

Occasional Paper No.2

Encouraging consumers to claim redress: evidence from a field trial

April 2013

Paul Adams, Stefan Hunt



FCA OCCASIONAL PAPERS IN FINANCIAL REGULATION

Disclaimer

All parties involved in this research ensured that individual consumer information was anonymous and protected at all times. Consumers were identified using a unique ID, and data files were encrypted.

The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Financial Conduct Authority (FCA). They do not. All errors or omissions are the authors' sole responsibility.

The Occasional Paper series

The FCA is committed to encouraging debate among academics, practitioners and policymakers in all aspects of financial regulation. To facilitate this, it publishes a series of Occasional Papers in financial regulation, extending across economics and other disciplines.

These papers cover topics such as the rationale for regulation, the costs and benefits of various aspects of regulation, and the structure and development of markets in financial services. Since their main purpose is to stimulate interest and debate, the FCA welcomes the opportunity to publish controversial and challenging material, including papers that may have been presented or published elsewhere.

The main factor in accepting papers, which are independently refereed, is that they should make substantial contributions to knowledge and understanding in the area of financial regulation. The FCA encourages contributions from external authors, as well as from its own staff. In either case, the papers will express the views of the author(s) and not those of the FCA.

If you want to contribute to this series, please contact Peter Andrews or Stefan Hunt at, respectively:

Telephone: 020 7066 3104; 020 7066 3554

Email: peter.andrews@fca.org.uk; stefan.hunt@fca.org.uk

These Occasional Papers are available on our website: www.fca.org.uk. Comments are welcome on these papers; please address them to the contacts listed above.

Biographical note

The authors work in the Chief Economist's Department of the Financial Conduct Authority.

Acknowledgements

We are thankful to the firm we worked with for their cooperation, without which this research would not have been possible. We are grateful to Peter Andrews, Robert Baker, James Bridge, Kristine Erta, Francesca Hopwood-Road, Zanna Iscenko, Mario Theodosiou and Charlotte Vowells for their help and comments. We also thank, from outside the FCA, Gharad Bryan and Steffen Huck for valuable remarks and astute critique, and Michael Hallsworth, David Halpern, Laura Haynes, Dean Karlan and Sam Nguyen for suggestions on further tests. Mistakes and omissions are solely the responsibility of the authors.

Contents

Foreword	3
Executive summary	4
1. Introduction	8
2. Research Context	10
3. Research Design	11
4. Results	14
5. Conclusions	23
Tables	24
Annexes	
1. Comparison of letters	34
2. Model equations	37
References	38
Endnotes	41

Foreword

At the Financial Conduct Authority we aim to put consumers at the heart of our regulation. This includes helping them to be informed about financial decisions, by carefully designing the information they receive. Effective communication is especially important when firms put things right after they have gone wrong, such as telling consumers about potential compensation (known as 'redress').

The public sector is focusing on how best to provide people with information using advances behavioural economics, which builds psychology into economics. In the US the Office of Information and Regulatory Affairs at the White House spearheads an initiative on disclosure and simplification as regulatory tools, underlining the critical role of testing and highlighting the benefits of field trials (also called randomised controlled trials or RCTs). In the UK, the Behavioural Insights Team at the Cabinet Office has led similar research, also focusing on trials.

By working in real settings and by repeating trials, RCTs build hard evidence that will likely apply to future regulatory interventions. RCTs have a number of benefits. The most relevant to this paper are that they produce results that are simple to understand and easy to apply, and that they provide precise quantitative information that allows us to pin down the impact of our interventions. This research delivers clear and unambiguous results that will be practically useful for future redress exercises and have already proved helpful.

While RCTs are increasingly used in the public sector, to our knowledge this is the first use by a regulator in the UK. This method could be used to build evidence on the impact of a variety of important regulatory interventions, especially where consumer behaviour makes these interventions difficult to design effectively, such as mandatory information disclosures.



Peter Andrews
Chief Economist, Financial Conduct Authority

Executive summary

Redress to consumers following the mis-selling of financial products is an important regulatory tool to secure an appropriate degree of protection for consumers, one of the three operational objectives of the Financial Conduct Authority (FCA). In 2011/12, excluding compensation for payment protection insurance (PPI), the FSA helped obtain over £150 million worth of redress for consumers. The FCA's focus on the consumer will continue to ensure that consumers receive appropriate redress.

While large redress exercises such as PPI receive considerable publicity, many instances where consumers are due redress understandably do not. In these cases, the firm alerts customers to a potential issue, often in the form of a letter that gives customers information, which they need to answer. We believe that some customers do not respond, even when they have been mis-sold and when it would be in their interest to act. One reason may be that the relevant information is obscured or more complex than necessary, or that consumers suffer from inertia. Firms alone may not have sufficient incentives to correct these issues.

This research focuses on how to encourage consumers who may be due redress to respond to letters. We worked on a real case, with a firm that was voluntarily writing to almost 200,000 customers about a failing in its sales process. Influenced by new and exciting advances in behavioural economics, we tested different potential improvements to the firm's letter using an RCT. And we used rigorous quantitative techniques to analyse the responses.

The results are clear, surprising in magnitude and direction, and challenge current accepted practice.

What did we test?

We begin with a simple, intuitive model of consumer behaviour, in part informed by findings from behavioural science. People receive a lot of mail in the post and have to sift through it in a limited amount of time. Much of this is marketing, which they may have little interest in. People have to decide which letters to open, which letters to look at and which letters to read in more detail. Even if a consumer reads a letter and is minded to act on it, there may be further barriers to responding, e.g. concern that calling a helpline may be time-consuming. People may intend to respond, but procrastinate or forget. Much of this decision-making is quick and automatic, rather than slow and deliberative.

This model suggests how firms might help consumers to pay suitable attention to communication about redress. The envelope, to be opened, needs to be appropriately distinctive. The key messages must be as salient and immediate as possible; firms need to reduce any excess verbiage. Firms need to reassure consumers that claiming redress will be as easy as possible. Consumers may respond better to letters sent by an authoritative individual or organisation. Firms should remind people to respond.

We developed changes to seven features of the communication (the ‘treatments’), and these are listed below in the order that consumers may receive them: changing the envelope, five changes to the firm’s letter and sending a reminder. The ‘control’ is the original envelope and letter designed by the firm with no reminder, see Annex 1.

Treatments

1. Envelope ▼	Adds a message to ‘act quickly’ to a plain envelope
2. FSA logo ▼	Uses the FSA logo in the letter head
3. Salient bullets ▼	Replaces the two bullet points at the top of the letter with more salient bullet points
4. Simplified ▼	Makes the body of the letter simpler and more concise, by reducing the text by 40%
5. Claims process ▼	Includes a sentence in bold explaining that the claims process would only take five minutes
6. CEO signature ▼	Uses the firm CEO’s signature to sign the letter, instead of a generic ‘Customer Team’
7. Reminder	Sends a second letter three to six weeks after the first

Over a five-week period the firm contacted different groups of consumers with various letters. Some were the firm’s original letter, while others included different combinations of the treatments. For each of the seven features listed, we randomly vary whether a particular customer received the treatment or control version. As we vary seven features, each with two versions, there were 128 possible combinations in total and so 128 groups of customers. We made sure that each of these groups was the same in terms of the important customer characteristics we could observe, such as age and amount of redress owed.

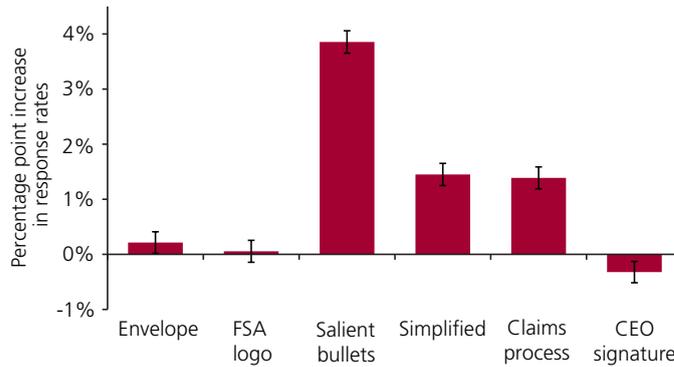
What were the results?

The firm’s original letter, the control with no treatments, received a 1.5% response rate, using our primary statistical model. Compared with other redress exercises undertaken by the FSA, this response rate is particularly low, although understandable in this particular setting.*

The changes we made had a marked effect on response rates. To begin, we use a basic statistical model with no interactions between the treatments. As Chart 1 shows, *Salient bullets* had the largest single effect, increasing response rates over the control by 3.8 percentage points, just over 2.5 times compared to the original letter. *Simplified* and *Claims process* each increase response by 1.4 percentage points, almost doubling the response rate. *Envelope* had only a small positive effect and there was no impact of *FSA logo*. Unexpectedly, there was a small but statistically significant decrease in response using *CEO signature*. As *Reminder* has clear interaction effects, we discuss it below.

* There are several reasons. First, many consumers had already been provided with a refund from the firm on their own initiative. Second, a number may also have been happy with the sales process and not felt in need of redress. Third, the potential value of redress was low, the average redress due was only £21. Fourth, the relationship between the firm and the consumer had already ended, which may mean the firm has an out-of-date address or that the consumer is less likely to open the envelope.

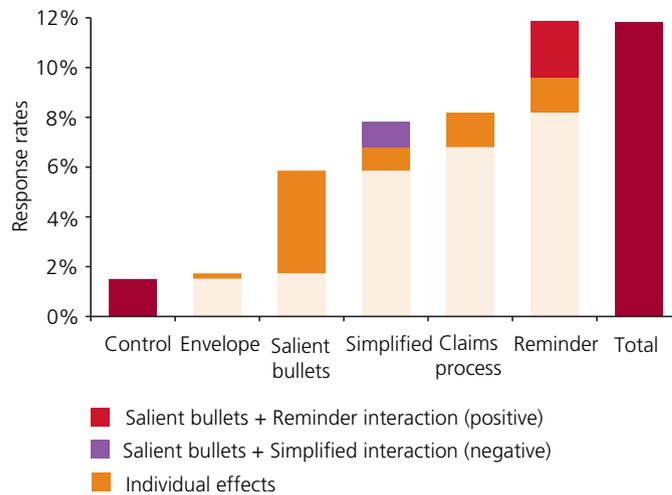
Chart 1: Effect of treatments relative to control, no interaction effects



Notes: Includes 5% confidence intervals

Because we vary each of the treatments independently (‘crossing’ treatments), we can carefully estimate the effect of multiple treatments, and so find the effect of the ‘best letter’, with the best combination of treatments. Our primary statistical model therefore allows for interactions between different treatments. As we can see in Chart 2, while *Salient bullets* and *Simplified* each encourage response separately, having both together has less of an effect than merely adding the two separate responses – a ‘negative interaction’. Alternatively, we find that *Reminder*, which was a copy of the original letter and so can interact with the other treatments, has much more effect if it has *Salient bullets* – a ‘positive interaction’. We are able to improve response rates to almost 12%. This is equivalent to an additional 20,000 people responding to claim redress. Overall, the best letter increases the response rate by over seven times compared to the control.

Chart 2: Effect sizes including interactions



We found a number of other relevant results:

- Reminder letters were sent out between three and six weeks after the original letter and responses are significantly higher when sent at the three week point.

- Overall, gender plays little role in response to the letter. There are two main areas where this is not the case: only women are affected negatively by *CEO signature* and women react significantly more than men to *Salient bullets*.
- In contrast, there are marked differences across different age groups. With the control letter the middle-aged respond the least and older age groups respond far more. But the pattern changes for the best letter: the young respond the least and response increases with age. So our treatments have the greatest relative effect on the middle-aged, who are arguably the busiest.
- There are fewer marked differences across those people due different amounts of redress. With the control letter there is little change in response between those who are due £50 or more and those who are due less than £10. But with the best letter, there is a stronger relationship between response and redress due; however, this variation is still less than the variation in response with age. The fact that response rates to the control letter did not vary much with the size of redress suggests that the control letter failed to focus consumers' attention on the amount of redress owed.

What did we learn?

Our results provide hard evidence on what did and did not work in this particular setting. Taken as a whole, the results are compatible with our simple model of busy people reviewing quickly the post that they receive.

The surprise in our findings is in the overall magnitude of the effects and the relative effect of different treatments. Our small improvements to what was already a clear letter had a dramatic impact. While we anticipated some impact based on our knowledge and experience of relevant research, the overall effect was much bigger than we expected. We also did not predict the overall pattern of results across the treatments; and, we did not expect any negative effects at all. Difficulty in predicting results when applying psychological insights is common. It underscores the importance of original research instead of extrapolating results from different contexts.

Knowing objectively and quantitatively which elements of a letter are the most important helps our supervisors' dialogue with firms regarding communication with customers about redress. Our findings underline the importance of the precise design of the communication.

By generating evidence from real settings, RCTs can build reliable evidence that will likely apply to future redress exercises, especially if we can repeat trials and validate their conclusions. The research highlights the potential for RCTs to provide evidence to aid regulatory decisions. The results are not only reliable, but simple to understand and easy to apply.

1. Introduction

Redress to consumers following the mis-selling of financial products is an important component of consumer protection. Large redress exercises receive considerable media attention. Financial firms paid out £11.8 billion in compensation to consumers after mis-selling private pension plans in the 1990s. Given the current level of provisions, compensation payments for Payment Protection Insurance (PPI) could overtake that figure. Both cases were, understandably, well-covered by the press.

However, most instances where consumers are due redress do not receive publicity. In these cases, the firm alerts customers to a potential issue, often in the form of a letter to which consumers must respond. Although response rates to these exercises vary widely, they are sometimes lower than we might expect. It is possible that consumers do not respond even though they are due redress. This research focuses on how to encourage consumers who are due redress to respond. Working with a firm that was voluntarily writing to customers about a failing in its sales process, we apply rigorous quantitative techniques in the form of a randomised control trial to test behaviourally-informed improvements to the firm's letter.

Traditional economic models assume that customers make a fully rational choice to respond or not based on the costs and benefits associated with each choice. Initially, the customer will make a decision to open the letter based on time constraints. Then, the customer will fully process the information in the letter and make a decision to respond that depends on an estimation of whether she has been mis-sold, the compensation she might be due and the costs she faces in responding. If expected compensation outweighs costs, then she will respond.

The simplest traditional models would allow consumers to process information effortlessly. One minor modification is to assume that consumers still make rational choices but face a variety of transaction costs in opening, reading and dealing with mail.¹ Consumers are busy and have a limited ability to sift through their mail. They need to prioritise. If letters do not appear important, they will not open them, not read them, or just skim them to detect whether it is worth reading in greater detail.

This modified rational model provides the basis for our research design. In addition, evidence about the behaviour of imperfectly rational consumers helps us consider how people process information and what causes them to act. Specifically, we focus on three areas to develop our research: how to highlight the key messages; who communication should come from and be associated with; and, the importance of reminding people to take action.

We vary seven features of the original letter designed by the firm. We randomly assign each customer to receive the treatment or control version of each of these features. This randomised controlled trial (RCT) allows us to estimate the causal impact of each of these different treatment variations and the potential interactions between treatments. In doing so, we are able to establish the best combination of treatments, the 'best letter'.

Recent use of RCTs in UK policy has been spearheaded by the Behavioural Insights Team (BIT) at the Cabinet Office.² Trials and the use of psychology in policy are tightly linked: once we allow for a richer model of consumer behaviour, we need empirical evidence to determine actual consumer responses to policy. To harness behavioural insights policymakers need to embrace new forms of research. Of particular relevance to this research, the BIT has used several RCTs to establish that making small changes to letters can have large effects on when and whether people pay their taxes.

We build on these studies by applying similar insights to a specific financial regulation problem, redress. We also advance the research methods used in UK policy and regulatory research. We test multiple treatments simultaneously and independently, allowing us to discover quickly the best combination of treatments. The large population in this study and careful stratification of the population allows us to analyse the results for specific sub-populations, e.g. different age groups, to add further detail and nuance to our findings. These are not merely technical considerations: our results will help firms and the FCA to develop letters that resonate with customers and ensure that redress is appropriately delivered.

We have written this paper for policymakers, other regulators, interested firms and consumers. We report only the headline results in the main text; however the underlying analysis is statistically rigorous and has been academically peer-reviewed. We include our econometric outputs and summary statistics in the Annex. Further details are available on request.

We organise this paper as follows. The first section provides more detail on the context in which the research took place, the second section briefly explains the research design we use, the third section provides the results from the study and the final section provides some high-level conclusions and indications for further research.

2. Research Context

When firms do not follow the Financial Conduct Authority's (or previously the FSA's) Rules or Principles, and consumers incur financial losses as a result, we will strive to ensure that customers receive adequate compensation, or redress. Redress can be considered as a particular subset of compensation, whereby consumers are returned to the financial position they would have been in without any wrongdoing. In some cases, the firm alerts customers to a potential issue requiring redress, through letters that customers need to answer. While we always strive to ensure these letters are as good as possible, we have never before systematically tested ways to improve them. This research shows how the FCA will use the most up-to-date research methods to improve outcomes for consumers.

The research is based on a redress exercise voluntarily undertaken by a financial services firm (the 'firm'), at the time under FSA supervision. The firm was planning to write to approximately 200,000 former customers to address past inadequacies in its sales process. The letters asked those who felt they had been misled when sold their product to call the firm's helpline to discuss the sale. 100% of those customers who felt they were misled received a refund. Working in partnership with the firm, we use this mass communication exercise to run a randomised controlled trial. Our main outcome measure is whether the customer responded to the letter to claim a refund. We also have data on those customers who called without requesting a refund but do not focus on it in this paper. We observe the response of each individual customer, their particular characteristics and which treatments they received.

The population of consumers included in our test had each bought a single product between 2010 and 2012 and had subsequently cancelled the product between one and six months after starting to pay the monthly premium.³ Customers were originally sold one of several products. Our sample covered all adult ages and just over half were male. Because customers on average cancelled their product after just over two months of payments, the level of total redress due was relatively small. Mean (median) total redress per customer was £21 (£16). A full set of summary statistics is in Table 1.

3. Research Design

Why use a randomised controlled trial?

Our research design is based on a randomised controlled trial (RCT). In RCTs, some individuals or groups are assigned to a 'control' intervention that is then compared to a 'treatment' intervention. RCTs have been used extensively in medical testing and increasingly in policy development. The use of control and treatment groups allows a direct comparison of outcomes. And by randomising the allocation of a medicine or a policy intervention, the approach removes problems relating to selection bias and therefore allows us to draw conclusions about the causal effect of interventions on the mean outcome. These features make it especially powerful in evaluating the impact of our interventions.

This research uses an RCT design conducted in a real-life setting; such trials are often called 'field trials'. RCTs conducted in this way have a number of benefits over other RCTs and other research designs. By working with real consumers, we can reliably estimate the impact of our treatments on real decisions. In contrast, in a focus group or other qualitative research, a consumer can focus far too much on the different letters, evaluating them in a completely different manner than in a natural setting when they might give cursory attention to the letter.⁴ Expert opinion also has a number of known problems: in particular in this setting, experts may know too much about a given area and what seems best to them may not lead to the most effective presentation of information for the average consumer.⁵ RCTs conducted in a laboratory experiment may provide clear and manageable environment for testing, but results may not extrapolate as well to the real world.⁶ We recognise that other forms of research also have their strengths, especially in situations where outcomes might be hard to measure in the real world.

By working in a real setting, and by repeating experiments to establish convincing results, RCTs build hard evidence that will likely apply to future redress exercises. RCTs therefore have a number of benefits. They can provide causal evidence that can rule out conventional wisdom.⁷ They can provide causal evidence in favour of things that most people would never think particularly plausible.⁸ And, most importantly for this research, they provide precise quantitative information that allows us to pin down effect sizes.⁹

Our 'control' is the letter the firm would have sent to customers in the absence of this research. We then split this letter into seven specific features. For each feature, we develop seven alterations or additions to the control version, the treatments. For each feature, either the customer receives the firm's version – the 'control' – or an alternative – the 'treatment'. We stratify customers based on their observable characteristics and assign each customer to either treatment or control versions of each feature. While the stratification is not essential given our sample size, it increases power and improves our ability to analyse the results of particular sub-groups, e.g. the elderly. The assignment of a particular feature to a specific customer is randomised and is independent of the assignment made for all other features.

What treatments did we test?

The modified rational model of consumer behaviour described in the introduction is the basis for our treatments. Behavioural insights help us consider how to reduce these transaction costs and help us develop three hypotheses which we will test in this research. First, if information is not salient it may not be attended to and may not affect economic choices. So we thought firms could highlight key messages by using simple language or by bolding the most relevant text.¹⁰ Second, who delivers a message can change its impact, over and above the content of the information given. Similarly, images included in a letter can have an important impact on response rates.¹¹ So we thought carefully about whom the letter was from and what logos to include. Third, people are likely to be busy and may forget, so we thought firms should remind them.¹²

These three considerations motivate the seven different treatments we tested, as summarised below. We order the treatments in terms of how the consumer may receive them – envelope first, then top to bottom as they read the letter.

Treatments

1. Envelope ▼	Adds a message to 'act quickly' to a plain envelope
2. FSA logo ▼	Uses the FSA logo in the letter head
3. Salient bullets ▼	Replaces the two bullet points at the top of the letter with more salient bullet points
4. Simplified ▼	Makes the body of the letter simpler and more concise, by reducing the text by 40%
5. Claims process ▼	Includes a sentence in bold explaining that the claims process would only take five minutes
6. CEO signature ▼	Uses the firm CEO's signature to sign the letter, instead of a generic 'Customer Team'
7. Reminder	Sends a second letter three to six weeks after the first. Reminder letters were identical to the original letter the customer received; apart from a message at the start of the letter stating that this was a reminder letter.

We expected that those treatments that acted most directly to reduce the transaction costs of responding would have the largest effect; so *Salient bullets*, *Simplified* and *Claims process* were expected to have the biggest effect. Since *Envelope* would act on the first barrier to responding – actually opening the letter – we expected it would also have a strong effect. We had fewer preconceptions about the impact of *FSA logo*, *CEO signature* and *Reminder*, simply because it was harder to second-guess how these different messengers would be perceived by consumers.

Our research is unable to pin down the precise channels whereby our treatments have an effect and we recognise that any effects of these treatments are not purely 'behavioural', in that they only act on imperfectly rational behaviour. Making information more salient can just be seen as reducing the transaction costs of opening or reading a letter. Images, logos, signatures and repeated letters might signal the importance of the communication to a rational consumer. So any effects could be explained by the model of rational consumers with limited time described above. The aim of this research is not to understand the precise mechanism. Rather, we aim to determine what works in practice, exclude what doesn't work and provide ideas for further research in this area. Context is important, so we will look to repeat the research in other settings to understand what works across situations.

How did we assign the different treatments?

Assigning the different treatments randomly is essential to ensure we are able to draw conclusions from our results. Randomisation ensures, especially with the large population that we work with, that we can consider the groups receiving different combinations of treatments to be balanced on unobservable characteristics (e.g. the likelihood that they were genuinely mis-sold), as well as observable characteristics. Therefore statistically significant differences in response rates between the groups, assuming that standard errors are correctly calculated, must be caused by the allocation of the treatment. This is the power of randomisation.

We control for observable characteristics of consumers by stratifying our population of just under 200,000 customers using the following six characteristics: age; gender; number of days since the purchased the product; type of product purchased; the number of monthly premiums paid by the consumer; and, the amount of potential redress (in pounds sterling, size of premium multiplied by number of months paid).

Stratifying across these variables generates over 1,800 different strata – homogenous sub-groups. Stratifying in this way allows us to look at the effect of our treatments on specific sub-groups, such as the elderly or women. It also increases our statistical power, which reduces our standard errors and allows us to test for differences between groups with more accuracy, though this is a minor concern given the sample size.

Our five letter treatments and the *Envelope* treatment are each assigned to 50% of the population. We 'cross' these treatments, generating 64 unique combinations (two to the power of six), which we assign across stratum. Because of capacity constraints, we are limited to sending approximately 20,000 reminder letters. These are randomly assigned across the 64 letter groups before the letters are sent. Table 2 shows mean and standard deviation figures for selected demographic variables for each of the seven treatments to demonstrate balance between our groups.

4. Results

Overall, our analysis shows that we can increase response rates by seven times to 11.9% by applying the right mix of treatments. Five out of seven of our treatments had a positive and significant impact on response rates, and this was robust to various different models and controls. *Salient Bullets* was our most effective treatment, followed by *Simplified*, *Claims process*, *Reminder* and *Envelope*. We also find some interaction effects between *Salient bullets* and *Simplified* and between *Salient bullets* and *Reminder*. This section discusses the analysis we undertake. It builds our statistical analysis from a simple model without interactions and looks in more detail at the impact of reminder letters before testing for the presence of interaction effects. It finishes by testing how the effect of our treatments varies with customer characteristics such as age and amount of redress owed. We use standard Ordinary Least Squares regressions throughout, as this is most appropriate when considering binary dependant variables. However, our results hold in general for other regression models such as probit.

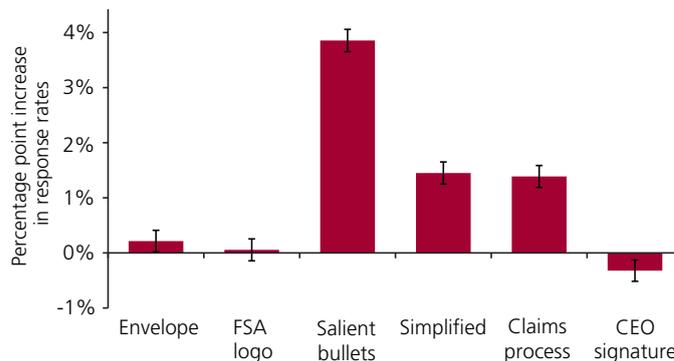
Basic statistical model

We first analyse the data assuming no interaction effects between treatments. Because the reminder letter is almost the same as the original letter, we omit the *Reminder* treatment from this analysis as we clearly expect interaction effects (e.g. people who receive a reminder letter with *Salient bullets* may respond more). This basic statistical model therefore captures some of the impact that in reality would be a result of the consumer being allocated to receive a *Reminder*. However, although the magnitude of impacts may not be precise, the relative ordering of effects should be valid, since the allocation of *Reminder* is random and independent to the allocation of all other treatments. See Annex 2 for the basic model specification.

Under this basic model, we find that the modelled response rate with no treatments is 1.9%. Compared with other redress exercises undertaken by the FSA, this response rate is particularly low. However, there may be a number of reasons why our particular sample may have very low response rates. First, some consumers had already been provided with a refund from the firm on their own initiative. Second, a number may also have been happy with the sales process and not felt in need of redress. Third, as they only held one product, the potential value of redress was low. Fourth, the relationship between the firm and the consumer had already ended, which may mean the firm has an out-of-date address or that the consumer is less likely to open the envelope. We also have data on average response rates from people outside our sample which shows that, for other customers, response rates were higher. For current customers, the response rate was approximately 20%. For cancelled customers with more than one product, this was approximately 5%. These figures support the view that our particular sample would be particularly difficult to reach, regardless of the letter that the firm sent.

Several of our letter treatments had a statistically significant impact. The relative impact, which we are interested in at this stage, varied greatly. The main treatment effects are reported in Chart 1 below, and the detailed regression outputs are included in Table 3.

Chart 1: Basic model of treatment effects



Notes: Includes 5% confidence intervals

Of our letter treatments, *Salient bullets* had the biggest positive effect (+3.9 percentage points), followed by *Simplified* (+1.4 percentage points) and *Claims process* (+1.4 percentage points). To put these effects into tangible terms, applying these three treatments across the entire sample would increase response rates by 7,500, 2,800 and 2,700 respectively.

Our results support other research findings showing that apparently minor changes to how information is presented can have a significant impact on response rates. What is most striking from our results is the magnitude of the effect and that some treatments performed significantly better than others. The impact of *Salient bullets* and *Simplified* highlight the importance of presenting the information clearly and succinctly to help the customer to quickly read and understand the decision that is required. They support the hypothesis that most readers are busy, that letters are skimmed and that even innocuous but irrelevant language can reduce engagement with the information. The strong impact of *Claims process* supports our initial hypothesis that readers need reassurance that making a claim will be quick and easy. In our control group, the absence of information about the process creates uncertainty. In the face of this uncertainty, customers may rationally substitute this missing information with past experiences or public perceptions of customer helplines and poor customer service in the financial services industry and decide it is not worth the effort.

Other treatments that we expected to have a strong effect had a significantly weaker or in some case no impact. *Envelope* has a small positive effect of 0.2 percentage points. Given the magnitude of our other results and the strength of the anti-procrastination message, we expected a stronger effect. One explanation for this is that in our treatment, we vary two different elements of the envelope at the same time and the effect of these two changes may be confounding each other. The treatment removed the firm's logo *and* added an additional message of importance and action. The presence of the logo could work to increase responses if consumers are aware of the brand and associate it with a product they previously purchased. Alternatively, it could decrease responses if consumers distrust the brand or fear that the envelope contains unwanted marketing material. Over this effect, we overlay the impact of the importance and anti-procrastination message. Unfortunately at this stage we are unable to disentangle these effects; this would be an interesting area for further work.

Interestingly, counter to our expectations, *CEO signature* reduced response rates by 0.3 percentage points. As noted, our research design does not enable us to provide a clear rationale for this. It could be that consumers' view the CEO's signature as disingenuous since they may perceive that he was not personally involved, reducing the credibility of the message. The CEO was a man, so it is possible that this led some people to reduce their trust in the letter.¹³

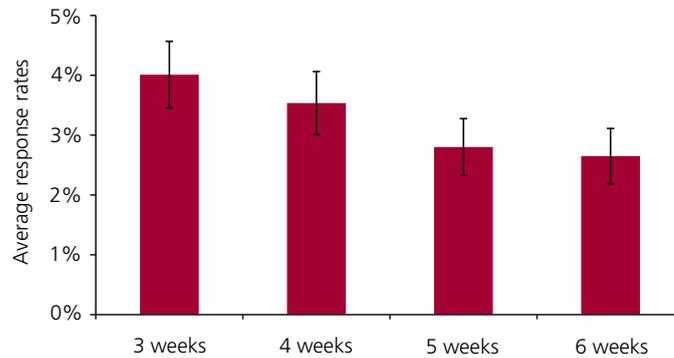
FSA Logo had no impact when compared to the average letter. Research conducted by the FSA shows that the brand is not very well known among the general public and it is likely that the logo itself is even less well known.¹⁴ This could to some extent be driving the results, although there may be other interactions with the logo that we cannot test for, for example between the Firm's logo and brand identity and the FSA's. Despite the possible interactions between the two organisations, the results provide evidence that can help decide when and how firms should use the FCA's logo in the future. Although firms operating in the UK must display information about whom they are authorised and regulated by, we have never required firms to include references to the FSA more prominently.¹⁵ Such results prove the importance of testing our policy proposals to ensure they are effective, before applying more generally.

Reminder letters

Having looked at the treatments we apply to the letters and envelope, we now look at the effect of sending consumers a letter reminding them of the situation and the need to respond. The *Reminder* treatment had the same combination of treatments as the first letter: randomly allocated to be sent at either three, four, five or six weeks after the initial letter; and, only if no response was recorded from the original mailing after three weeks.

This analysis seeks to answer the policy question: 'what is the impact of sending a reminder letter for those people who have not responded after three weeks?' Receiving a reminder letter increases response rates by 2.6 percentage points. It is worth highlighting that this is a slightly different sample group than the previous analysis and that the effect sizes are therefore not directly comparable. This is because the analysis only looks at the treatment effect for those people who had not responded to the initial letter after three weeks. As in the previous section, it also ignores any interactions that may occur between *Reminder* and the other treatments. The treatment effect is therefore only the effect of sending a notional 'average' reminder letter. We would expect an even stronger effect if we were to look at reminder letters with our most effective letter treatments.

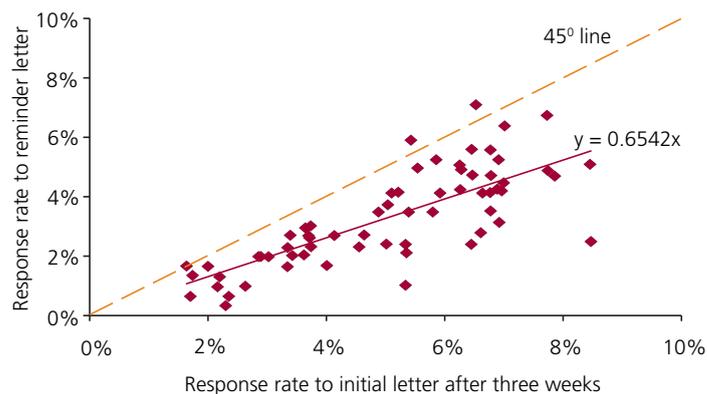
Looking at only those people who received reminder letters, we can also test the optimal number of weeks' delay before sending the reminder letter. This is possible because we also randomly varied the time delay between the initial mailing and the reminder mailing, from between three and six weeks. Among those who did not respond after three weeks, we find that an additional week's delay reduces response rates by 0.5 percentage points and that this is statistically significant. We can see this relationship if we look at the average response rates to the reminder letter, for those people who did not respond after three weeks, shown in Chart 2. One explanation for these findings is that for longer delays, consumers are less likely to recall the original letter, therefore reducing the strength of the importance signal that a second letter might provide. We were unable to test the impact of reminder letters at shorter intervals, although we suspect that sending letters earlier would increase response rates further; this would be an interesting area for further work.

Chart 2: Average response rates to reminder letters at different time intervals

Notes: Includes 5% confidence interval bars

Overall, our analysis of reminder letters indicates that they are a cost-effective way of increasing the amount of redress claimed by consumers. The reminder letters prompted a meaningful increase in response rates. Using information provided from the firm we are able to estimate the costs of each mailing letter. Based on this, and the estimated redress paid out as a result of the reminder letters, we can calculate the rate of return on the costs of mailing. Based on the notional average response rate from reminder letters, a reminder resulted in extra compensation worth 233% of the firm's mailing costs. Combining this analysis with the analysis of the most effective variations to the letter would increase the cost effectiveness of letters even further.

Chart 3 plots the response rate to the first mailing against response rate to the reminder letter for each of the 64 different letters (two versions of six treatments). It clearly shows that response to *Reminder* is correlated with the response rate to the original letter. Remembering that *Reminder* includes the same treatments as the original letter, this implies that there is some interaction between *Reminder* and one or some of the other treatments. The reminder letter provided a prompt, but better letters provided a better prompt. The next section will investigate this interaction in more detail.

Chart 3: Response rates by letter in initial and reminder mailings

Robustness and interaction effects

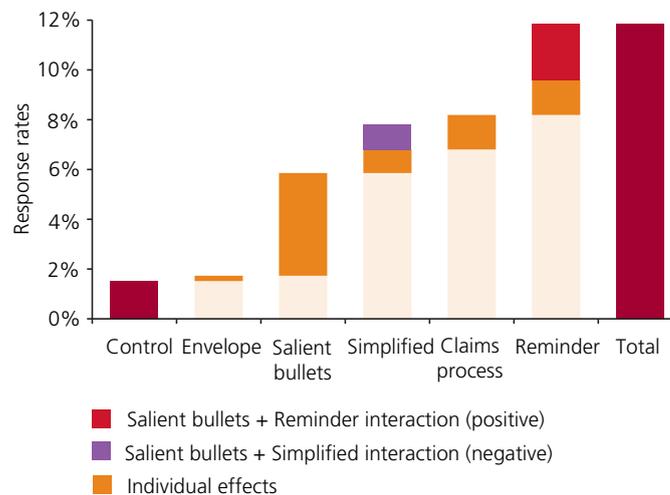
This section discusses in more detail the robustness of our results and the presence of interaction effects. We incorporate the reminder letter treatment into the analysis of the six letter treatments so that we can look precisely at the relative impact of these, and the potential interaction effects between them.

We test our basic statistical model against several alternative specifications to ensure they are robust in order to develop our analysis. First, as we would expect, our treatment effects do not materially change when we include control variables used in the randomisation of treatments (column 4 in Table 4). Second, we test several alternative specifications with different interaction effects (columns 2 and 3 in Table 3).

Interaction effects are when two or more treatments have a different effect on the outcome together, than when applied separately. In other words, the total effect of two interventions at the same time is either more or less than the sum of the individual effects. We do not have any strong expectations on the possible interactions between different treatments since we felt many of our treatments could act through multiple channels. However, we did reasonably expect that within the two broad types of changes to the letters – those that reduce transaction costs and those that signal importance – there could be negative interactions. For example, two signals of importance such as *FSA logo* and *CEO signature* might increase responses by 0.1% each, but in combination they only increase response rates by 0.15%. In addition, and as noted above, we expect a positive interaction between *Reminder* and other treatments.

We discard an extended statistical model including all interactions at all levels since we believe the higher-level interactions are not realistic and that these reduce the predictive power of our model. We then test the impact of including all interactions at the two-way level (see column 2 in Table 3). For the seven treatments, we find only two significant interactions. The interaction of *Salient bullets* and *Simplified* is significant and slightly negative; while the interaction of *Salient bullets* and *Reminder* is significant and positive. The interaction of *Salient bullets* and *Simplified* we interpret as indicating that they could work through a similar channel, namely making clear to the reader the precise nature of the letter. Having them individually increases response rates by significant amounts, but having them together does not increase response rates by as much as the simple addition of the two treatments. The positive interaction of *Salient bullets* and *Reminder* indicates that making the bullet points salient increased response rates for those people who received a first letter, but also made those people who received a second similar letter respond more as well. It is worth repeating here that the reminder letters were identical to the original letter the customer received in every way, apart from a message at the start of the letter saying that it was a reminder letter.

Finally, we use this to develop our primary statistical model (column 3 in Table 3), in which we include the negative interaction of *Salient bullets* and *Simplified* and the positive interaction of *Salient bullets* and *Reminder*. The results of this can give us an estimate of the cumulative impact of our effective treatments (see Chart 4).

Chart 4: Best letter treatment effects including interactions

The relative impact of treatments does not change significantly from those illustrated in the basic model discussed above, and we find all other treatments to be additive. Under this analysis we are able to assess the total combined impact of our effective treatments. We are able to improve response rates to almost 12%. This is equivalent to an additional 20,000 people responding to claim redress. This includes the average effect of sending a reminder letter. As seen in the previous section, we can improve this even further if we were to send everyone a letter after three weeks. This analysis uses our data to provide an estimate of effect sizes, but because of the way we have crossed treatments, we are also able to compare the subset of people who received the control version of the letter and those who received this precise combination of treatments. These two groups had response rates of 1.3% and 11.5% respectively. This comparison validates our modelled results.

When applying the individual treatments to the control letter (Table 8), *FSA logo* has a positive impact at the 5% significance level. While we cannot place too much emphasis on this result, given the negligible impact the treatment has elsewhere, it may be that the logo is acting to increase the perceived importance the reader places on the letter. When that letter is less direct the logo may act as a signal to persuade readers that they need to expend the extra effort to read the letter carefully. When combined with other treatments that improve the salience, readability or informational content of the letter, then this signal no longer has any significant impact.

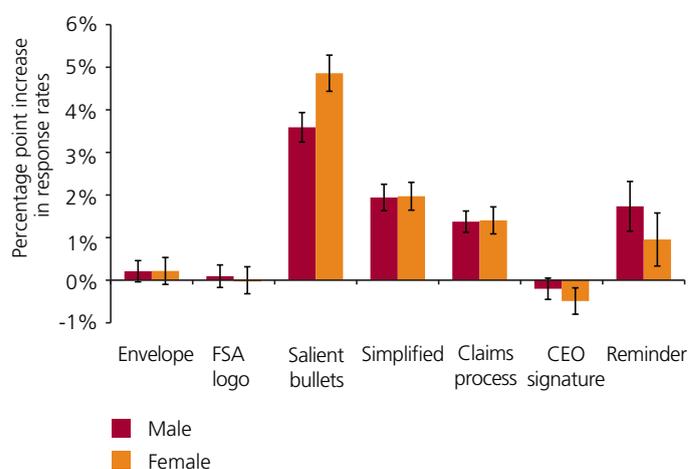
Testing for relationships with individual characteristics

As discussed, the richness of our dataset and the process of stratification allow us to investigate the particular characteristics and circumstances of individuals and how they respond. In this section we use our primary statistical model to look at the relationships with gender, age and refund amount, as well as attempting to find some proxies for unobserved characteristics such as wealth, consumer types and poverty.

We find that our treatment effects are broadly similar between **genders**. When we look at the sample as a whole, women were affected significantly more by *Salient bullets* than men (column 1 in Table 4). For women, *Salient bullets* increased response rates by 4.9 percentage points, while for men this was 3.6 percentage points.

If we look at each gender group separately, then we find only minor differences (Table 5). For men, we cannot statistically reject that the effect of *CEO signature* is zero; it has no impact on response rates. However, if we look at women, the treatment has a significant and negative effect (see Chart 5). On the particular impact of *CEO signature*, it is worth noting that the CEO is a man. There are a number of possible explanations. First, women could be negatively affected by introducing a personalised signature in general – that it is perceived as false or disingenuous and therefore reduces propensity to respond. Second, women could be less likely to trust or respond to messages delivered by men. There is some literature on gender and trust, although this does not provide conclusive evidence on the level of trust that women place in men.¹⁶ However, the literature indicates that the level of trust exhibited by women is variable and dependant on the precise experimental setting.¹⁷ It could be that under real life trials, gender differences do make a difference to the level of trust. Unfortunately in our research we are unable to determine whether this is a general or gender-specific effect. Further research that also varied the gender identity of the signatory would be able to provide some answers.

Chart 5: Gender differences in treatment effects

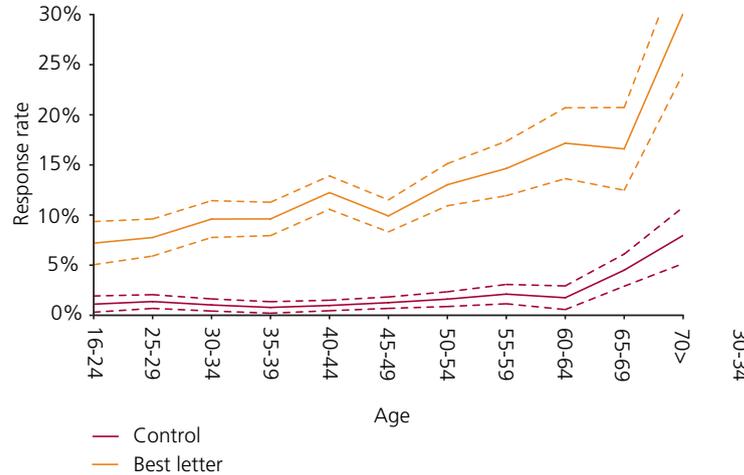


Notes: Percentage point effect sizes with 5% confidence interval bars

We are also able to investigate how response rates might change with **age** and test to see if any of our treatments have different effects at different ages. We find a convex relationship between average response rates and age. Age is negatively related to response rates while age-squared is positively related to response rates. Those over retirement age are most likely to respond, followed by the young and then lastly the middle-aged. Our hypothesis is that the middle-aged are most likely to be time-constrained by family and/or work commitments. This makes them less likely to open, read, consider or respond to the letters.

We find that *Salient bullets*, *Simplified*, *Claims process* and *Reminder* all have a positive and statistically significant interaction with age (Table 4), suggesting that the older the recipient, the bigger the effect of the treatment. Chart 6 plots a line of estimated response rates at discrete intervals based on age under two circumstances: the response rate without any treatments and the response rates for letters that include these four treatments. The regression is run at discrete intervals and this means that the confidence intervals are wider than if we were to plot a smooth relationship of age and response rates. Despite this, we think the chart clearly illustrates how our treatments reverse the low response rates seen in the middle-aged. We interpret this as possible support for our theory that by increasing the speed with which busy readers can assess the relevant information we are able to improve the decisions made by this group.

Chart 6: Control vs. best letter, response rates by age

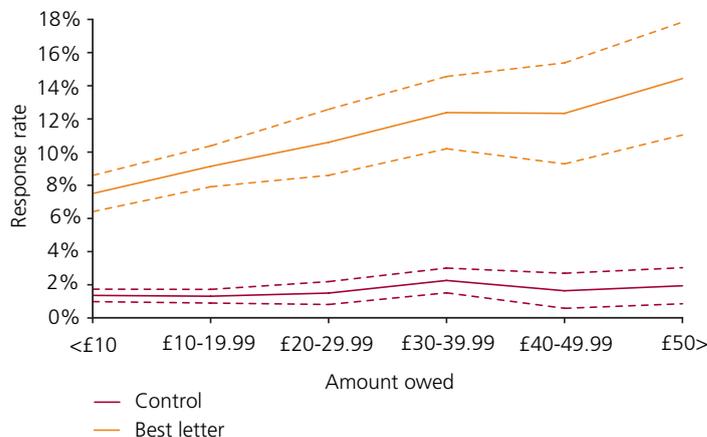


Notes: Average response rates at discrete intervals with 5% confidence bands

Economic theory would suggest that for any given level of transaction costs, higher **pay-offs** result in higher response rates. There is some significant correlation between the redress amount owed and the response rate from the average letter, each £10 worth of redress is correlated with an increase in response rates of 0.05 percentage points. We are unable to determine causation, since there may be countervailing unobserved variables that increase with total amount of redress (such as income) that would make people less likely to respond.

We find that some of our treatments improved outcomes more for those who were due more redress. This was true for *Salient bullets*, *Simplified* and *Reminder*, where the interaction with total amount owed is statistically significant and positive (Table 4). From a policy perspective, this is an important finding, since firms’ letters should rightly target those consumers who are owed the most redress. As in the analysis for age, Chart 7 displays estimated response rates on a number of different subsets of the population comparing response rates in the absence of treatments and with the three treatments we know interact with age. There is a similar pattern of interactions between response rates and monthly premiums but the monthly premiums do not significantly affect response rates in general. This is interesting, since the monthly premium is displayed on the letter and is therefore potentially more salient, but does not affect average response rates, only increasing response rates through our treatments.

Chart 7: Control vs. best letter, response rates by total amount owed



Notes: Average response rates at discrete intervals with 5% confidence bands

Finally, we are able to match each individual consumer to several **other pieces of demographic and consumer data**. By doing so we attempt to identify specific characteristics that may help explain whether certain groups respond more or less to these letters and whether our treatments have stronger effects on certain groups. We look at the Experian Financial Strategy Segments (FSS), Indices of Multiple Deprivation (IMD) and average mortgage value (Table 9). For each of these, we note that the data are averages based on postcode or other geographic regions. This may lead to significant measurement errors, especially when the characteristic of interest is highly variable in the postcode area. Our analysis is therefore speculative at this stage, although this provides a potential area for future work. We would like to test for further relationships with, for example, average educational attainment, but the data we have found relates to the 2001 census. When the full 2011 census data is available, it may be possible to extend our analysis.

FSS groups segments the population based on age, income, wealth and other indicators.¹⁸ We can match FSS segments to individuals based on postcode information. We find that there appears to be no strong relationship between particular FSS groups and response rates or treatment effects. Belonging to the 'Sunset Security' group has a significant correlation with increased response rates; being part of this group would indicate an increase in response rate of 1 percentage point. However, this effect disappears if we also control for age, suggesting that the FSS adds nothing over the variables that we already have.

The Indices of Multiple Deprivation (IMD) is a government index that looks at the relative deprivation in small geographical areas based on a number of factors affecting living standards including income, employment, education and crime.¹⁹ We test whether deprivation may make consumers more or less likely to respond. There may be several reasons why deprivation may alter response rates. Areas with high levels of unemployment or low levels of income may have higher response rates in general, as they may have lower time costs of responding relative to those who are employed or on higher incomes. Areas with lower levels of education may respond better to treatments that reduce the cognition costs of processing the information. As with the FSS groups, although we find some relationships between these indices and response rates, if we also control for basic characteristics such as age and total amount of redress, then such effects disappear.

We match our results to average mortgage values from 2012 for the first section of each individual's postcode. An increase in average mortgage values of £10,000 is correlated with a reduction in response rates of approximately 0.06 percentage points. This fits with a rational model where we would expect the wealthy to respond less due to their higher value of time reducing the relative net benefit of responding.

5. Conclusions

Our research has confirmed other recent findings in UK public policy that show that subtle changes to the presentation of information can have large effects. The cumulative effect of our treatments is statistically and economically meaningful, even in this particular context when the payoffs are small. Of the three broad areas our research tested, our results show that tailoring the letters to increase the readability of letters is the most effective. This, combined with the evidence that the middle-aged respond the least, suggests that when people do open these letters, they process the information very quickly. Reminding people helps and we think this is likely to work as a prompt as well as a signal of importance. Finally, personalising the messenger or including a credible third party does not improve response rates. We believe that the broad messages from our results can be applied in other redress exercises and we are already doing so to help improve consumer outcomes.

The broad range in consumers' response rates to our different treatments supports the need for testing and experimentation in the development of public policy. As has been found in numerous other trials, it is very difficult to predict what will have the biggest impact or indeed any impact at all. Similarly, in a regulatory context, proposed policy interventions do not always have the expected consequences. RCTs can allow us to identify unexpected consequences among a smaller population (though there are some well-known caveats) and therefore avoid any wider detrimental impacts of misguided interventions. In our opinion, RCTs especially field trials where we are able to observe real world responses, are especially useful in overcoming conventional wisdom and challenging policymakers' own prior beliefs about useful interventions.

In the context of communications and the provision of information, there are a number of areas where we can extend our analysis. We aim to test further changes to consumer communications, in particular methods to help consumer reduce procrastination and the effect of different channels of communication (such as emails or text messages). Where possible, we also aim to test and improve our own communications to firms to understand how firms respond to our letters. We also believe that applying the same methods to disclosures about products could also make this information more effective and ensure that financial consumers are informed, and therefore empowered, as best as possible.

Tables

Table 1: Summary statistics

(means or proportions with standard deviations in parentheses)

	Full sample	Responded and received refund	Responded, but did not request refund
Age in years	43.56 (12.57)	47.56 (13.62)	51.49 (14.99)
Proportion male	0.58	0.56	0.60
Total amount owed £	21.33 (16.92)	23.97 (18.49)	21.70 (16.90)
Monthly premium £	9.60 (3.77)	9.82 (3.81)	9.08 (3.53)
Number of monthly payments	2.27 (1.55)	2.48 (1.63)	2.43 (1.63)
Days since purchase	496 (197)	493 (199)	473 (186)
Mailing wave			
1	0.20	0.20	0.21
2	0.20	0.20	0.20
3	0.20	0.20	0.21
4	0.20	0.20	0.19
5	0.20	0.20	0.19
Observations	195140	10213	2580
Proportion of total sample	1.00	0.05	0.01

Table 2: Experimental summary

(means or proportions with standard deviations in parentheses)

	Full sample	Envelope	FSA logo	Salient bullets	Simplified	Claims process	CEO signature	Reminders
Age in years	43.56 (12.57)	43.56 (12.57)	43.58 (12.59)	43.58 (12.60)	43.56 (12.57)	43.55 (12.55)	43.57 (12.60)	43.57 (12.58)
Proportion male	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58
Total amount owed £	21.33 (16.92)	21.31 (16.88)	21.35 (16.96)	21.35 (16.95)	21.34 (16.95)	21.35 (16.94)	21.32 (16.89)	21.45 (16.69)
Monthly premium £	9.60 (3.77)	9.59 (3.77)	9.60 (3.78)	9.59 (3.76)	9.59 (3.77)	9.60 (3.77)	9.59 (3.77)	9.66 (3.76)
Number of monthly payments	2.27 (1.55)	2.28 (1.55)						
Days since purchase	496 (197)	499 (196)						
Mailing wave								
1	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
2	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
3	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
4	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
5	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Observations	195140	97584	97529	97506	97608	97542	97553	20232
Proportion of total sample	1.00	0.50	0.50	0.50	0.50	0.50	0.50	0.10

Table 3: Main treatment effects

	Basic model (1)	Two-way interactions model (2)	Primary model (3)	Primary model with Controls (4)
Envelope	0.21** (0.10)	-0.04 (0.22)	0.21** (0.10)	0.22** (0.10)
FSA logo	0.06 (0.10)	0.07 (0.21)	0.05 (0.10)	0.05 (0.10)
Salient bullets	3.86*** (0.10)	4.10*** (0.24)	4.13*** (0.14)	4.11*** (0.14)
Simplified	1.45*** (0.10)	1.71*** (0.24)	1.95*** (0.12)	1.95*** (0.12)
Claims process	1.39*** (0.10)	1.25*** (0.22)	1.39*** (0.10)	1.39*** (0.10)
CEO signature	-0.32*** (0.10)	-0.06 (0.22)	-0.32*** (0.10)	-0.32*** (0.10)
Reminder		1.15*** (0.45)	1.41*** (0.22)	1.41*** (0.22)
Salient bullets*Simplified		-1.01*** (0.20)	-1.01*** (0.20)	-1.00*** (0.20)
Reminder * Salient bullets		2.25*** (0.40)	2.25*** (0.40)	2.26*** (0.40)
Age				0.13*** (0.00)
Total refund				0.06*** (0.01)
Gender=Male				-0.71*** (0.10)
Constant	1.91*** (0.12)	1.63*** (0.19)	1.52*** (0.12)	-5.87*** (0.38)
Observations	195,140	195,140	195,140	195,140
Number of strata	1,811	1,811	1,811	1,811

Notes: OLS regressions i.e. linear probability. Dependent variable is responded and received refund. All coefficients and standard errors multiplied by 100, so coefficients on treatments dummies are percentage points. *** denotes significance at 1%, ** at 5%, and * at 10%. Robust standard errors adjusted for clusters in stratum. Effect sizes with clustered standard errors in parentheses.

Table 4: Interactions between treatments and consumer characteristics

Gender interactions (1)	Age interactions (2)	Premium interactions (3)	Total redress interactions (4)
Male* Envelope (-0.0066 (0.21)	Age* Envelope 0.013 (0.0088)	Premium* Envelope 0.00002 (0.027)	Total redress *Envelope -0.0066 (0.007)
Male* FSA logo (0.095 (0.21)	Age* FSA logo -0.00619 (0.0091)	Premium* FSA logo 0.0136 (0.028)	Total redress *FSA logo 0.00263 (0.007)
Male* Salient bullets (-1.00*** (0.21)	Age* Salient bullets 0.0232** (0.0092)	Premium* Salient bullets 0.090*** (0.028)	Total redress *Salient bullets 0.045*** (0.0063)
Male* Simplified (0.36* (0.21)	Age* Simplified 0.0487*** (0.0090)	Premium* Simplified 0.090*** (0.027)	Total redress *Simplified 0.016** (0.0068)
Male* Claims process (-0.03 (0.20)	Age* Claims process 0.0446*** (0.0090)	Premium* Claims process 0.0156 (0.027)	Total redress *Claims process 0.016 (0.0063)
Male* CEO Signature (0.29 (0.20)	Age* CEO Signature -0.00131 (0.0092)	Premium* CEO Signature -0.036 (0.027)	Total redress * CEO Signature 0.0032 (0.0066)
Male* Reminder (0.17 (0.40)	Age* Reminder 0.065*** (0.017)	Premium* Reminder 0.138*** (0.0508)	Total redress *Reminder 0.041*** (0.012)
Constant (1.67*** (0.19)	5.2*** 0.72	1.69*** (0.339)	1.25*** (0.19)
Observations 195,140	195,140	195,140	195,140
Number of strata 1,811	1,811	1,811	1,811

Notes: Only interaction coefficients shown. OLS regressions, i.e. linear probability. Dependent variable is responded and received refund. All coefficients and standard errors multiplied by 100, so coefficients on treatments dummies are percentage points. *** denotes significance at 1%, ** at 5%, and * at 10%. Robust standard errors adjusted for clusters in stratum. Effect sizes with clustered standard errors in parentheses. For each regression, the interacted variables are also included as individual variables.

Table 5: Treatment effects across gender

	Male (1)	Female (2)
Envelope	0.21* (0.13)	0.21 (0.16)
FSA logo	0.09 (0.13)	-0.0014 (0.16)
Salient bullets	3.60*** (0.18)	4.86*** (0.22)
Simplified	1.94*** (0.16)	1.97*** (0.17)
Claims process	1.37*** (0.13)	1.40*** (0.16)
CEO signature	-0.20 (0.13)	-0.49*** (0.16)
Reminder	1.7*** (0.30)	0.96*** (0.32)
Salient bullets* Simplified	-0.67** (0.27)	-1.47*** (0.30)
Reminder * Salient bullets	1.75*** (0.51)	2.97*** (0.62)
Constant	1.47*** (0.16)	1.59*** (0.19)
Observations	113,035	82,105
Number of strata	1,292	1,003

Notes: OLS regressions i.e. linear probability. Dependent variable is responded and received refund. All coefficients and standard errors multiplied by 100, so coefficients on treatments dummies are percentage points. *** denotes significance at 1%, ** at 5%, and * at 10%. Robust standard errors adjusted for clusters in stratum. Effect sizes with clustered standard errors in parentheses.

Table 6: Treatment effects across age groups

	16-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	>70
Envelope	0.157 (0.36)	-0.211 (0.277)	-0.0409 (0.258)	0.267 (0.249)	-0.0353 (0.236)	0.35 (0.251)	-0.023 (0.304)	0.684* (0.403)	1.17** (0.548)	1.79*** (0.682)	-1.34 (0.954)
FSA logo	-0.0219 (0.349)	-0.169 (0.285)	0.00185 (0.268)	0.442* (0.239)	0.084 (0.231)	-0.309 (0.252)	0.614* (0.313)	-0.3 (0.419)	1.21** (0.55)	-1.03 (0.715)	-1.72* (0.959)
Salient bullets	3.97*** (0.495)	3.69*** (0.386)	3.82*** (0.361)	4.26*** (0.331)	3.86*** (0.308)	3.98*** (0.357)	4.67*** (0.407)	3.60*** (0.546)	4.90*** (0.693)	4.89*** (0.955)	4.68*** (1.459)
Simplified	1.23*** (0.383)	1.40*** (0.299)	1.21*** (0.283)	1.25*** (0.274)	1.78*** (0.251)	1.87*** (0.279)	3.12*** (0.365)	2.19*** (0.475)	3.32*** (0.633)	4.15*** (0.964)	2.06 (1.327)
Claims process	0.645* (0.33)	0.698** (0.285)	1.02*** (0.271)	1.23*** (0.25)	1.19*** (0.237)	1.52*** (0.245)	1.19*** (0.317)	2.01*** (0.416)	2.12*** (0.55)	2.19*** (0.703)	4.93*** (0.989)
CEO signature	-0.622** (0.316)	-0.603* (0.309)	-0.138 (0.265)	-0.28 (0.247)	-0.236 (0.229)	-0.451* (0.25)	-0.308 (0.296)	0.579 (0.404)	-0.596 (0.536)	-0.863 (0.704)	-1.14 (1.039)
Reminders	0.787 (0.638)	1.28** (0.61)	0.906* (0.52)	1.47*** (0.517)	1.20** (0.479)	1.87*** (0.565)	0.486 (0.639)	1.11 (0.863)	3.02** (1.174)	-0.23 (1.45)	6.13*** (2.3)
Salient bullets* Simplified	-1.54** (0.711)	-1.07* (0.579)	-0.921* (0.536)	-1.05** (0.492)	-0.472 (0.469)	-0.554 (0.506)	-2.10*** (0.622)	0.0378 (0.801)	-1.18 (1.06)	-2.52* (1.488)	-1.15 (2.116)
Reminders* Salient bullets	0.993 (1.262)	0.386 (1.099)	2.53** (0.992)	1.68* (0.984)	3.70*** (0.917)	-0.035 (0.992)	4.04*** (1.171)	3.59** (1.581)	3.25 (2.05)	3.61 (2.45)	5.5 (3.487)
Constant	1.11*** (0.413)	1.37*** (0.349)	1.03*** (0.308)	0.782*** (0.292)	0.978*** (0.267)	1.26*** (0.29)	1.61*** (0.375)	2.11*** (0.491)	1.75*** (0.601)	4.50*** (0.817)	7.97*** (1.43)
Observations	11668	16632	21620	25214	30579	29697	22842	14813	10553	6973	4549
Adjusted R-squared	0.00893	0.00868	0.011	0.0117	0.0142	0.0109	0.012	0.0112	0.0148	0.00972	0.018

Notes: OLS regressions, i.e. linear probability. Dependent variable is responded and received refund. All coefficients and standard errors multiplied by 100, so coefficients on treatments dummies are percentage points. *** denotes significance at 1%, ** at 5% and * at 10%. Robust standard errors adjusted for clusters in stratum. Effect sizes with clustered standard errors in parentheses. All coefficients and standard errors multiplied by 100. For each regression, the interacted variables are also included as individual variables.

Table 7: Treatment effects across total amount of redress

	<£10	£10-19.99	£20-29.99	£30-39.99	£40-49.99	£50>
Envelope	0.389** (0.1595)	0.181 (0.1778)	0.353 (0.2904)	-0.0713 (0.3155)	-0.00338 (0.4644)	-0.16 (0.4855)
FSA logo	-0.0677 (0.1595)	0.00786 (0.1778)	0.289 (0.2906)	0.387 (0.3153)	-0.25 (0.4642)	0.143 (0.4851)
Salient bullets	3.36*** (0.2154)	3.98*** (0.2362)	4.77*** (0.3932)	4.46*** (0.4318)	5.64*** (0.6212)	5.68*** (0.633)
Simplified	1.57*** (0.184)	2.21*** (0.2024)	2.26*** (0.3326)	1.25*** (0.352)	2.39*** (0.5121)	3.12*** (0.5588)
Claims process	1.08*** (0.1595)	1.58*** (0.1777)	1.72*** (0.2905)	1.09*** (0.3155)	1.48** (0.4654)	1.85*** (0.485)
CEO signature	-0.209 (0.1595)	-0.549*** (0.1778)	-0.246 (0.2904)	-0.578* (0.3155)	0.444 (0.465)	-0.237 (0.4852)
Reminders	1.14*** (0.3473)	1.12*** (0.3746)	1.33** (0.602)	2.30*** (0.6906)	1.73* (0.979)	2.67** (1.1268)
Salient bullets* Simplified	-0.919*** (0.3188)	-1.15*** (0.3561)	-1.28** (0.5801)	-0.0774 (0.6318)	-1.33 (0.9317)	-1.55 (0.9668)
Reminders* Salient bullets	1.49** (0.6077)	2.30*** (0.6729)	2.55** (1.0747)	2.34* (1.2186)	3.18* (1.7207)	3.85** (1.9324)
Constant	1.36*** (0.1916)	1.31*** (0.2102)	1.50*** (0.352)	2.26*** (0.381)	1.64*** (0.5393)	1.94*** (0.5546)
Observations	65363	58945	26574	21938	11094	11226
Adjusted R-squared	0.00797	0.0113	0.0122	0.0132	0.0156	0.0167

Notes: OLS regressions, i.e. linear probability. Dependent variable is responded and received refund. All coefficients and standard errors multiplied by 100, so coefficients on treatments dummies are percentage points. *** denotes significance at 0.1%, ** at 5% and + at 10%. Robust standard errors adjusted for clusters in stratum. Effect sizes with clustered standard errors in parentheses.

Table 8: Treatment effects when applied individually to control letter

	Envelope	FSA logo	Salient bullets	Simplified	Claims process	CEO signature	Reminders
Envelope	0.659** (0.31)						
FSA logo		0.795** (0.32)					
Salient bullets			4.85** (0.47)				
Simplified				2.51*** (0.39)			
Claims process					1.49*** (0.36)		
CEO signature						0.209 (0.3)	
Reminders							2.60** (1.09)
Constant	1.17*** (0.20)	1.17*** (0.20)	1.17*** (0.20)	1.17*** (0.2)	1.17*** (0.2)	1.17*** (0.2)	1.17*** (0.2)
Observations	5800	5791	5794	5778	5778	5785	3055
Adjusted R-squared	0.000552	0.000836	0.0162	0.00631	0.00273	-0.0000869	0.00413

Notes: OLS regressions, i.e. linear probability. Dependent variable is responded and received refund. All coefficients and standard errors multiplied by 100, so coefficients on treatments dummies are percentage points. *** denotes significance at 0.1%, ** at 5% and + at 10%. Robust standard errors adjusted for clusters in stratum. Effect sizes with clustered standard errors in parentheses.

Table 9: Additional demographic controls

	Primary model with Controls and IMD	Primary model with Controls and Experian FSS	Primary model with Controls and Mean mortgage value
Envelope	0.334*** (0.11)	0.234** (0.10)	0.221** (0.10)
FSA logo	0.0153 (0.11)	0.049 (0.10)	0.046 (0.10)
Salient bullets	4.05*** (0.15)	4.08*** (0.14)	4.11*** (0.14)
Simplified	2.0*** (0.13)	1.97*** (0.12)	1.95*** (0.12)
Claims process	1.38*** (0.11)	1.41*** (0.11)	1.39*** (0.10)
CEO signature	-0.316*** (0.11)	-0.359*** (0.10)	-0.320*** (0.10)
Reminders	1.36*** (0.24)	1.40*** (0.22)	1.40*** (0.22)
Salient bullets * Simplified	-1.08*** (0.23)	-0.982*** (0.21)	-0.999*** (0.20)
Reminders * Salient bullets	2.53*** (0.45)	2.28*** (0.42)	2.27*** (0.40)
Age	0.138*** (0.01)	0.136*** (0.01)	0.135*** (0.005)
Total refund	0.0484*** (0.0036)	0.0488*** (0.0033)	0.0480*** (0.0032)
Gender=Male	-0.745*** (0.12)	-0.719*** (0.11)	-0.716*** (0.10)
IMD 2011	-0.00335 (0.0035)		
FSSG11==2		-0.616* (0.36)	
FSSG11==3		-0.61 (0.43)	

	Primary model with Controls and IMD	Primary model with Controls and Experian FSS	Primary model with Controls and Mean mortgage value
FSSG11==4		-0.249 (0.41)	
FSSG11==5		0.139 (0.38)	
FSSG11==6		-0.659 (0.53)	
FSSG11==7		-0.0594 (0.41)	
FSSG11==8		-0.226 (0.34)	
FSSG11==9		-0.538 (0.35)	
FSSG11==10		0.116 (0.39)	
FSSG11==11		-0.0255 (0.40)	
FSSG11==12		-0.112 (0.46)	
FSSG11==13		0.301 (0.37)	
FSSG11==14		-0.00551 (0.36)	
Mean Mortgage value (£1000)			-0.000554** (0.0003)
Constant	-5.03*** (0.30)	-4.83*** (0.41)	-4.84*** (0.27)
Observations	155247	183638	194931
Adjusted R-squared	0.0184	0.0187	0.0183

Annex 1:

Comparison of letters and envelopes

Telephone [REDACTED]
Mon-Fri: 9am-5.30pm
Sat: 9am-1pm

Policy reference: [REDACTED]

- **You bought [REDACTED] with [REDACTED] which you cancelled**
- **This letter contains important information, please read it carefully and respond if necessary**

Dear [REDACTED]

In September 2010, you took advantage of our offer of [REDACTED] when you purchased our [REDACTED]. You cancelled the policy in February 2011 and the premium you were paying at the time was £3.99.

We are committed to the ongoing improvement of the products we offer and our service and we welcome all feedback from our customers. Therefore, we would like to know if you cancelled your policy because you were in any way dissatisfied with the product or the way it was sold to you. It may be that you are entitled to a refund.

What you need to do

If you would like to discuss this with us please feel free to speak to one of our friendly team on [REDACTED] between 9am - 5.30pm, Monday to Friday or 9am - 1pm on Saturdays quoting your policy reference [REDACTED].

Yours sincerely,

Customer Services Team

Telephone [REDACTED]
Mon-Fri: 9am-5.30pm
Sat: 9am-1pm

 2

Policy reference: [REDACTED]

- You could be entitled to a refund on the [REDACTED] you bought
- If you have any concerns about the product or how we sold it to you, please call [REDACTED]

Dear

In May 2011, you bought our [REDACTED], including our opening offer of [REDACTED]. You cancelled the policy in December 2011 and the premium you were paying at the time was £4.99.

We would like to know if you cancelled your policy because you were in any way dissatisfied with the product or the way it we sold it to you. If so, you could be entitled to a refund.

What you need to do

If you have any concerns, please call us on [REDACTED] between 9am - 5.30pm, Monday to Friday or 9am - 1pm on Saturdays quoting your policy reference [REDACTED]

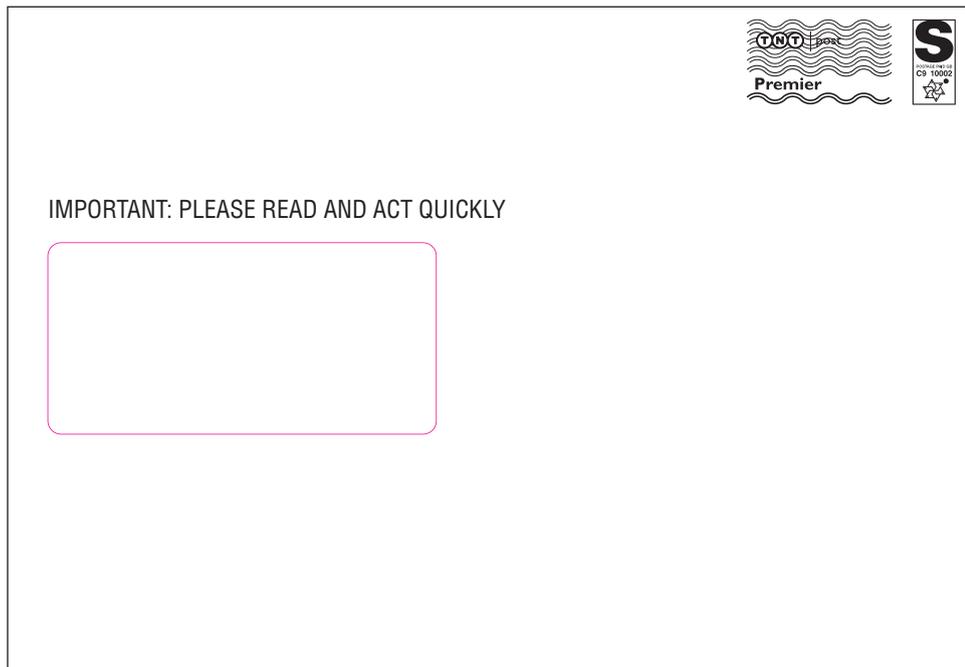
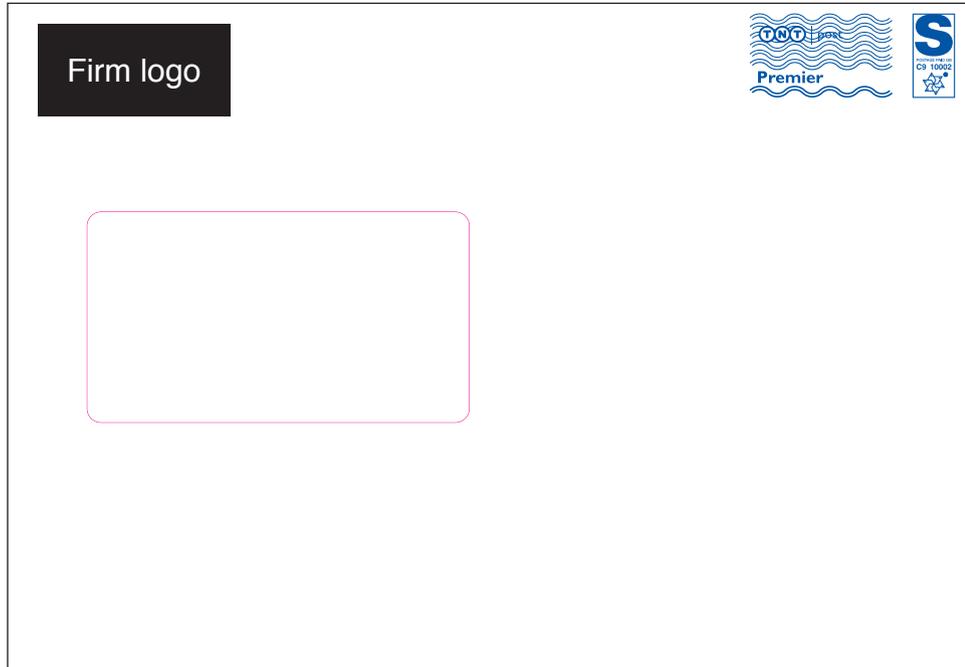
We will review your case on the telephone and you will know if you are due a refund within five minutes.

Yours sincerely,

[REDACTED]

CEO 6

2. FSA logo
3. Salient bullets
4. Simplified
5. Claims process
6. CEO signature



1. Envelope

Annex 2: Model equations

Basic regression (excluding the effect of reminder)

$$y_i = \sum_{t=1}^6 \beta^t T_i^t + \varepsilon_i$$

Where y_i is either 1 if individual i responded or 0 otherwise; t is the set of treatments we tested; T_i^t is 1 if individual i was assigned treatment t , and 0 otherwise; and, ε is the stratum clustered error term.

Regression with two-way interactions

$$y_i = \sum_{t=1}^7 \beta^t T_i^t + \sum_{t=1}^7 \sum_{\substack{s=2 \\ s>t}}^7 \beta^{t,s} T_i^t T_i^s + \varepsilon_i$$

Regression with interactions at all levels

$$y_i = \sum_{t=1}^7 \beta^t T_i^t + \sum_{t=1}^7 \sum_{\substack{s=2 \\ s>t}}^7 \beta^{t,s} T_i^t T_i^s + \sum_{t=1}^7 \sum_{\substack{s=2 \\ s>t}}^7 \sum_{\substack{r=3 \\ r>s}}^7 \beta^{t,s,r} T_i^t T_i^s T_i^r + \dots$$

$$\dots + \beta^{1,2,3,4,5,6,7} T_i^1 T_i^2 T_i^3 T_i^4 T_i^5 T_i^6 T_i^7 + \varepsilon_i$$

References

- Amos, C. and Paswan, A. (2009). Getting past the trash bin: Attribution about envelope message, envelope characteristics, and intention to open direct mail. *Journal of Marketing Communications*, 15(4), 247-265
- Bargh, J.A. (1982). Attention and automaticity in the processing of self-relevant information. *Journal of Personality and Social Psychology*, 43(3), 425-436
- Bateman H., Eckert, C., Geweke, J., Louviere, J., Satchell, S., and Thorp, S. (2011). Financial competence, risk presentation and retirement portfolio preferences. *UNSW Australian School of Business Research Paper*, No. 2011ACTL03
- Bertrand, M., Karlan, D., Mullainathan, S., Shafir, and Zinman, J. (2010). What's advertising content worth? Evidence from a Consumer Credit Marketing Field Experiment. *Quarterly Journal of Economics*, 125 (1), 263-306
- Bettinger E, Long B, Oreopoulos P and Sanbonmatsu L (2012). The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment. *Quarterly Journal of Economics*, 127 (3), 1205-1242
- Buchan, N. R., Croson, R.T.A, Solnick, S. (2008). Trust and gender: An examination of behaviour and beliefs in the investment game. *Journal of Economic Behavior & Organization*, 68 (3-4), 466-476
- Cabinet Office (2012). Applying behavioural insights to reduce fraud, error and debt.
- Card, D., DellaVigna, S., Malmendier, U. (2011). The Role of Theory in Field Experiments. *Journal of Economic Perspectives, American Economic Association*, 25(3), 39-62
- Chetty, R, Looney, A. and Kroft, K. (2009). Salience and Taxation: Theory and Evidence. *American Economic Review*, 99(4), 1145-77.
- Cialdini, R, B. (2007). *Influence: The Psychology of Persuasion. Revised Edition*. HarperBusiness
- Croson, R., Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47 (2), 448-474
- Deaton, A. (2010). Instruments, Randomization, and Learning about Development. *Journal of Economic Literature*, 48(2), 424-455
- DellaVigna, S. and Gentzkow, M. (2010). Persuasion: Empirical Evidence. *Annual Review of Economics, Annual Reviews*, 2(1), 643-669
- DellaVigna, S., (2009), Psychology and Economics: Evidence from the Field, *Journal of Economic Literature, American Economic Association*, 47(2), 315-372

- Dohrmann, T. and Pinshaw, G. (2009). The Road to Improved Compliance: A Benchmarking Study of Tax Administrations 2008–9. *McKinsey & Company*
- Eckel, C., Wilson, R. (2000). Whom to trust? Choice of partner in a trust game. Unpublished. Available at http://www.ruf.rice.edu/~rkw/RKW_FOLDER/Eckel_Wilson_ESA_00.pdf. [Accessed 15 January 2013]
- Feld, S., Frenzen, H., Krafft, M., Peters, K. and Verhoef, P. C., (2013). The Effect of Mailing Design Characteristics on Direct Mail Campaign Performance. *International Journal of Research in Marketing*, 30(2), forthcoming
- Financial Services Authority (2012). Journey to the FCA.
- Fryer R. (2013). Teacher Incentives and Student Achievement: Evidence from New York City Public Schools. *Journal of Labor Economics*, 31(2), forthcoming
- Garner R., (2005). Post-It[®] note persuasion: a sticky influence. *Journal of Consumer Psychology*, 15(3), 230-237
- Graves, P. (2010). *Consumer.ology: The Market Research Myth, the Truth about Consumers and the Psychology of Shopping*. London: Nicholas Brealey Publishing
- Heckman, J. and Urzua, S. (2009). Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics*, 156(1), 27-37
- Imbens, G. (2009). Better late than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature*, 48(2), 399-423
- Kahneman, D. (2012). *Thinking Fast and Slow*. London: Allen Lane
- Karlan, D., McConnell, M., Mullainathan, S. and Zinman, J., (2010). Getting to the Top of Mind: How Reminders Increase Saving. *Yale University Economic Growth Center Discussion Paper No. 988; Yale Economics Department Working Paper No. 82*.
- Laibson, D., (2001). A Cue-Theory of Consumption. *The Quarterly Journal of Economics*, 116(1), 81-119
- Langton W.E., (1992). Comprehension of illustrated text: pictures help to build mental models. *Journal of Memory and Language*, 31(2) 129-151
- Leah-Martin, V. (2012). The role of gender and race in cooperation and trust: experimental examination of identity based strategic interaction. *University of California*. Available at [http://www.econ.ucsd.edu/~rstarr/191AFall2012/191%20Final%20Paper%20-%20Leah-Martin%20\(2012\).pdf](http://www.econ.ucsd.edu/~rstarr/191AFall2012/191%20Final%20Paper%20-%20Leah-Martin%20(2012).pdf). [Accessed 15 January 2013]
- Leslie, T.F., Griffin, D.H., McGovern, P.D. and Raglin, D.A. (2004). Measuring the Impact of Questionnaire and Envelope Messages on Respondent Behavior. *U. S. Bureau of the Census*. Presented at American Statistical Association conference
- Leventhal H., Singer R. and Jones S. (1965). Effects of fear and specificity of recommendation upon attitudes and behaviour. *Journal of Personality and Social Psychology*, 2(1), 20-29

- Levie WH and Lentz R (1982). Effects of text illustrations: a review of research. *Education and Communication Technology Journal*, 30(30), 195-232
- Levitt, S. and List, J. (2009). Field Experiments in Economics: The Past, The Present, and The Future. *European Economic Review*, 53(1), 1-18.
- Mann T. and Ward A. (2007). Attention, self-control, and health behaviors. *Current Directions in Psychological Science*, 16(5), 280-283
- Orbell, J., Dawes, R. and Schwartz-Shea, R. (1994). Trust, social categories, and individuals: The case of gender. *Motivation and Emotion*, 18(2), 109-128
- Postma O.J. and Brokke M. (2002). Personalisation in practice: the proven effects of personalisation. *Journal of Database Marketing* 9(2), 137-142
- Rasul, I. and Huck, S., (2010). Transactions Costs in Charitable Giving: Evidence from Two Field Experiments. *The B.E. Journal of Economic Analysis & Policy*, 10(1), 1935-1682
- Shapiro K.L., Caldwell J. and Robyn E. (1997). Personal names and the attentional blink: a visual "cocktail party" effect. *Journal of Experimental Psychology: Human Perception and Performance*, 23(2), 504-514
- Vriens, M., van der Scheer, H.R., Hoekstra, H.C. and Bult, J.R. (2012). Conjoint experiments for direct mail response optimization. *European Journal of Marketing*, 32(3), 323-339

Endnotes

1. For such a model, see Huck and Rasul (2010).
2. See Cabinet Office (2012).
3. As each product triggered a separate letter in the redress exercise, we limited our trial to consumers who only had one product and therefore would only receive one letter. This avoids potential confusion from customers who would receive two or more letters with potentially different features.
4. See Graves (2012) for a cogent critique.
5. See Kahneman (2012) for a summary of research on expert opinion. For a particular example of the failure of expert opinion on information disclosure for financial products, see Bateman et al (2011).
6. See Levitt and List (2009) on limits to what can be learned from laboratory studies.
7. See Fryer (2011) on the lack of effect of teacher incentive pay on student tests score.
8. See Bertrand, Karlan, Mullainathan, Shafir and Zinman (2010) who established that placing a photo of an attractive woman on a direct mail solicitation for a consumer loan in South Africa increased direct mail loan take-up as much as reducing the interest rate by 2% a month.
9. There are a number of drawbacks of RCTs including cost and whether the results will hold when a particular policy is implemented more widely, for example if this changes the market equilibrium. For the debate on the use of RCTs (and causal identification more generally) in economics, see Deaton (2010), Heckman and Urzua (2009), Imbens (2009) and Levitt and List (2009).
10. See DellaVigna (2009) and Chetty, Looney and Kroft (2009) for examples of the importance of salience in economic decisions. See Kahneman (2012) and Cabinet Office (2012) on the use of highlighting key messages by the use of simple language or bolding relevant text.
11. See Cabinet Office (2012), Cialdini (2007) and DellaVigna and Gentzkow (2010) on the role of the messenger in the impact of a message. See Bertrand, Karlan, Mullainathan, Shafir and Zinman (2010) for an example the importance of images.
12. See Karlan, McConnell, Mullainathan and Zinman (2011) on the role and importance of reminders.
13. Orbell, Dawes, and Schwartz-Shea (1994) show, using a prisoner's dilemma experiment, that men are less trusted in general than women.

- 14.** Just over a third are aware of the FSA according research conducted in February 2012 www.fsa.gov.uk/static/fsaweb/shared/documents/pubs/consumer-research/crpr86.pdf
- 15.** See FSA Handbook GEN 4.1.
- 16.** Although Orbell, Dawes, and Schwartz-Shea (1994) show that men are less trusted in general than women, more detailed studies have not found that women specifically trust men less, for example Buchan, Croson and Solnick (2008).
- 17.** Croson and Gneezy (2009) investigate a range of experiments and find that the level of trust displayed by women varies according to the particular experimental setting.
- 18.** See this link for details on the Experian FSS – www.experian.co.uk/business-strategies/financial-strategy-segments.html
- 19.** See this link for details on the IMD – www.gov.uk/government/publications/english-indices-of-deprivation-2010

Financial Conduct Authority

