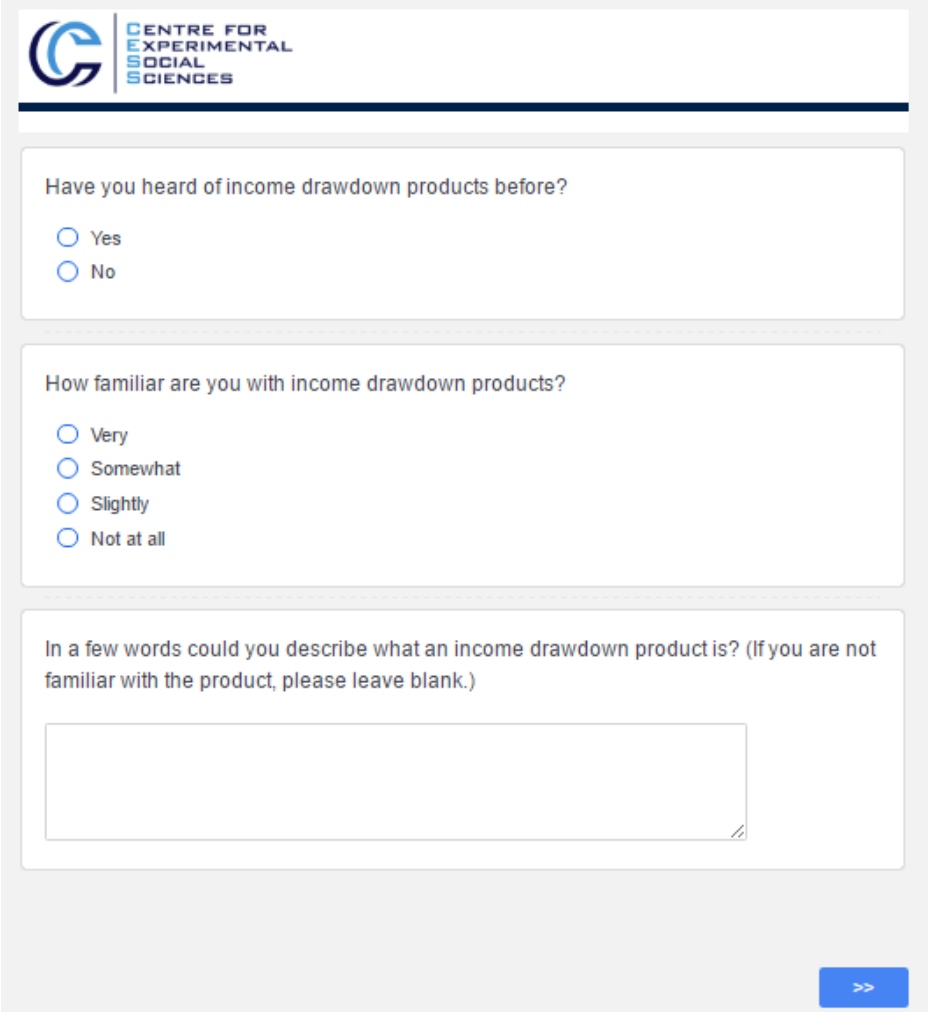


Figure A2.5 Retirement explanation screen

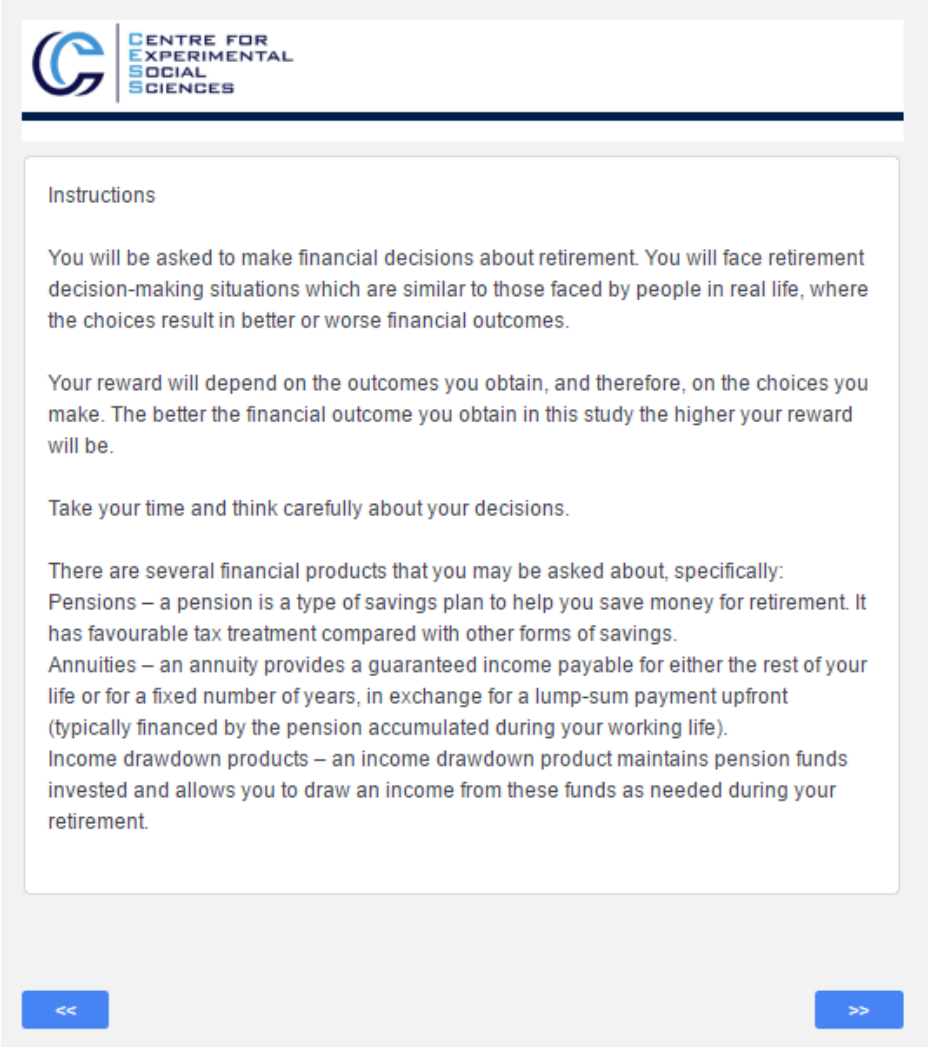


Figure A2.6 Income drawdown explanation screen

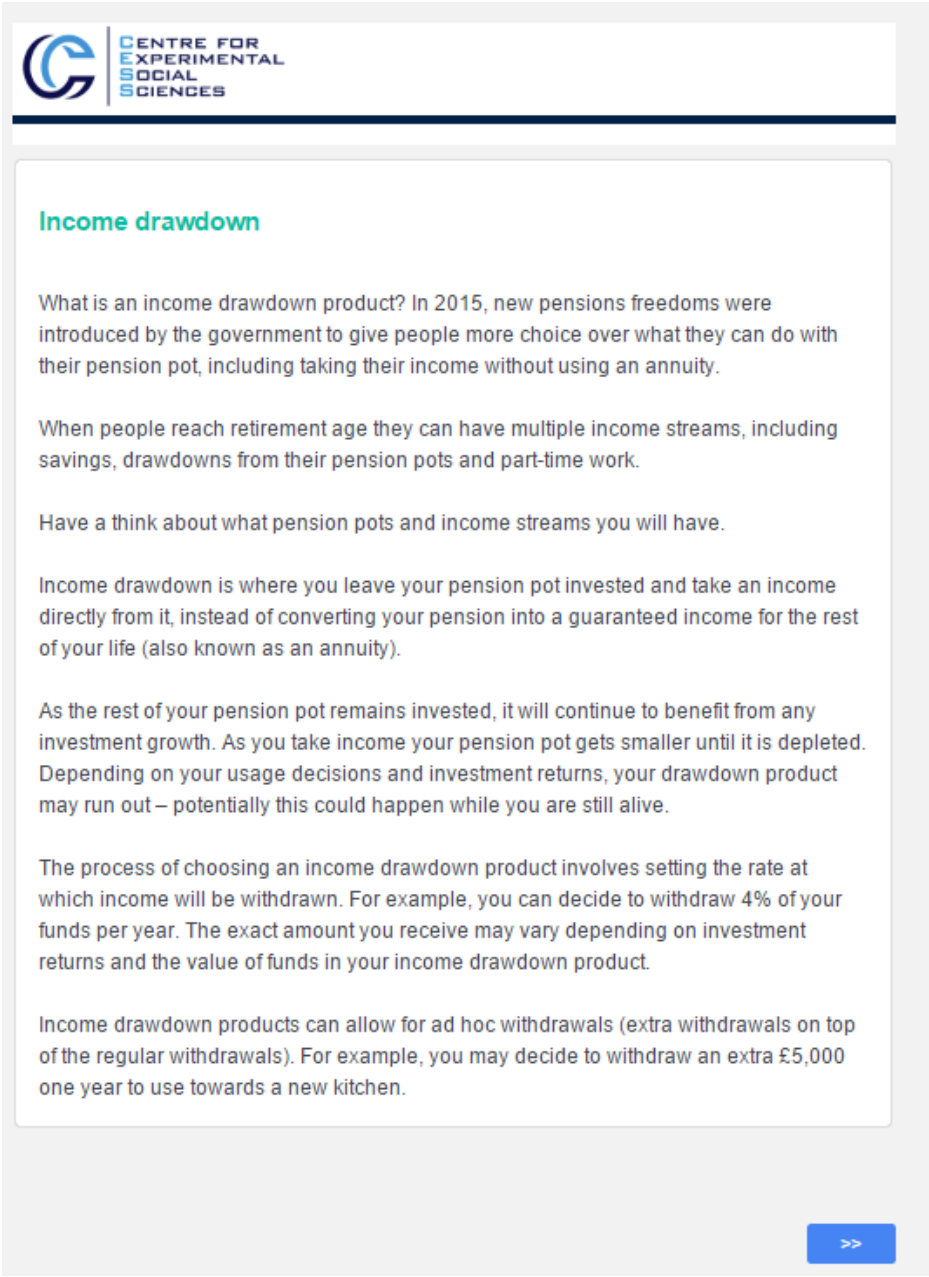


Figure A2.7 Income drawdown decision factors screen

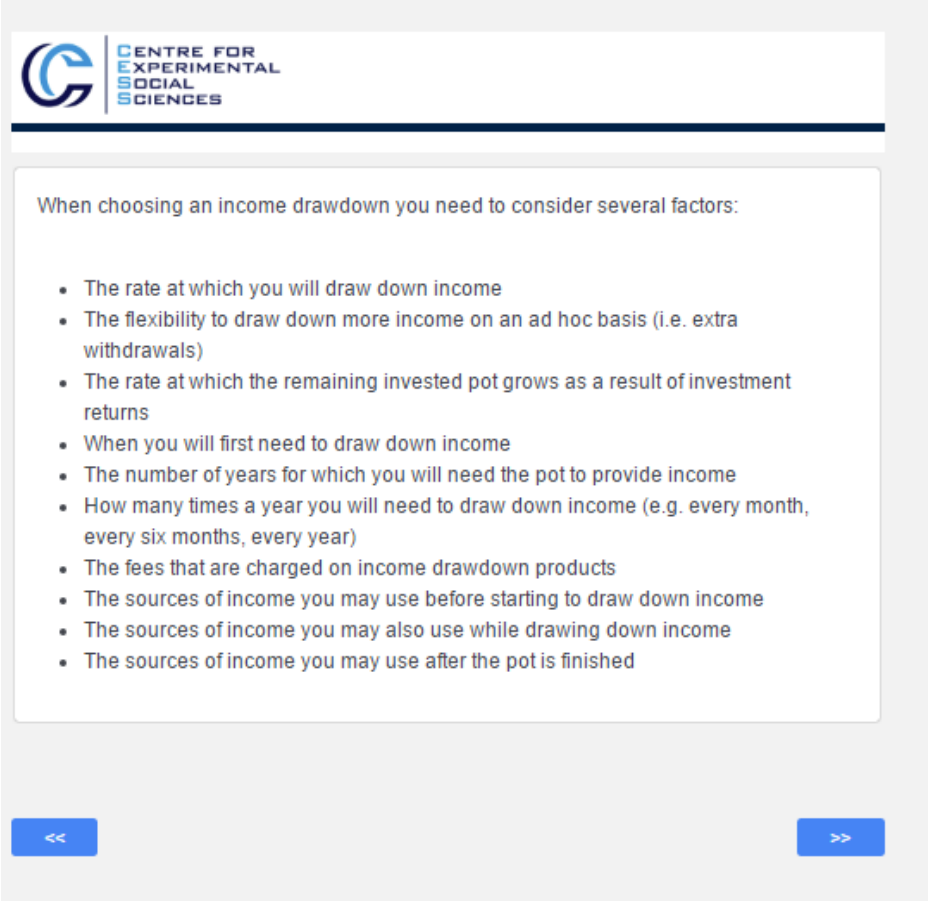
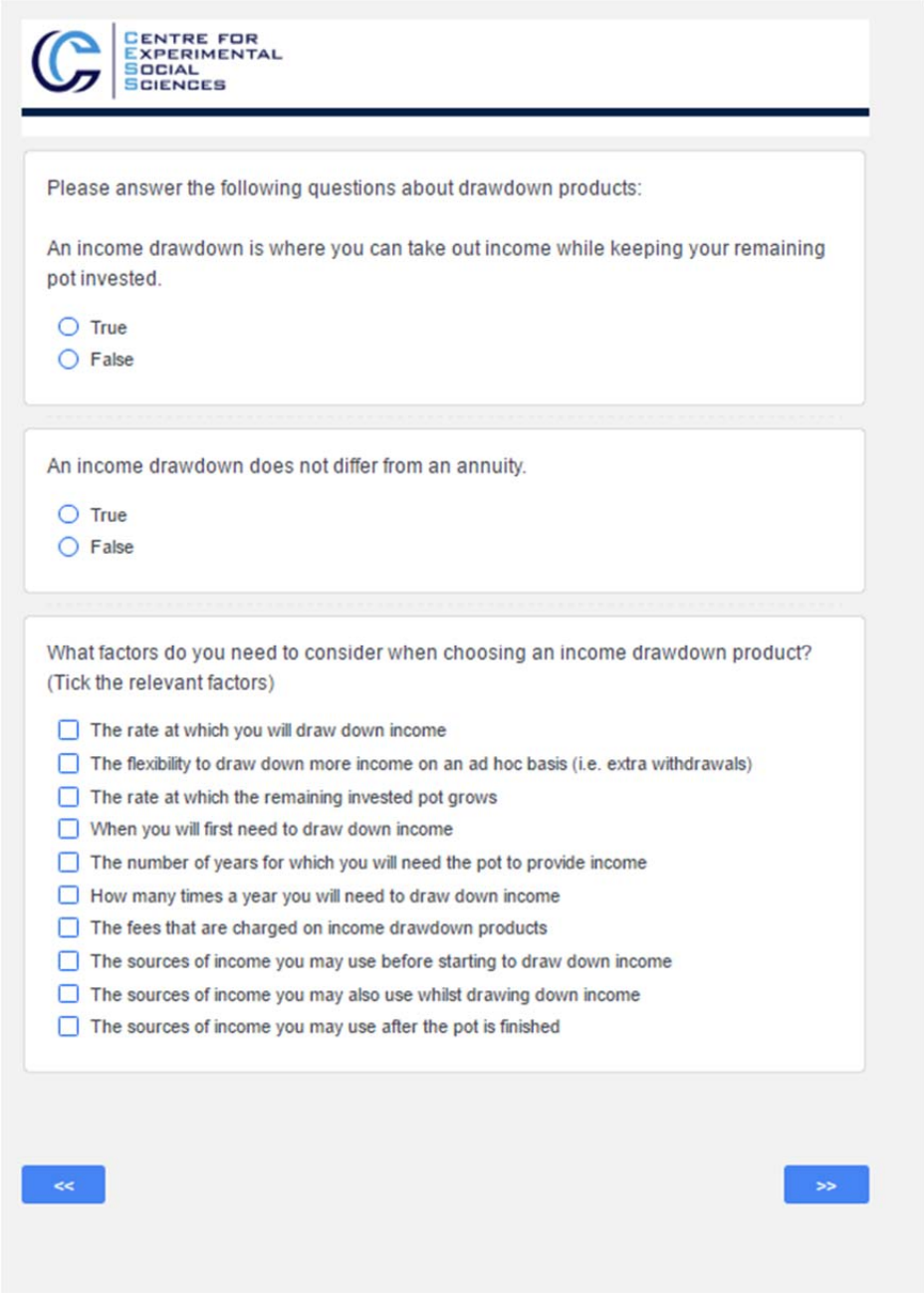


Figure A2.8 Questions 8, 9 and 10



Each participant then faced six profile vignettes. The names were male for male participants and female for female participants. Each profile was displayed on a separate screen, so each participant saw six screens describing them. The figures below show all 12 vignettes for male and female participants. The first vignette (David/Susan) appeared on the same screen as the task description.

Figure A2.9 Profile 1, David

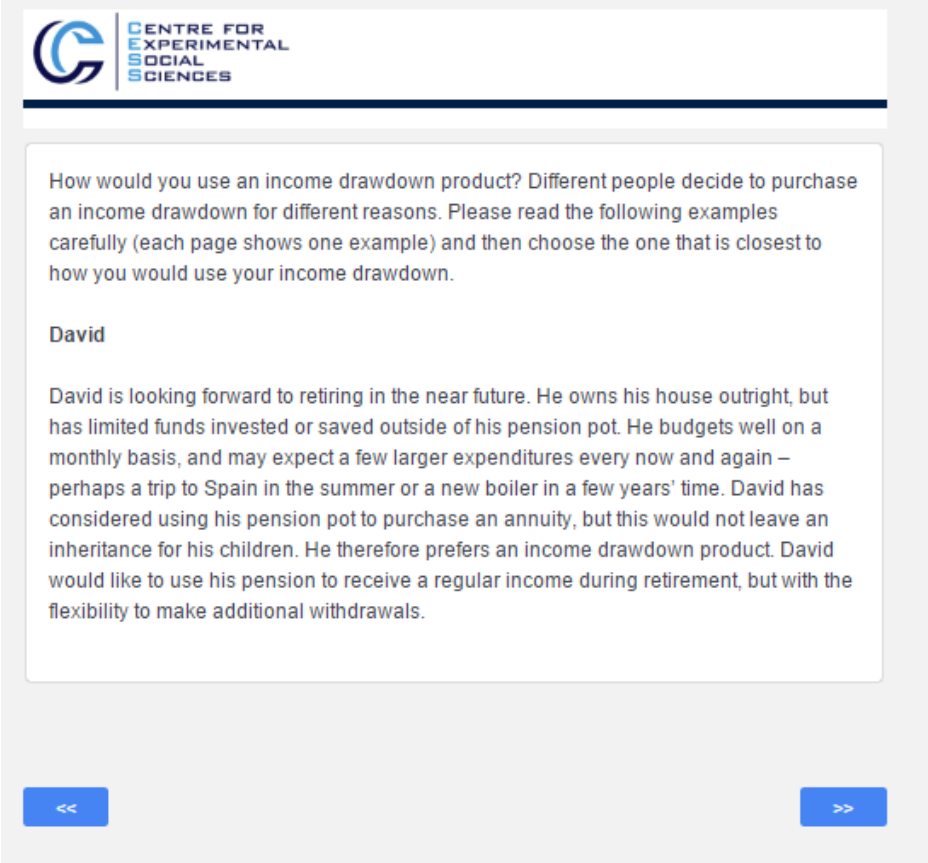


Figure A2.10 Profile 1, Susan

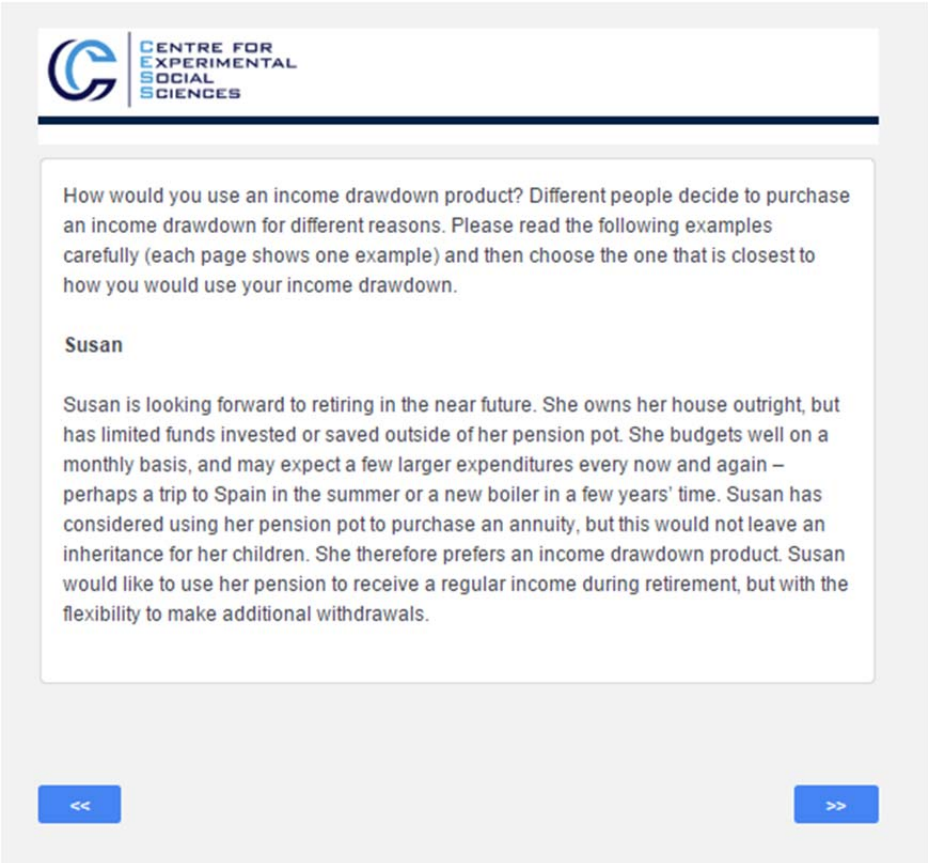


Figure A2.11 Profile 1, Paul

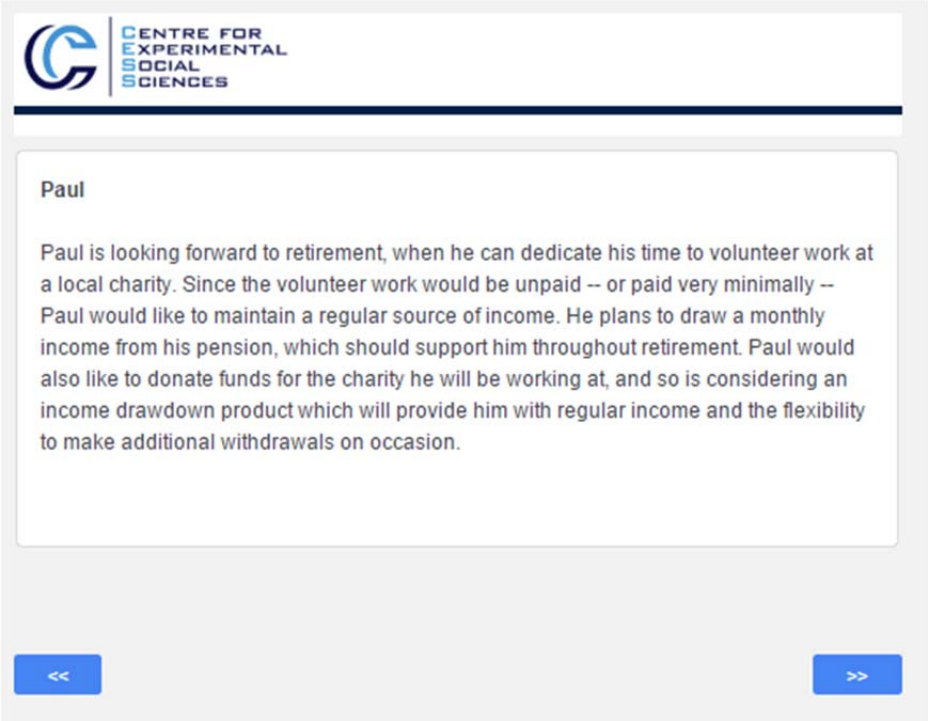


Figure A2.12 Profile 1, Julie

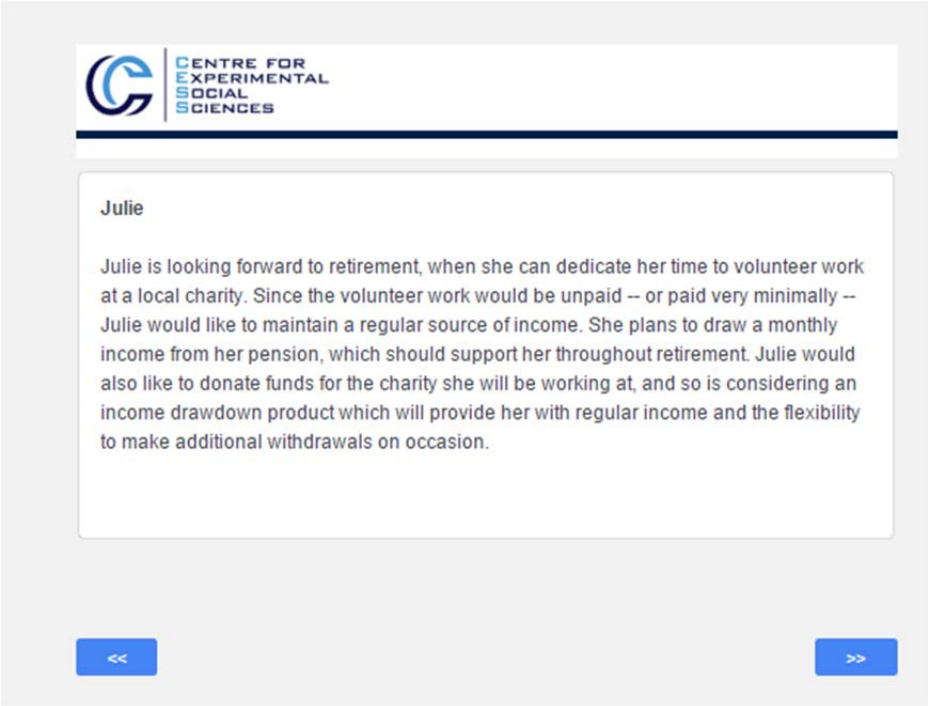


Figure A2.13 Profile 1, Andrew

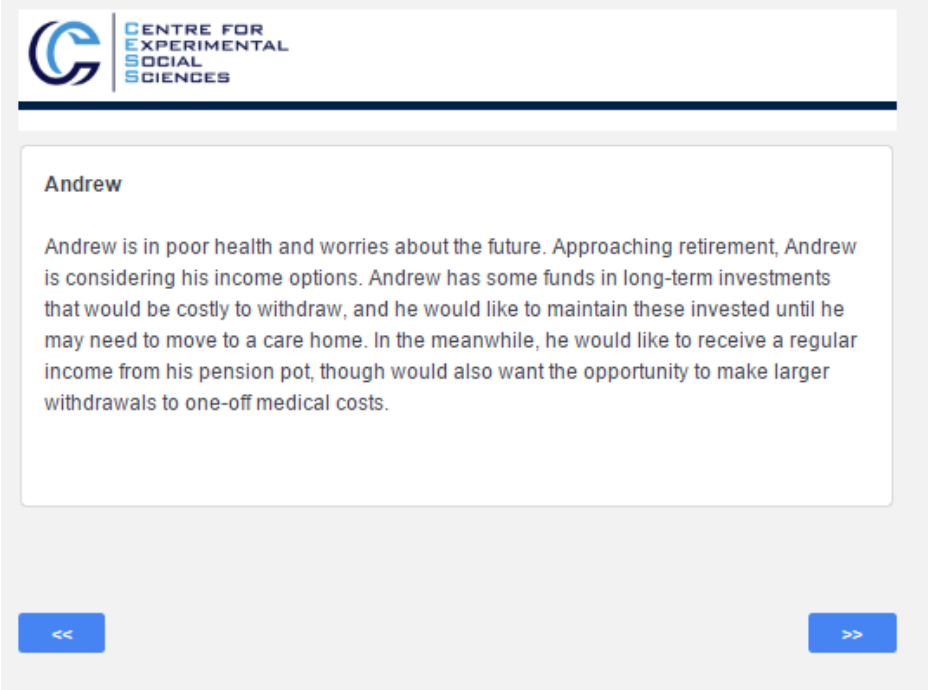


Figure A2.14 Profile 1, Karen

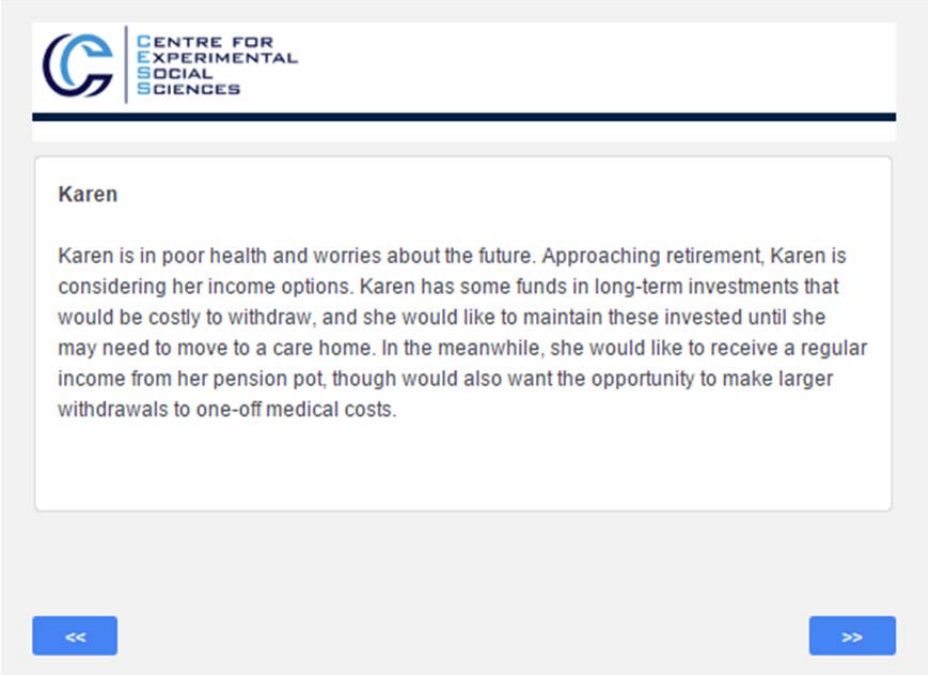


Figure A2.15 Profile 2, Mark

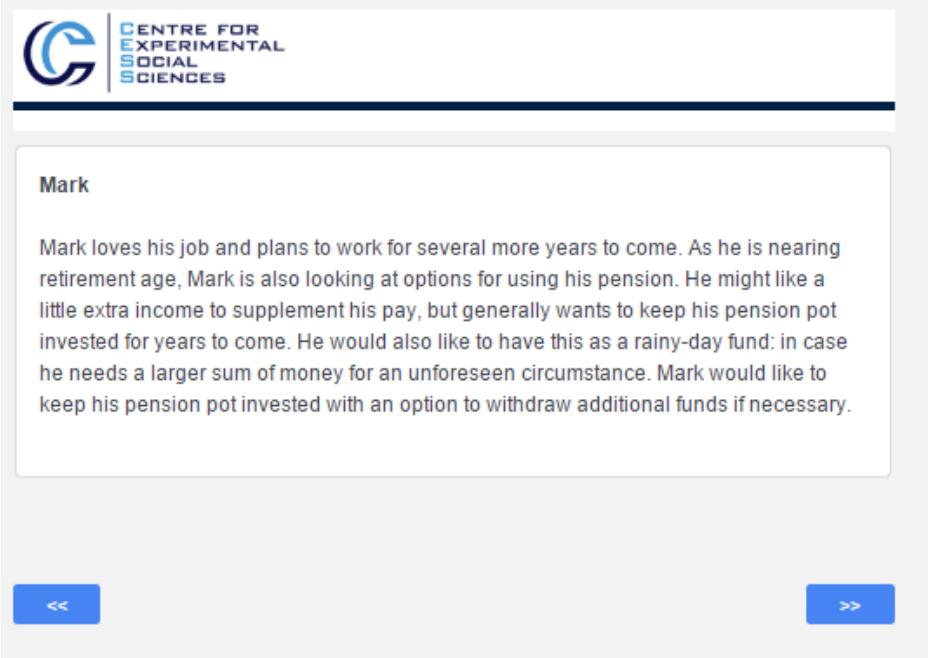


Figure A2.16 Profile 2, Helen

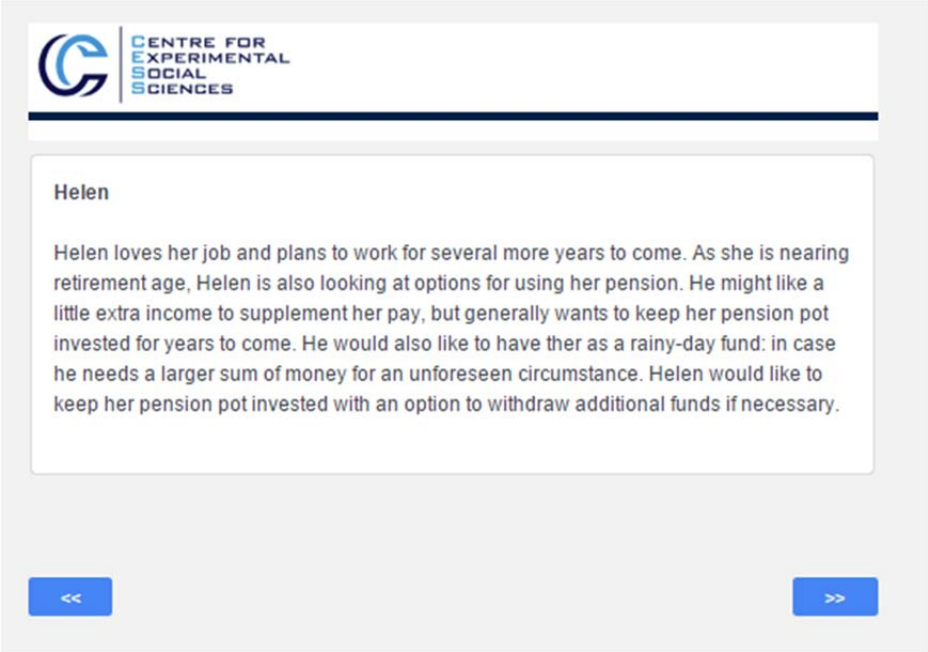


Figure A2.17 Profile 2, John

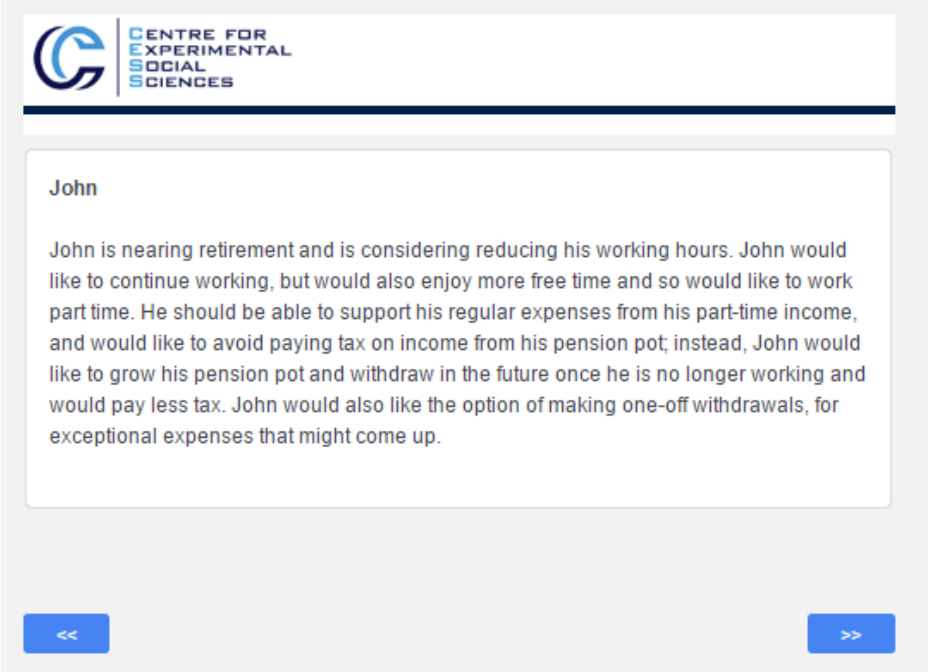


Figure A2.18 Profile 2, Deborah

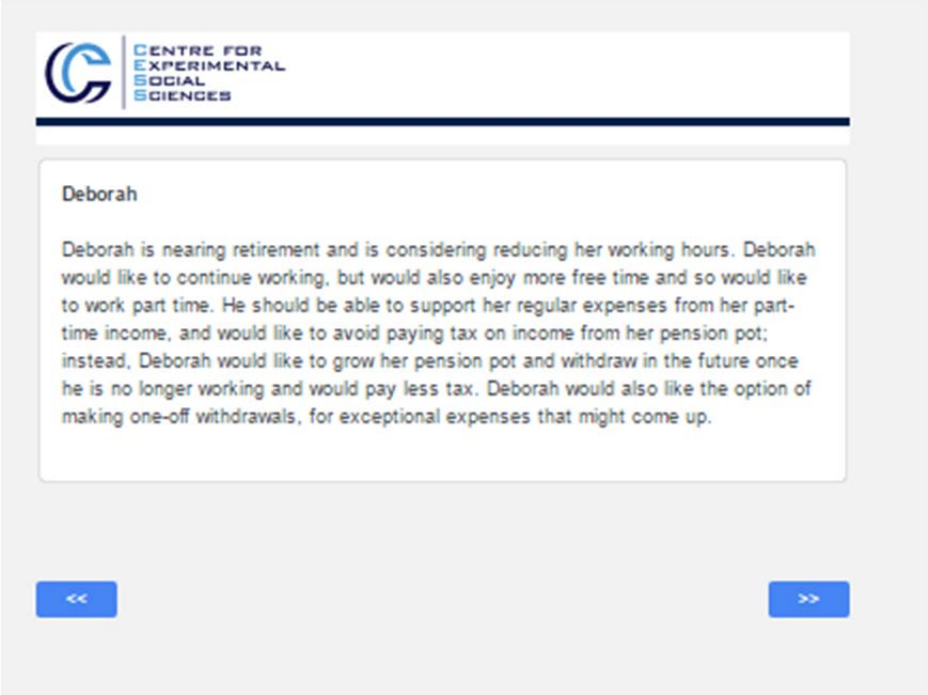


Figure A2.19 Profile 2, Stephen

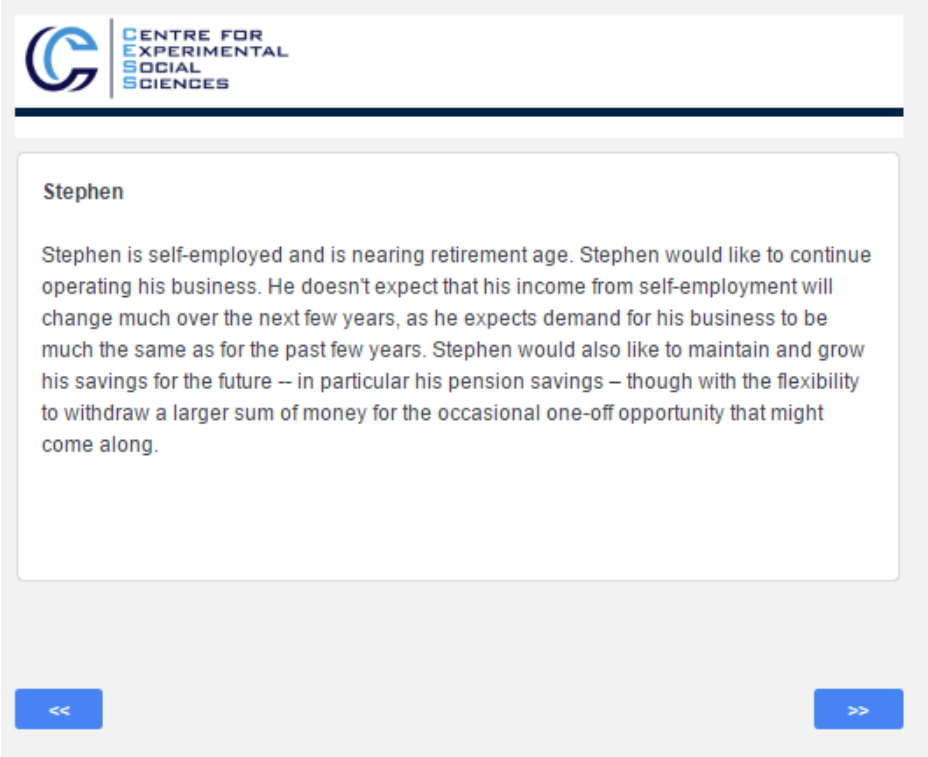


Figure A2.20 Profile 2, Tracy

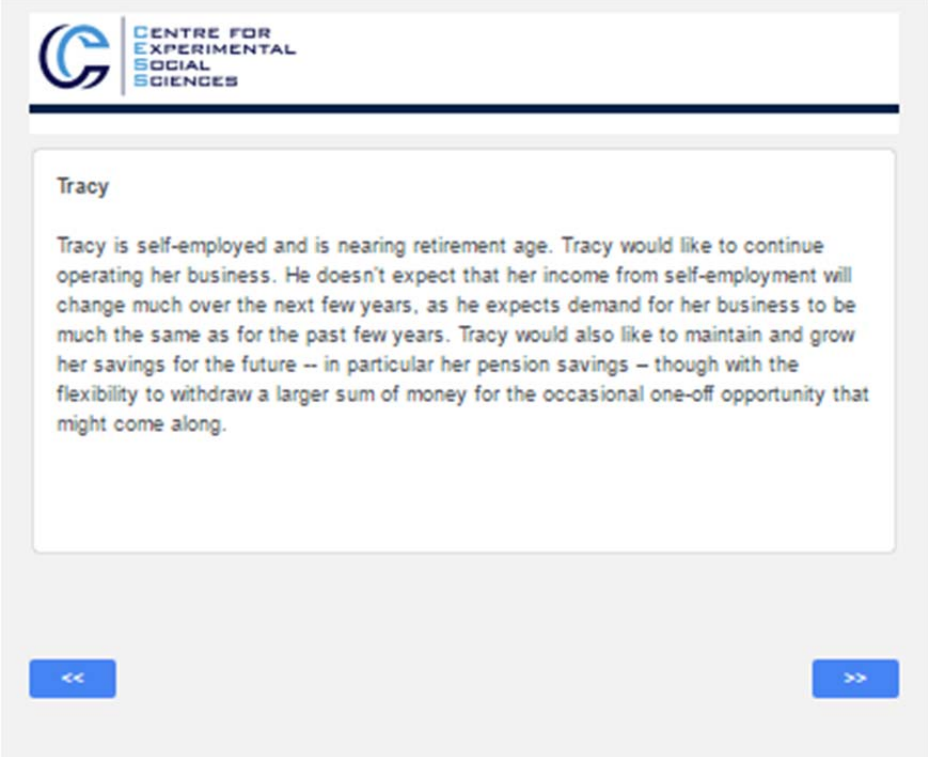


Figure A2.21 Questions 11 and 12

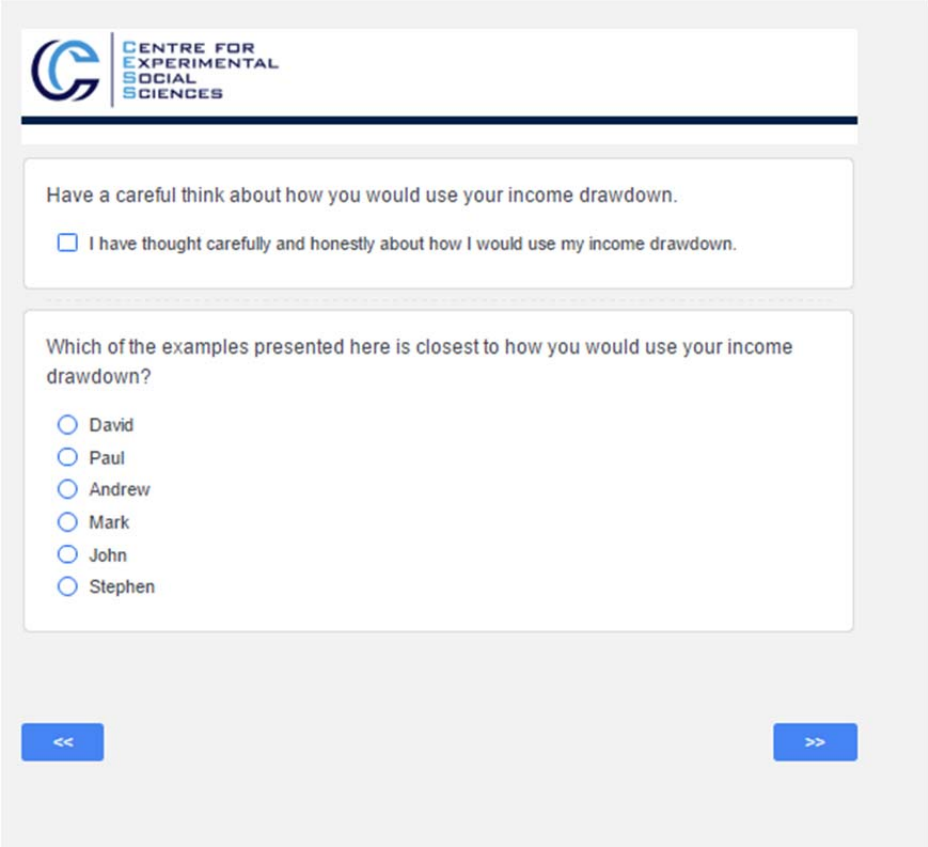


Figure A2.22 Retirement task payoff explanation

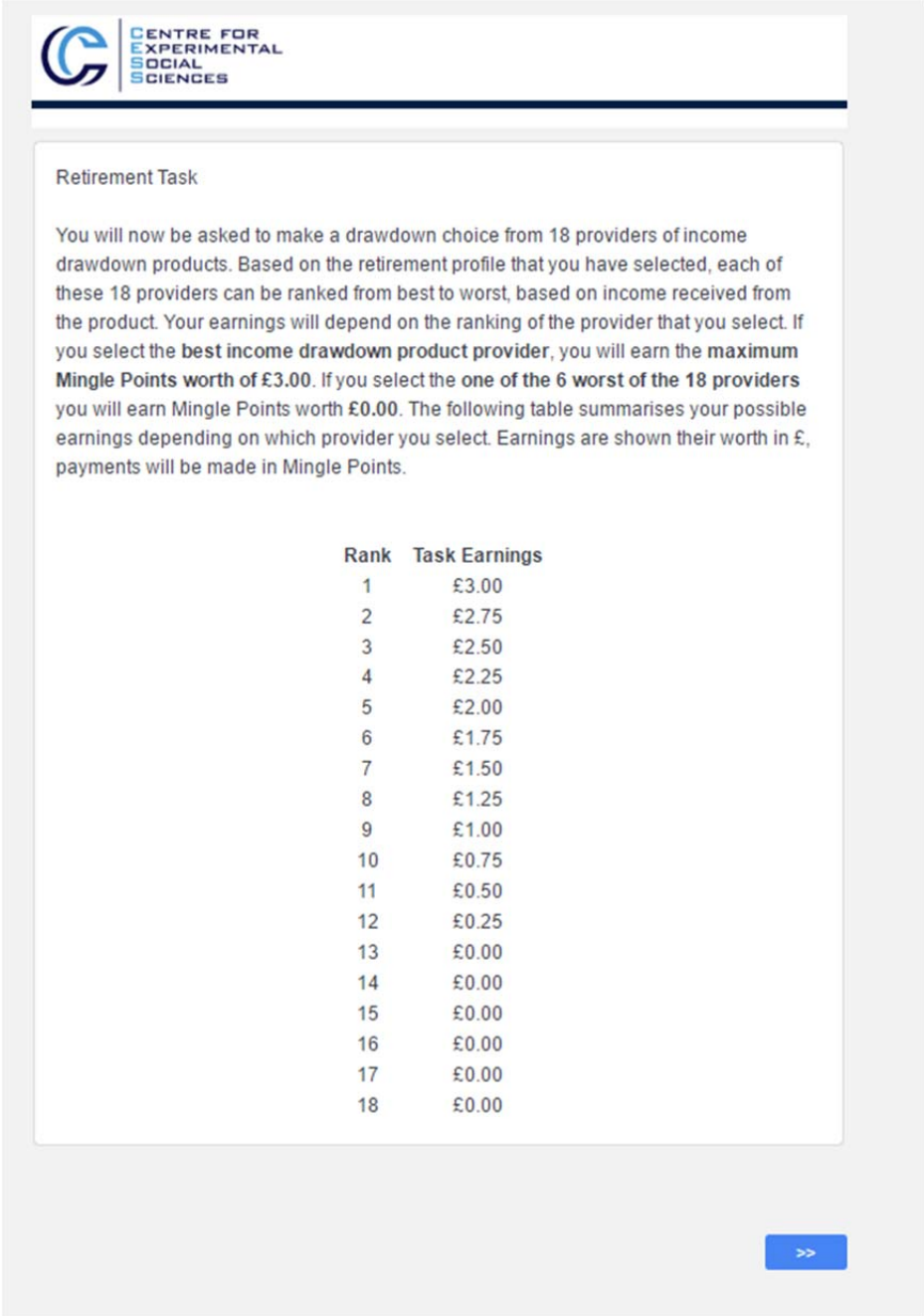


Figure A2.23 Price comparison task explanation

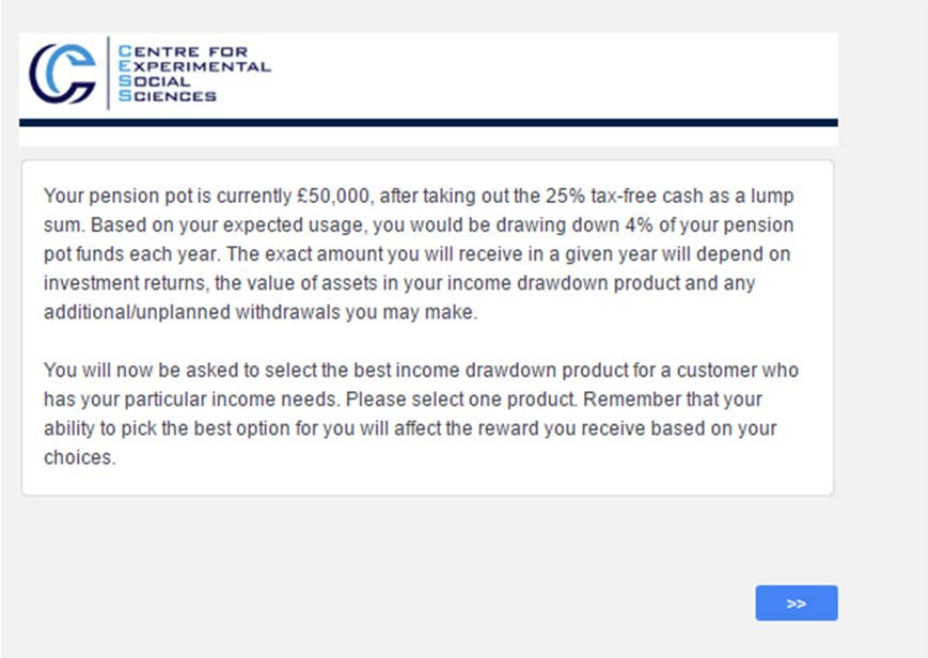
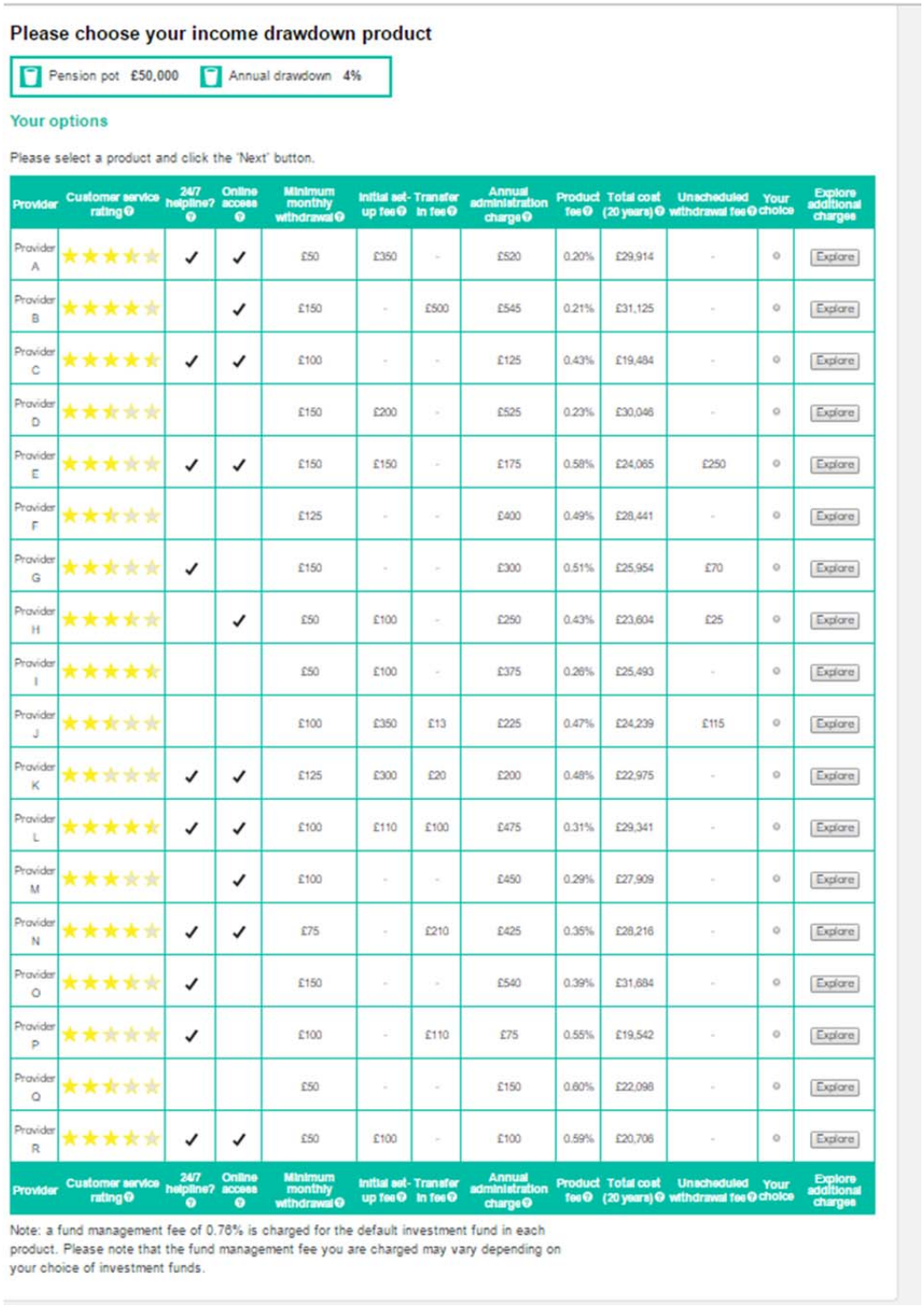


Figure A2.24 Price comparison website page, with total cost (20 years) treatment shown



Please choose your income drawdown product

Pension pot £50,000 Annual drawdown 4%

Your options

Please select a product and click the 'Next' button.

Provider	Customer service rating	24/7 helpline?	Online access	Minimum monthly withdrawal	Initial set-up fee	Transfer in fee	Annual administration charge	Product fee	Total cost (20 years)	Unscheduled withdrawal fee	Your choice	Explore additional charges
Provider A	★★★½☆	✓	✓	£50	£350	-	£520	0.20%	£29,914	-	○	Explore
Provider B	★★★★☆		✓	£150	-	£500	£545	0.21%	£31,125	-	○	Explore
Provider C	★★★★½	✓	✓	£100	-	-	£125	0.43%	£19,484	-	○	Explore
Provider D	★★½☆☆			£150	£200	-	£525	0.23%	£30,046	-	○	Explore
Provider E	★★★☆☆	✓	✓	£150	£150	-	£175	0.58%	£24,065	£250	○	Explore
Provider F	★★½☆☆			£125	-	-	£400	0.49%	£28,441	-	○	Explore
Provider G	★★½☆☆	✓		£150	-	-	£300	0.51%	£25,954	£70	○	Explore
Provider H	★★★½☆		✓	£50	£100	-	£250	0.43%	£23,604	£25	○	Explore
Provider I	★★★★½			£50	£100	-	£375	0.26%	£25,493	-	○	Explore
Provider J	★★½☆☆			£100	£350	£13	£225	0.47%	£24,239	£115	○	Explore
Provider K	★★☆☆☆	✓	✓	£125	£300	£20	£200	0.48%	£22,975	-	○	Explore
Provider L	★★★★½	✓	✓	£100	£110	£100	£475	0.31%	£29,341	-	○	Explore
Provider M	★★★☆☆		✓	£100	-	-	£450	0.29%	£27,909	-	○	Explore
Provider N	★★★★☆	✓	✓	£75	-	£210	£425	0.35%	£28,216	-	○	Explore
Provider O	★★★½☆	✓		£150	-	-	£540	0.39%	£31,684	-	○	Explore
Provider P	★★☆☆☆	✓		£100	-	£110	£75	0.55%	£19,542	-	○	Explore
Provider Q	★★½☆☆			£50	-	-	£150	0.60%	£22,098	-	○	Explore
Provider R	★★★½☆	✓	✓	£50	£100	-	£100	0.59%	£20,706	-	○	Explore

Note: a fund management fee of 0.76% is charged for the default investment fund in each product. Please note that the fund management fee you are charged may vary depending on your choice of investment funds.

Figure A2.25 Questions 13, 14, 15 and 16

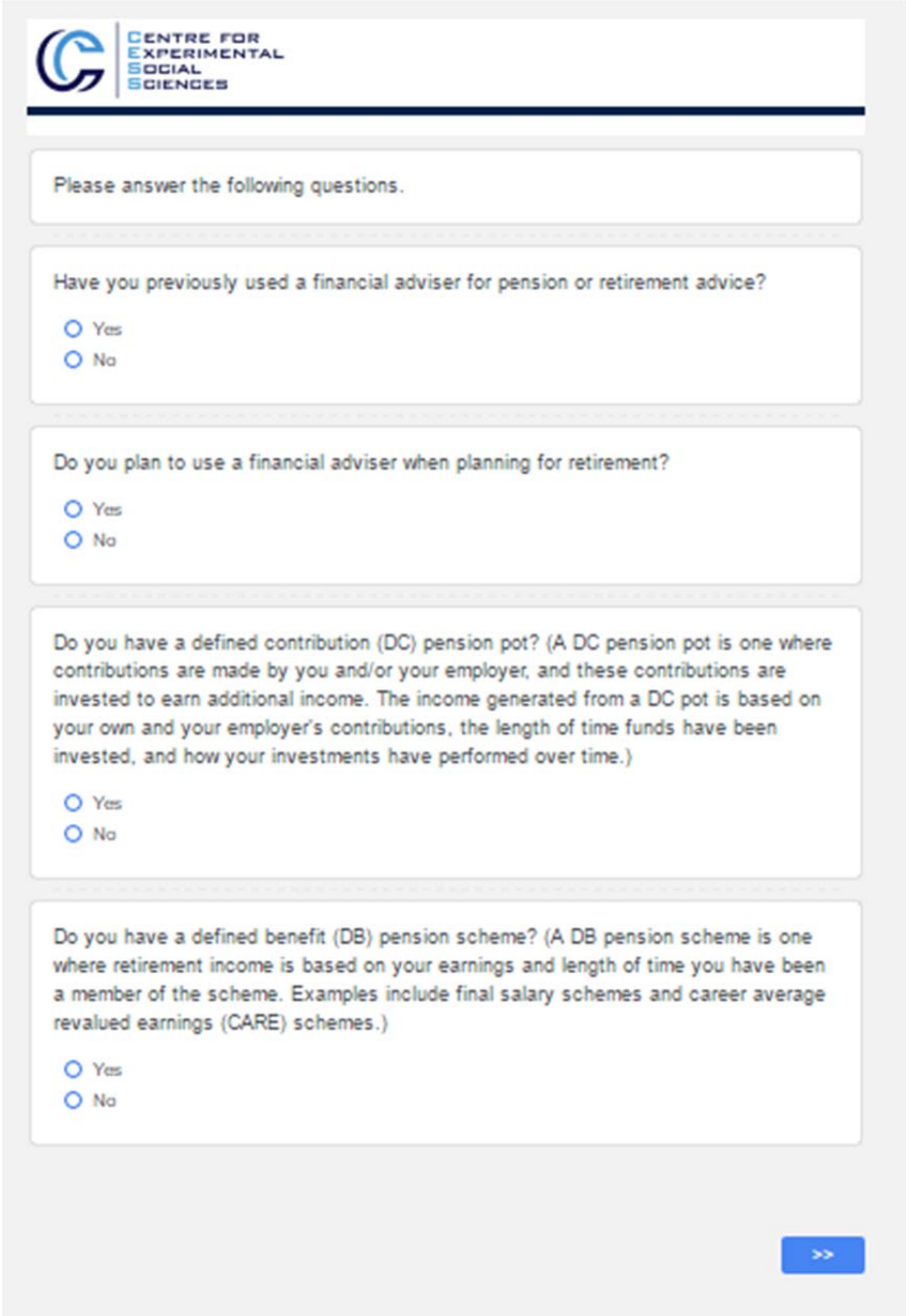


Figure A2.26 Questions 17, 18, 19, 20 and 21

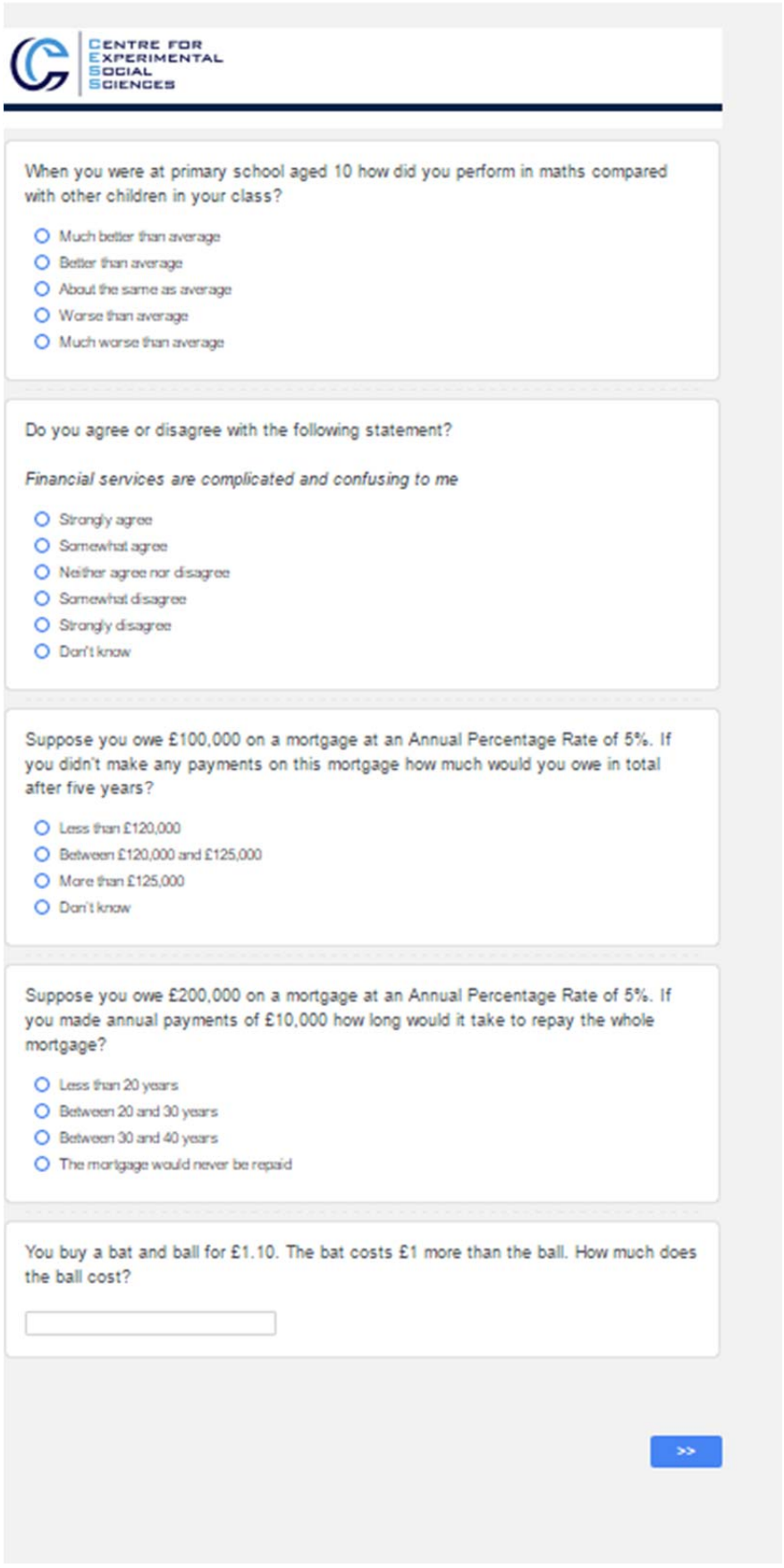


Figure A2.27 Question 22

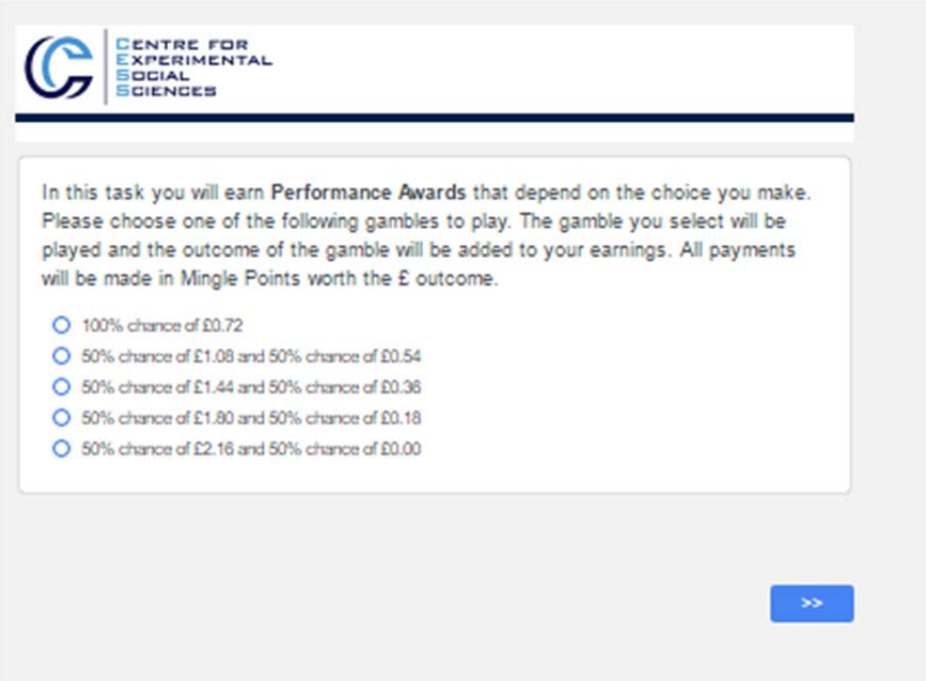


Figure A2.28 Questions 23 and 24

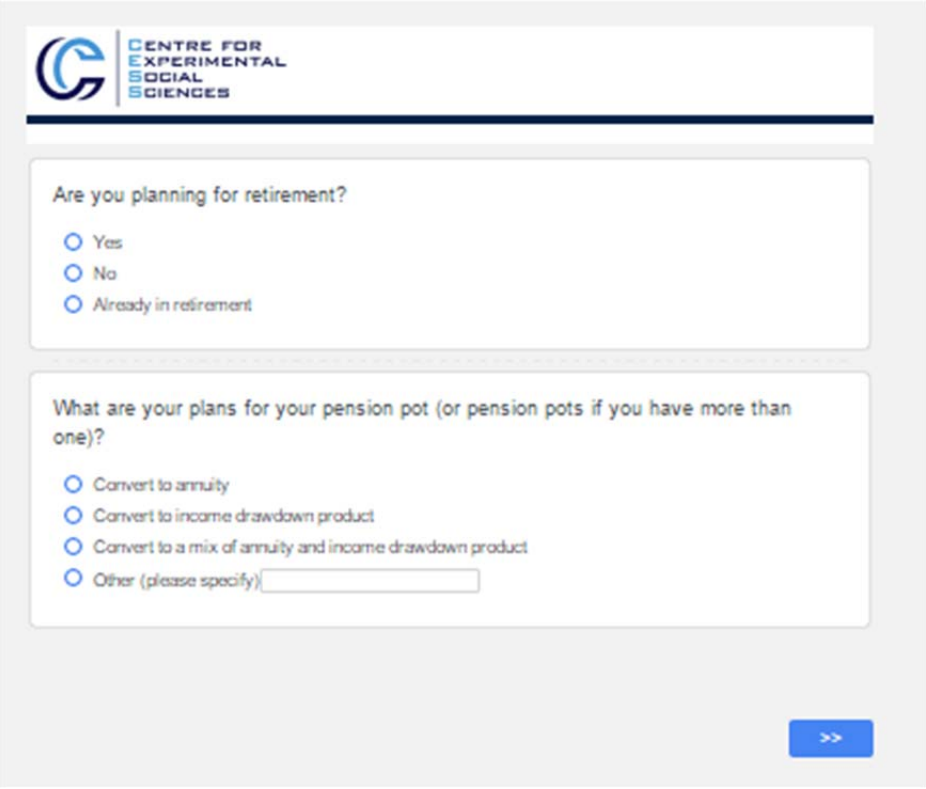
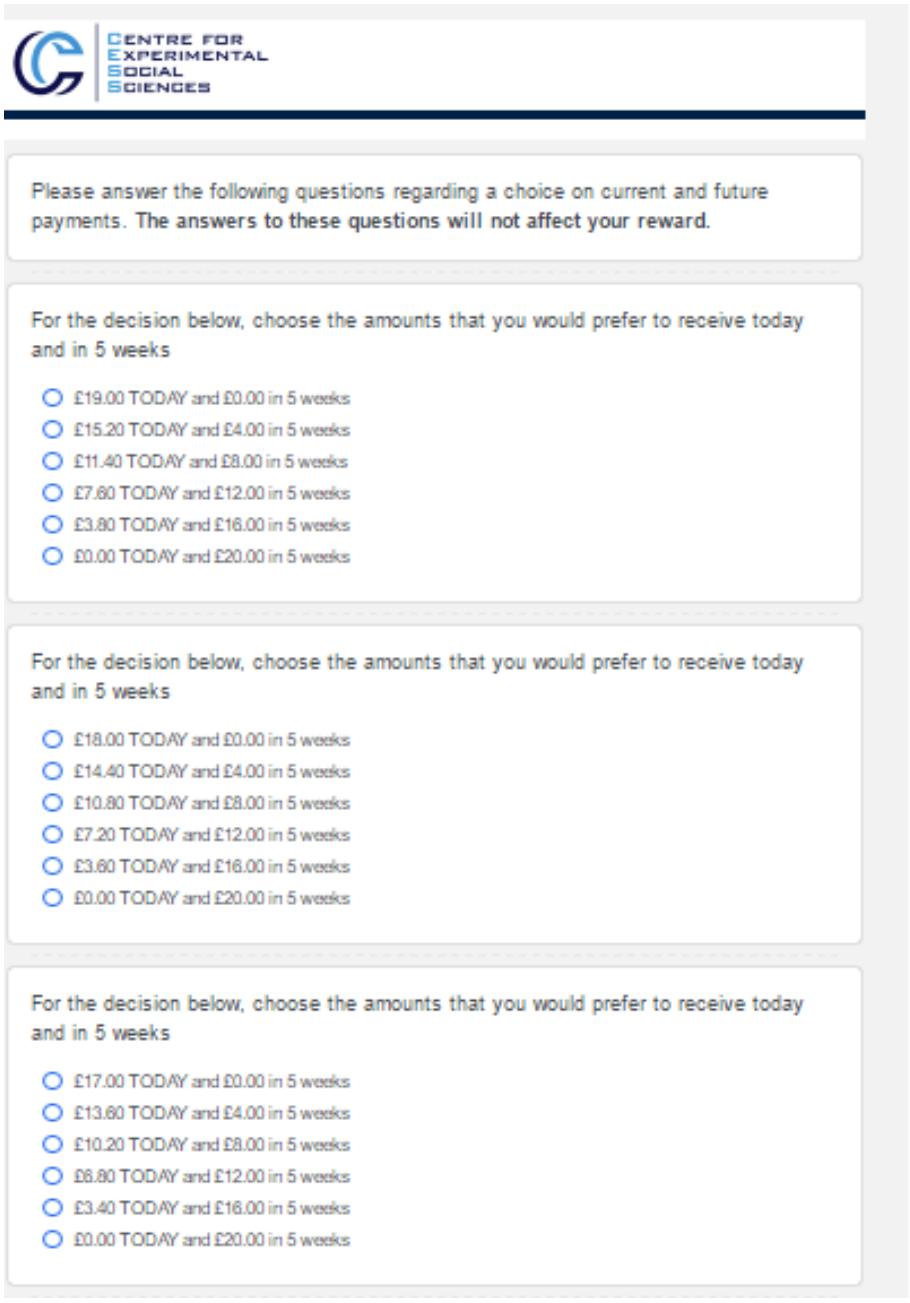


Figure A2.29 Questions 25, 26, 27, 28, 29 and 30



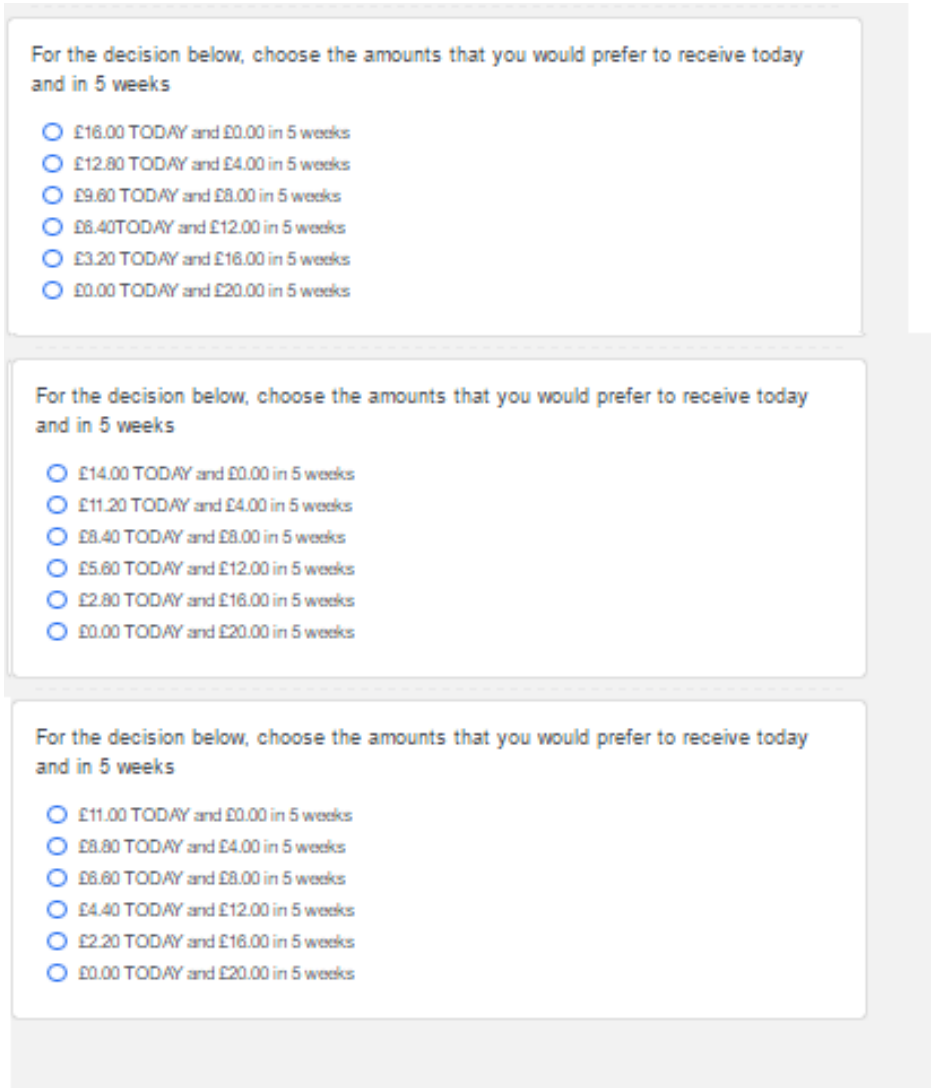

For the decision below, choose the amounts that you would prefer to receive today and in 5 weeks
£16.00 TODAY and £0.00 in 5 weeks
£12.80 TODAY and £4.00 in 5 weeks
£9.60 TODAY and £8.00 in 5 weeks
£6.40TODAY and £12.00 in 5 weeks
£3.20 TODAY and £16.00 in 5 weeks
£0.00 TODAY and £20.00 in 5 weeks
For the decision below, choose the amounts that you would prefer to receive today and in 5 weeks
£14.00 TODAY and £0.00 in 5 weeks
£11.20 TODAY and £4.00 in 5 weeks
£8.40 TODAY and £8.00 in 5 weeks
£5.60 TODAY and £12.00 in 5 weeks
£2.80 TODAY and £16.00 in 5 weeks
£0.00 TODAY and £20.00 in 5 weeks
For the decision below, choose the amounts that you would prefer to receive today and in 5 weeks
£11.00 TODAY and £0.00 in 5 weeks
£8.80 TODAY and £4.00 in 5 weeks
£6.60 TODAY and £8.00 in 5 weeks
£4.40 TODAY and £12.00 in 5 weeks
£2.20 TODAY and £16.00 in 5 weeks
£0.00 TODAY and £20.00 in 5 weeks

Figure A2.30 Questions 31, 32, 33, 34, 35 and 36

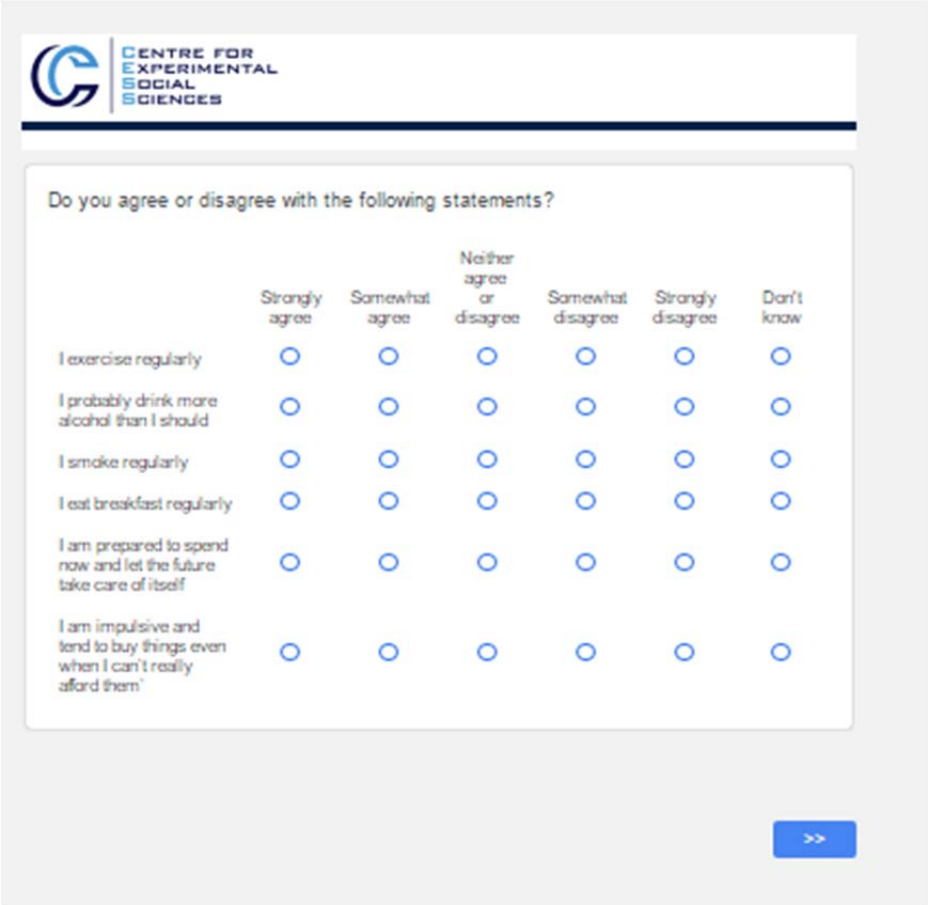


Figure A2.31 Questions 37 and 38

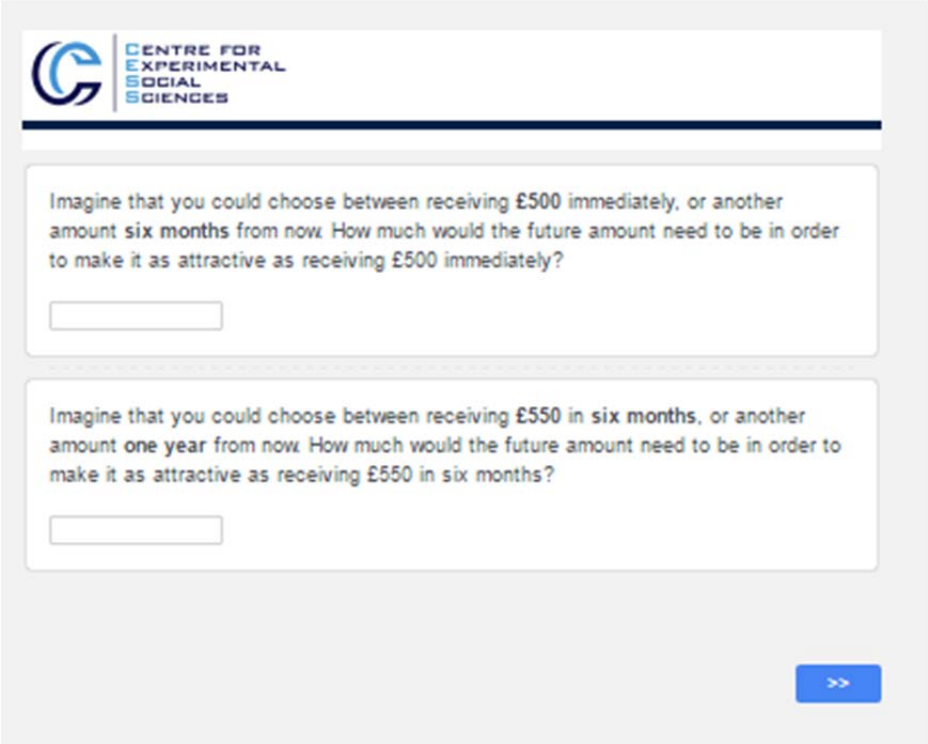
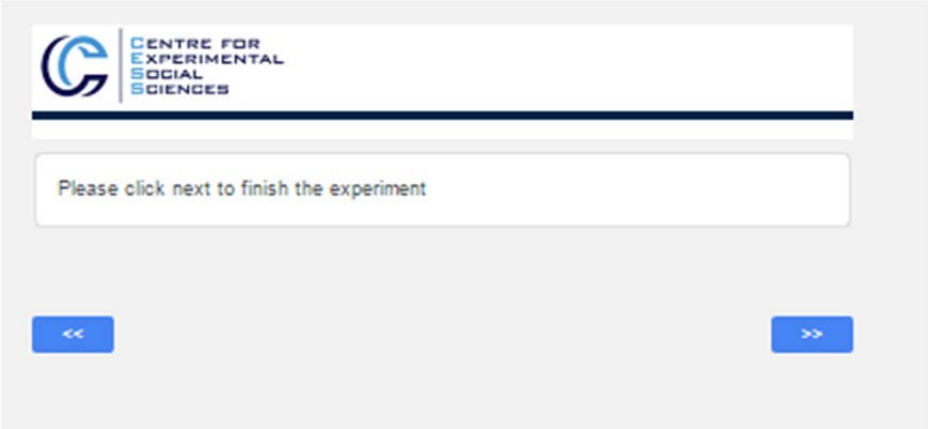


Figure A2.32 Final screen



A3 Regression tables

Regression analysis was conducted to assess the impact of the treatments on the different product selection metrics, controlling for other factors. This appendix sets out the results of the regression analysis.

In the tables in this appendix, t statistics are shown in parentheses, and the confidence intervals are denoted as follows: * significant at the 10% level; ** at the 5% level; *** at the 1% level. The omitted dummy variable in the regressions is the complex control treatment. Model (1) includes only the treatment variables and the constant; Model (2) includes more variables; and Model (3) includes all the variables.

Some of the variables are explained below.

Table A3.1 Variable definitions

Variable	Explanation
Impulsivity	Claimed extent of impulsiveness
Bat and ball	Correct on bat and ball question
Discount assessment	Sensible discount rate for six months' wait
Interest rates	Correct on interest rate question
Risk preferences	Preference for payment (low score means prefers now)
Spend now	Preference to spend (low score means prefers now)
Planning for retirement	Dummy for planning retirement
In retirement	Dummy for in retirement
Understand FS	Claims to understand Financial Services
Financial adviser	Has used a financial adviser

Source: Oxera/CESS.

A3.1 Ordinary least squares regressions

The product selection payoff, in monetary terms, varied from zero to £3. This was regressed against dummy variables for the treatments plus other explanatory variables which may determine product selection performance. The regression results are presented in Table A3.2 below.

A similar ordinary least square (OLS) regression was also conducted excluding participants who spent less than ten minutes on the experiment, which was deemed to be the minimum amount of time that the experiment could be completed while answering each question with some thought. Results are presented in Table A3.3.

Logistic regressions were also conducted on whether participants chose the best product C (Table A3.4) or chose the worst thirteen products (Table A3.5), regressed on the same set of explanatory variables. The worst thirteen products was chosen as an arbitrary cut-off point that was not found to affect the results significantly.

Regression coefficients report the difference in the payoff (£) between the complex control and the treatments.

Table A3.2 OLS regression on whole sample

	Model (1)	Model (2)	Model (3)
Complex control	-0.16* (-1.72)	-0.18* (-1.89)	-0.16* (-1.76)
PS after costs	0.4*** (4.35)	0.38*** (4.02)	0.37*** (3.98)
Average cost	0.21** (2.23)	0.2** (2.16)	0.2** (2.19)
Total cost	0.13 (1.39)	0.15 (1.61)	0.16* (1.71)
RiY	0.07 (0.79)	0.07 (0.77)	0.08 (0.84)
Cost rating	0.1 (1.03)	0.1 (1.06)	0.12 (1.34)
Profile		0.05 (1.02)	0.04 (0.77)
Gender		0.04 (0.84)	0.11** (2.15)
Age		0* (1.68)	0 (1.03)
Income		-0.03 (-0.64)	0.02 (0.38)
Education		0.21*** (3.89)	0.09 (1.57)
Time			4.19** (1.97)
Impulsivity			0.06** (2.13)
Bat and Ball			0.15** (2.46)
Discount rate			0.19*** (2.87)
Interest q.			0.2*** (3.64)
Risk pref.			0.05*** (2.58)
Spend now q.			0 (0.09)
Planning for ret.			-0.03 (-0.45)
In retirement			-0.12 (-1.49)
Understand FS			0.11* (1.72)
Financial advisor			-0.04 (-0.61)
Constant	2*** (31.08)	7.35*** (3.49)	0.77*** (2.89)
Number of obs	2,020	1,953	1,953

Source: Oxera/CESS.

Table A3.3 OLS regression on sample excluding those who spent ten minutes or less on the experiment

	Model (1)	Model (2)	Model (3)
Complex control	-0.16* (-1.72)	-0.16* (-1.67)	-0.15 (-1.58)
PS after costs	0.4*** (4.35)	0.35*** (3.53)	0.34*** (3.46)
Average cost	0.21** (2.23)	0.25** (2.5)	0.24** (2.47)
Total cost	0.13 (1.39)	0.19* (1.89)	0.2** (2.03)
RiY	0.07 (0.79)	0.13 (1.31)	0.13 (1.34)
Cost rating	0.1 (1.03)	0.15 (1.46)	0.17* (1.73)
Profile		0.03 (0.63)	0.02 (0.4)
Gender		0.05 (0.91)	0.13** (2.31)
Age		0 (1.15)	0 (0.93)
Income		-0.02 (-0.3)	0.04 (0.69)
Education		0.22*** (3.89)	0.1* (1.76)
Time		4.11* (1.88)	2.4 (1.1)
Impulsivity			0.05* (1.72)
Bat and Ball			0.14** (2.25)
Discount rate			0.18*** (2.74)
Interest q.			0.2*** (3.46)
Risk pref.			0.04** (2.04)
Spend now q.			0 (0.1)
Planning for ret.			0 (-0.07)
In retirement			-0.13 (-1.42)
Understand FS			0.11* (1.68)
Financial advisor			-0.03 (-0.55)
Constant	2*** (31.08)	1.53*** (6.74)	0.83*** (2.93)
Number of obs	1,789	1,722	1,722

Source: Oxera/CESS.

A3.2 Logistic regressions

Table A3.4 Logistic regression on whether the participants chose the best product (product C)

	Model (1)	Model (2)	Model (3)
Complex control	-0.27 (-1.59)	-0.34** (-1.97)	-0.32* (-1.81)
PS after costs	0.73*** (4.34)	0.67*** (3.83)	0.68*** (3.81)
Average cost	0.59*** (3.5)	0.56*** (3.27)	0.57*** (3.21)
Total cost	0.26 (1.51)	0.27 (1.54)	0.29 (1.62)
RiY	0.08 (0.51)	0.06 (0.35)	0.06 (0.38)
Cost rating	0.22 (1.3)	0.21 (1.2)	0.26 (1.49)
Profile		0.01 (0.14)	-0.01 (-0.12)
Gender		0.19** (2.08)	0.32*** (3.25)
Age		0.01 (1.12)	0 (-0.64)
Income		-0.02 (-0.25)	0.09 (0.93)
Education		0.36*** (3.65)	0.12 (1.11)
Time		8.62** (2.19)	1.98 (0.49)
Impulsivity			0.17*** (3.27)
Bat and Ball			0.21* (1.87)
Discount rate			0.3** (2.41)
Interest q.			0.39*** (3.66)
Risk pref.			0.07** (2.13)
Spend now q.			0.06 (1.05)
Planning for ret.			0.18 (1.38)
In retirement			0.07 (0.43)
Understand FS			0.11 (0.9)
Financial advisor			-0.02 (-0.22)
Constant	-0.34*** (-2.88)	-1.28*** (-3.21)	-2.51*** (-4.88)
Number of obs	2,020	1,953	1,953

Source: Oxera/CESS.

Table A3.5 Logistic regression on whether the participants chose one of the worst products

	Model (1)	Model (2)	Model (3)
Complex control	0.11 (0.66)	0.13 (0.76)	0.11 (0.64)
PS after costs	-0.85*** (-4.56)	-0.83*** (-4.36)	-0.84*** (-4.35)
Average cost	-0.43** (-2.44)	-0.44** (-2.43)	-0.45** (-2.46)
Total cost	-0.3* (-1.73)	-0.36** (-1.97)	-0.39** (-2.09)
RiY	-0.26 (-1.54)	-0.28 (-1.64)	-0.3* (-1.73)
Cost rating	-0.17 (-0.99)	-0.19 (-1.05)	-0.23 (-1.26)
Profile		-0.03 (-0.29)	0 (-0.05)
Gender		0.01 (0.11)	-0.1 (-0.93)
Age		-0.01** (-1.97)	-0.01* (-1.74)
Income		0.04 (0.42)	-0.04 (-0.38)
Education		-0.29*** (-2.88)	-0.12 (-1.12)
Time		-15.35*** (-3.32)	-10.02** (-2.17)
Impulsivity			-0.12** (-2.33)
Bat and Ball			-0.27** (-2.16)
Discount rate			-0.35*** (-2.59)
Interest q.			-0.27** (-2.46)
Risk pref.			-0.05 (-1.56)
Spend now q.			0.02 (0.29)
Planning for ret.			0.09 (0.72)
In retirement			0.32* (1.89)
Understand FS			-0.18 (-1.36)
Financial advisor			0.04 (0.32)
Constant	-0.42*** (-3.57)	0.59 (1.41)	1.69*** (3.2)
Number of obs	2,020	1,953	1,953

Source: Oxera/CESS.


www.oxera.com
oxera
compelling economics