Financial Conduct Authority

Occasional Paper No. 3

How does selling insurance as an add-on affect consumer decisions?

A practical application of behavioural experiments in financial regulation

March 2014

FCA OCCASIONAL PAPERS IN FINANCIAL REGULATION

Disclaimer

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.

Authors

Zanna Iscenko, Charlotte Duke, Steffen Huck, Brian Wallace.

Bibliographical note

Zanna Iscenko works in the Chief Economist's Department of the Financial Conduct Authority. Charlotte Duke works for London Economics. Steffen Huck and Brian Wallace are specialist consultants with London Economics and are based at University College London. Steffen Huck also holds a position at WZB Berlin Social Science Centre.

Acknowledgements

The authors are grateful to Peter Edmonds and the other members of the FCA's add-on insurance market study team for their support and input throughout this project, and also to Peter Andrews and Paul Adams at the Chief Economist's Department for their help and comments. We also thank Miriam Sinn and Mark Lewis from London Economics, and Gavin Ellison from YouGov for their contributions to designing and carrying out this research. From outside of our respective organisations, we are very grateful to Matthew Rabin, David Laibson and Vincent Crawford for generously sharing their thoughts and ideas related to this work, and especially to Marieke Huysentruyt for her valuable ongoing feedback on this research. All errors and omissions are our own.

Contents

Foreword

Behavioural economics tells us that consumers do not always make decisions in the ways that regulators have traditionally assumed. And consumers' reactions to well-intentioned regulations are not always what the regulator expected. These messages need to be taken into account to design regulatory interventions that effectively identify and address the root causes of poor consumer outcomes.

As demonstrated in its Occasional Papers 1 and 2, the FCA has been exploring practical ways of using behavioural economics to improve its regime. As part of this exploration, this paper discusses the pros and cons of using behavioural experiments in creating regulatory policy. It finds that these experiments have an important role to play, often providing insights into the way markets work that may not be possible to obtain otherwise.

The paper illustrates this by describing an innovative behavioural experiment developed for the FCA's first competition market study, which examined how consumer behaviour is affected when general insurance is sold as an add-on product to a primary purchase.

The paper also describes the results of this research, which show that the 'add-on mechanism' weakens demand-side discipline on firms. This in turn has implications for possible remedies.

Peter Andrews
Chief Economist

Executive Summary

The Occasional Paper on behavioural economics published by the Financial Conduct Authority (FCA) in April 2013 concluded that the effective regulation of retail financial markets requires a deeper and more realistic understanding of how consumers behave and what choices they find particularly difficult.[1]

Over the past year, the FCA has made considerable progress in embedding behavioural economics in its regulatory practice. For instance, more sophisticated analysis of consumer behaviour is being used to help identify problems in markets, design policy interventions, and analyse drivers of market outcomes. More practical results of this work will emerge over time - for example, in market studies conducted in pursuit of the FCA's objective to promote effective competition in the interests of consumers.

Practical behavioural analysis needs to be grounded in good empirical evidence on how consumers actually make decisions. This is particularly important for identifying the underlying drivers of consumer behaviour that in turn lead to poor market outcomes – for example, features of the market that might impede consumers' ability to drive effective competition. Regulators need this understanding to be able to target the root causes of problems, rather than tackle symptoms. This granular investigation of consumer behaviour often requires more complex behavioural research techniques, which have not traditionally been in the regulators' analytical toolkit.

Behavioural experiments as a regulatory tool

Behavioural experiments are one of the tools regulators can use to support practical behavioural analysis. This methodology makes it possible to study how consumer behaviour varies across different contexts, and to identify the causes of these variations with far higher level of confidence than simply observing consumers' choices in the market.

In contrast with field trials, which also explore the effects of particular changes in a controlled way but in real-world situations, experiments normally study decision-making in more stylised environments, which, nonetheless, capture the main aspects of the decision that is being investigated.

This level of abstraction has important practical advantages because it allows researchers to:

- identify causes and effects more precisely, by stripping out irrelevant features which might influence consumer behaviour; and

1 For more information see: Erta, Hunt, Iscenko and Brambley. (2013). Applying behavioural economics at the Financial Conduct Authority. *FCA Occasional Paper 1*.

- investigate whether or not there are general problems that affect consumers across multiple specific contexts—for example, markets for different insurance products.

Establishing general patterns is particularly valuable for regulatory analysis as interventions can then be designed to affect a number of markets.

A common perception is that because experiments take place in abstract and simplified settings, they cannot be used to make reliable predictions for the real world. This is not so. In fact, a well-designed experiment can be structured in a way that uses simplicity as an advantage. For example, if consumers struggle to make good comparisons in a very simple setting where the only difference between products is price, we can be confident that their ability to make good decisions will be no better in similar real-life settings that involve comparisons of a much more diverse range of products.

In general, there is good evidence to suggest that experimental results can reliably predict the presence and relative size of the effects of particular features of interest on consumer behaviour in real-world markets – for example, whether lower transparency of prices makes it harder for consumers to compare products effectively, and whether this effect is greater than for other potential barriers to shopping around.

However, the fact that experiments involve some simplifying departures from reality means that they are often not the best way to measure the absolute size of the effects of interventions on consumer behaviour in specific real-world situations. Like any other methodology, they might not be the right tool to use in some contexts. Section 3 of this paper explores the strengths and limitations of experiments (with the additional detail provided in Annex 1), and concludes that they, nonetheless, are applicable to a wide range of important regulatory questions. Experiments can be particularly useful for exploring topics such as barriers to consumers assessing information about products, and in regulatory analysis that supports diagnosing root causes of problems and the early stages of filtering out less effective remedies.

The FCA's experiment on add-on general insurance products

The FCA's first experience of using experimental methods to investigate consumer behaviour formed part of its market study on general insurance products commonly sold as add-ons (the GIAO study). This research demonstrates the value of these methods for regulatory analysis.[2]

One of the main objectives of the GIAO study was to investigate whether the inherent structure of transactions where insurance is offered as an add-on to another (primary) product has common effects on consumer behaviour across different insurance markets, and whether those effects impede effective competition.

The experiment the FCA designed with London Economics and academics from University College London (UCL) has been an integral part of the GIAO study. It made it possible to:

- directly test for common patterns of consumer behaviour that arise from the structure of the add-on transaction (the add-on mechanism) itself and hold across different products

2 *The full report on the provisional findings on the GIAO study* (FCA, 2014), as well as the detailed technical report on the experimental consumer research it involved (London Economics, 2014) are available at: http://www.fca.org.uk/news/general-insurance-add-ons-market-study

- identify which of the aspects of a typical add-on transaction are particularly important in driving these effects.

This research investigated the effects of a number of important current and potential features of an add-on transaction:

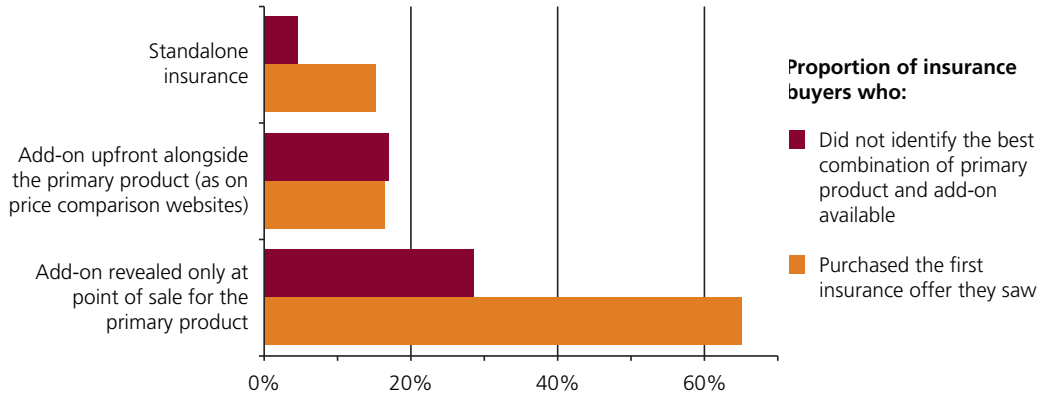
- complexity of comparing offers that involve individual prices for several products and no clear total price for the bundle as a whole
- reduced transparency due to the price of the add-on product being revealed only at the point of sale of the primary product (POS)
- a potential option to search for stand-alone offers from independent insurance providers alongside the add-on offer at the point of sale
- barriers to searching for stand-alone alternatives during an add-on transaction
- perceived lower prices due to annual contracts being presented in monthly terms.

The controlled experimental environment allowed a careful exploration of how these individual elements – and their typical combinations found in the real market – influence consumer behaviour and choices. Understanding these effects was very important for the GIAO study, but is also highly relevant for analysing how add-ons might affect consumers' ability to drive effective competition in other markets. The behaviours the experiment explored included:

- how consumers shop around for insurance and primary products
- how consumers decide whether to purchase a product and how much they are willing to pay
- how difficult it is for consumers to identify the best deal.

The experiment demonstrated that the structure of the transaction has a very big effect on consumers. For example, as some of the headline results in **Figure 1** show, being allocated to one of the three most common ways of selling insurance – stand-alone, transparently advertised alongside the primary product, or drip-fed during the purchase of the primary product – made a major difference to consumers' willingness to shop around and their ability to do this effectively.

Figure 1: The effects of the common insurance sales formats on consumer behaviour



The controlled experimental setting made it possible to break down these headline results into the effects of the individual elements of the add-on mechanism. As shown in Figure 2, this revealed that many of the individual elements of the add-on transaction can contribute to poor consumer outcomes, but they affect behaviour in different ways.

For example, on its own the complexity of comparing multiple prices (of the primary and the add-on products) at the same time without a clear total did not discourage consumers from shopping around for insurance, although it made them slightly less willing to look at primary product alternatives. However, added complexity – even without delaying the introduction of the add-on until the point of sale – made consumers much more likely to be unable to identify the best deals available to them. These difficulties are even more likely in the real world, where shopping around is more complex because products vary in quality as well as price, and multiple add-ons are often present at once. This suggests that consumers could need help in comparing options even when add-on prices are given up-front.

The experiment also highlighted that lower transparency due to add-ons being revealed only at the point of sale significantly exacerbates poor consumer outcomes. This element of the add-on mechanism was the most consistent in producing material detrimental effects on the outcome metrics studied in the experiment – including shopping around, prices paid and the quality of consumers' decisions.

Monthly pricing for annual contracts was another mechanism that had consistent, if smaller, detrimental effects across most types of outcomes. Surprisingly, however, even though presenting prices in monthly terms reduced the extent and effectiveness of shopping around, the fact that insurance 'looked cheaper' did not have an effect on consumers' willingness to buy it overall.

Finally, the research suggested that consumers are, in principle, willing to engage with stand-alone alternatives when those are easy to access at the point of sale, and benefit considerably from doing this. However, the beneficial effects start to diminish quickly when there are barriers to accessing those alternatives, even if these barriers are much smaller than the ones consumers face in real life.

Overall, the experimental research showed that the structure of the transaction has important and nuanced effects on consumer behaviour, many of which appear relevant for add-ons beyond insurance as well. Some of the elements with the worst effects on consumers also made them more willing to accept higher prices, suggesting that those ways of selling add-ons could be particularly attractive to opportunistic firms in markets that do not function well.

Figure 2: Selected effects of the elements of the add-on mechanism on behaviour and outcomes

	Complexity of multiple prices	Lower transparency from delaying add-on offer until POS	Access to stand-alone alternatives at POS	Small barrier to accessing stand-alone offers	Monthly pricing for an annual policy
Change compared to:	Insurance only	Add-on prices up-front with the primary product	Only add-on revealed at POS	Immediate access to alternatives at POS	Annual pricing
Shopping around					
Number of insurance offers seen	↑	↓	↑	↓	↓
Frequency of buying the first insurance product seen without looking further	=	↑	↓	↑	↑
Number of primary product offers seen		↑	=	=	↓
Purchasing behaviour					
Likelihood of buying insurance	↑	↓	↑	↓	=
Average price paid for insurance	↑	↑	↓	↑	↑
Proportion of very expensive insurance offers accepted (price over 4.5x expected payout)	↑	↑	↓	↑	↑
Quality of decisions					
Likelihood of not choosing the best deal available	↑	↑	↓	=	=
Average amount of money lost per consumer because of not identifying best deals	↑	↑	↓	↑	↑

Legend:

↑/↓ A material adverse effect on consumers (the element increases/decreases the metric in the direction that is likely to be associated with worse outcomes by 25% or more).

↑/↓ A small adverse effect on consumers (the change in the metric is towards worse outcomes for consumers, but by less than 25%.)

↑/↓ A material positive effect on consumers (the element increases/decreases the metric in the direction that is likely to be associated with better outcomes by 25% or more).

↑/↓ A small positive effect on consumers (leading to change in the metric of less than 25%).

= No statistically significant effect identified.

↑/↓ Indicate effects where consumer outcomes are uncertain. Larger arrows show stronger effects of 25% or more.

Empty cell: Metric is not applicable or comparable for the given element (e.g. number of primary offers seen relative to stand-alone insurance).

1. Introduction

Applying behavioural economics requires the use of a variety of empirical techniques to understand what drives consumer choices. A year ago the FCA published its first paper on field trials – a way of testing the effects of particular regulatory (or other) changes in a controlled way but in real-world situations – that demonstrated the value of this methodology for learning which remedies work in practice in financial regulation. The current paper focuses on behavioural experiments, another empirical methodology with a lot of potential value for regulatory analysis. This type of research explores how consumers make decisions in more stylised environments that capture the essential features of the consumer decision being investigated but leave out some of the detail. This in turn makes it easier to identify the important effects precisely.

Background on behavioural economics at the FCA

In April 2013, the FCA published its first Occasional Paper, 'Applying behavioural economics at the FCA' (OP1), which observed that that errors arising from consumers' behavioural biases are a major driver of outcomes in many retail financial markets. OP1 also set out initial thinking on how better understanding of consumer behaviours can contribute to all aspects of regulatory analysis: identifying problems in markets, diagnosing root causes of problems and designing effective and proportionate interventions. The paper also acknowledged that many practical challenges still remained around how to integrate behavioural economics into regulatory practice in a pragmatic and effective way.

Alongside the paper on the conceptual applications of behavioural economics, the FCA published its first applied behavioural research (OP2). OP2 described the methodology and findings of the FCA's field trial that used behavioural insights to improve consumers' responses to redress letters, demonstrating how more behaviourally-informed communications can significantly improve consumer outcomes. This research also showed the value of field trials as a methodology for testing the effectiveness of interventions when consumer behaviour is difficult to predict (as is almost always the case).

Although OP2 is the only example of the FCA's practical behavioural research published to date, over the past year the FCA has made considerable progress in integrating behavioural analysis in its day-to-day activities. Behavioural drivers of consumer detriment, for example, are a significant part of the FCA's approach to identifying and prioritising problems in markets it regulates. This approach involves systematically assessing each market to look for the typical drivers of poor market outcomes, including the main types of distortions arising from behavioural biases. Behavioural analysis has also been integrated into the standard framework for economic impact assessment for policy projects.

One other area where behavioural economics has made a significant contribution is the market studies the FCA carries out in pursuit of its objective to promote effective competition in the interests of consumers. The experiment discussed in this paper, for example, formed a part of the FCA's first competition study (the GIAO study), which concerned markets for general

insurance products sold as add-ons to another product or service (primary product)—for example, travel insurance offered alongside a holiday package.

Understanding how consumer behaviour might be affected by selling insurance as an add-on, rather than separately, played a key role in analysing how competition works in these markets. Insights from behavioural economics contributed throughout the study – for example, in developing hypotheses about ways in which the add-on mechanism can affect competition and gathering evidence effectively, such as designing surveys that were robust to the possibility that consumers might not pay much attention to the add-on purchase.

Some important subconscious effects of the add-on mechanism on consumer behaviour, however, could not be reliably tested using the evidence that is traditionally gathered in market studies. As a result, the FCA worked with London Economics and academics from University College London (UCL) to carry out an online behavioural experiment (the add-on insurance experiment), exploring some of the subtle ways in which the add-on sale might affect consumers' ability to drive effective competition and get good outcomes in markets.

Section 4 provides more detail about the GIAO study and discusses how the experiment was designed to support it.

The purpose of this paper

This occasional paper aims to share some of the methodological lessons learned, as well as specific findings, from the add-on insurance experiment. It forms a part of the FCA's wider programme of behavioural work, which is focused on developing, learning about and sharing pragmatic ways of integrating insights from behavioural economics into regulatory practice.

Behavioural experiments are an area where further learning and regulatory applications are particularly warranted since, despite their high potential value in exploring consumer behaviour, experiments are very rarely used in policy development. The GIAO study is among the first cases of financial regulators using experiments to diagnose common underlying drivers of problems that affect a number of markets. Experiments that test the effects of specific remedies—although more common—are still few and far between. This occasional paper reflects some of the FCA's ongoing work in applying behavioural experiments in the hope of stimulating regulatory debate about their use.

The paper is split into two parts:

- Part I introduces behavioural experiments as a tool regulators can use to explore drivers of consumer behaviour in financial markets, and assesses the strengths and limitations of this methodology in regulatory contexts.

- Part II describes the specific application of behavioural experiments as part of the GIAO study. It provides insights into how experimental design was developed to best support the objectives of the market study as a whole, and presents the main findings from the research that are relevant to the GIAO study and beyond.

The full technical report by London Economics on the add-on insurance experiment, containing a more detailed discussion of the methodology, the statistical analysis and the findings, is available as a separate online appendix to this paper.[3]

3 London Economics (2014), available at: http://www.fca.org.uk/static/documents/market-studies/gi-add-ons-experimental-consumer-research-report.pdf

Part I: Experiments in the regulatory context

2. Introduction to experiments

What is an experiment?

Economic experiments are a way of systematically studying how people's behaviour varies across different conditions. In essence, this methodology involves assigning participants to various different circumstances and observing how the choices they make change. The controlled nature of this process allows the researchers or policymakers to identify causal links between the circumstances and their associated differences in behaviour with far greater confidence than can normally be achieved by simply observing consumers' choices in real life.

The extent to which the setting of the experiment resembles the natural environment for the choices being studied varies greatly among experimental methodologies. Typically, these settings are stylised at least to some extent —giving participants tasks that mirror the essential features of the decision being studied, while removing some of the detail to increase simplicity and control. Stripping out factors that are not relevant reduces the risk that participants' behaviour will be swayed by features of the context (such as liking the brand of the product) in ways that will make it difficult to isolate the effects of genuine interest (such as complexity of different price formats). Participants in these experiments often perform tasks in specialised computer laboratories run by universities, or online (for example, through invitations to participate in research being sent to members of a survey panel).

However, some experimental methodologies—field trials—involve testing changes in real-life situations (for example, varying the disclosure forms given to clients by a financial adviser) and observing the consequences without consumers being aware that they are taking part in an research. For example, the FCA's field trial on improving firms' letters to encourage consumers to claim redress tested the effectiveness of changes to the letters as part of an actual redress exercise, but assigned consumers to receive different versions of the letters in a randomised and controlled way (a key feature of experiments).

Field trials are a very important tool for studying consumer behaviour, in regulatory contexts and otherwise. However, their real-world setting often presents very distinct challenges from most other experimental methods. For simplicity, in the rest of the paper the term 'experiment' is used to refer only to experimental methodologies that involve at least partly artificial, controlled settings, in contrast with field trials.[4]

4 More formally, using the typology of controlled experimental methods proposed in Harrison and List (2004), the term 'experiment' in this paper refers to laboratory experiments, artefactual field experiments and framed field experiments only. 'Field trials' refers to the remaining category: natural field experiments. Natural experiments (looking at how circumstances differ in the presence of the interest but are otherwise similar) are out of scope for this paper, as they do not allow intentional variation of conditions by researchers.

Table 1 below summarises the features that distinguish experiments and field trials from other methods of studying consumer behaviour.[5]

Table 1: Key features of experiments and field trials

Control	Experiments control the environment within which individuals make choices, in order to isolate the effects of interest and minimise the risk that changes in behaviour are driven by unrelated variations in unobservable factors that are not of interest. For example, if we are interested in exploring how information affects behaviour, an experiment can specify precisely what information the consumer sees for the decision of interest, how it is presented, and which attributes change and which stay constant between different experimental conditions being tested.
Treatments	An experiment typically involves holding constant most factors that could influence behaviour across different conditions, and only varying one feature at a time. Comparing consumer behaviours or outcomes in two conditions that only differ in one feature (for example, whether prices of a one-year subscription are provided in monthly or annual terms) isolates the effect of that feature. The incremental changes tested in the experiment are called 'treatments'. Although individual treatments are often small changes, this need not be the case: for example, treatments in an experiment can be two alternative designs for a disclosure form.
Observation	Experiments specify outcomes of interest in advance, and can be set in a way that allows the researchers to observe and measure how treatments affect those outcomes in a consistent and detailed way. Many experiments also monitor participants' behaviour leading up to the decision (for example, the information they look at) as well as observing the outcomes of their choices.
Randomisation	Participants in an experiment are generally allocated to specific treatments randomly to increase control. This minimises the common real-world problem that differences between outcomes in different conditions are driven by consumers of a certain type tending to choose some conditions over others. For example, more sophisticated consumers might recognise some price formats as misleading and avoid them, whereas less sophisticated consumers would not. If only less sophisticated consumers, who are more likely to make errors in any format, end up buying from dealers with very complex prices, just comparing consumer outcomes between for different price formats would overstate the inherent detrimental effects of complexity itself.

5 Criteria adapted from Lowenstein (1999) and the overview of experimental methods in Office of Fair Trading and Competition Commission (2009).

What are the main contributions of experiments?

Although not without their own limitations, experiments can be a valuable source of evidence about the underlying drivers of consumer behaviour that may be difficult or impossible to obtain through other means.

Some important advantages of experiments over other methodologies such as field trials or consumer surveys are:

- reliability in establishing causal effects
- the possibility of observing the effects on consumer behaviour, not just changes in outcomes
- the ease of investigating general underlying principles of behaviour.[6]

Establishing causal effects

The high degree of control over the environment, which allows the causal effects of specific changes to be isolated, is a defining feature of experimental methodologies. Identifying these causal links is very important for regulatory analysis, both in diagnosing the root causes of consumer behaviour in the market and in predicting the effects of regulatory interventions.

In the real world, although we may often see certain things happening together (for example, a complex pricing tariff being used by sellers and consumers failing to get the best deal), it is very difficult to establish that one causes the other. In principle, both could be driven by an unobserved third factor. In other words, just observing consumer choices in the market often does not allow us to identify causality, just correlation. Traditional consumer research that relies on consumer self-reporting might also be of limited use in exploring causality, because many behavioural biases have subconscious effects of which consumers may not even be aware.

Compared to self-reporting, field trials are more effective in exploring causal effects because, like the experiments discussed in this paper, they can also involve randomisation and control that directly link treatments being tested with the associated changes in behaviour. However, because they explore behaviour in more complex natural settings, field trials can have less control over the environment. As a result, there can be features which influence the observed outcomes but are not monitored and cannot be taken into account in assessing causality (for example, the likability of the adviser assigned to use a particular disclosure form, or differences in the pitches given to consumers).

Observing the effects on consumer behaviour

It is much easier to directly monitor how participants' behaviour changes across different treatments in experiments than with other methodologies. As a result, not only do experiments help establish causality between treatments and outcomes; they can also help uncover the mechanisms that drive changes in behaviour. In the FCA's add-on insurance experiment, for example, it was possible to explore whether consumers shopped around in different ways (such as whether they were less likely to look at add-on prices) depending on the structure of the add-on transaction, as well as simply measuring outcomes such as price paid for insurance. This information about the underlying mechanisms can be very valuable in designing remedies that target the root causes of consumer detriment, rather than just their symptoms.

6 A comprehensive comparative survey of different consumer research methodologies is out of the scope of this paper. OFT and CC (2009) provides a useful overview of the relative merits of different methods specifically for road-testing remedies.

In contrast, it may be very difficult to observe more than outcomes when using other methodologies. Because field trials take place in the real world, with consumers not aware they are participating in the research, the most researchers normally get to observe are the outcomes associated with each of the different treatments being tested – for example, how many consumers responded to a letter or rolled over their payday loan when different formats for information were used.[7] However, it is generally not possible to monitor exactly how consumer behaviour changed to produce different outcomes. Consumer surveys or focus groups can reveal more information about how consumers make choices, but consumers may not be able to reliably report what drove their actions or predict how their behaviour would have changed if circumstances were different.

Investigating general underlying principles of behaviour

Experiments make it possible to establish general principles that drive behaviour across specific contexts more easily and reliably than other consumer research methods. A well-designed experiment provides a stylised task that captures the key elements of a particular consumer decision, stripping the problem down to basic mechanisms that are common across decisions of this type and removing context-specific variations. Shopping around in the presence of complex multi-part pricing, for instance, at its core involves identifying different offers on the market, adding up multiple prices for the elements of the bundle offered by each seller and comparing the totals for different bundles. This is complex to explore in real-world settings. However, a stylised experimental task featuring two unlabelled products, or with a range of differently labelled pairs of products, can help explore the general principles that affect consumers' decision-making when they are confronted with these bundles across a range of real-life markets.

So far, experiments have been rarely used in regulatory contexts to test and diagnose the effects of general mechanisms (for example, transaction structures or pricing formats) on consumer behaviour, despite the significant potential to do so cheaply and efficiently. One notable exception has been the behavioural research by the Office of Fair Trading (OFT) on different price formats, which has since been successfully applied in interventions on airline charges and supermarket pricing - despite its fully abstract format. The FCA's research on the underlying effects of the different elements of an add-on transaction is, we believe, among the first applications of experiments in financial regulation that involves diagnosing general behavioural effects that may be distorting a wide range of markets.

7 For example, in discussing their field trial on usage of payday loans Bertrand and Morse (2011) specifically note the limitations of not being able to observe what actions the borrowers took to change their payday loan borrowing behaviour as a result of the different information treatments tested in the research.

Some recent uses of experiments by regulators

Office of Fair Trading – Price Formats (OFT, 2010a)
Topic: Effects of different price formats on consumers' ability to shop around effectively. This is a rare example of a regulatory use of experiments to diagnose problems (i.e. to identify pricing features that impede effective decision-making) rather than test specific remedies.

Format: An abstract lab-based experiment with a highly stylised choice environment (two alternative providers and unlabelled products).

Use of findings: The research contributed to the OFT's 'Advertising of prices' study, which identified particular risks associated with a number of specific promotional practices. The results of the study have been applied in challenging the lack of transparency of airlines' payment card surcharges and in agreeing principles for fair advertising of discounts with supermarkets (OFT, 2012a and 2012b).

European Commission – Retail Investments (Chater, Huck and Inderst, 2010)
Topic: Exploring the factors that affect consumers' investment decisions in advised and non-advised sales, and testing potential remedies (for example, information standardisation or disclosure of adviser's conflicts of interest).

Format: A combination of highly simplified investment choices online (advised and non-advised), and a laboratory experiment involving direct computer-based interaction between (non-specialist) participants randomly allocated to roles of consumers or financial advisers.

Use of findings: The European Commission explicitly referred to the research findings about the beneficial effects of simplified and standardised of information in its proposals on key information documents for investment products (EC, 2012). Other regulators, including the FCA, also drew on this research in their own behavioural work (Erta et al., 2013; Oxera, 2013).

Financial Services Authority – PPI Disclosure (de Meza, Irlenbusch and Reyniers, 2010)
Topic: Testing the effects of additional different types of disclosure on consumer decisions with respect to Payment Protection Insurance (PPI) sold as an add-on to credit products.

Format: A face-to-face simulated sales environment, using professional salespersons and consumers. Both sides of the transaction were very highly incentivised.

Use of findings: Although the experiment was originally focused on testing the effects of different types of disclosure on consumers' ability to assess PPI purchases, the most critical finding was that in a face-to-face setting the main factor that influenced consumers' decisions on whether or not to buy insurance was the skill of the salesperson they were paired up with. This result has been used as a cautionary notice when considering the viability of disclosure as the sole tool for addressing problems in face-to-face transactions for PPI and other work, including the GIAO study (FCA, 2014).

3. Reliability of experiments in regulatory contexts

Are experimental results relevant in real-life settings?

As discussed in the previous section, one of the greatest benefits of experiments is that they can identify the causal link between a treatment and changes in behaviour. Reliable causal identification of this kind is also referred to as high internal validity of the findings.

To be useful for answering questions with policy or academic relevance, it is also important for a study to have external validity. This means the effects established in the experimental setting still hold true in the real world. This generalizability of results is where the stylised and artificial nature of experimental environments is often perceived to be a downside. The external validity of experimental results is a topic of ongoing academic debate, but there are four important points that need to be borne in mind when considering the relevance of this debate to practical regulatory research.

1. What it means for experimental results to be externally valid depends on the purpose of the research, and is often different in an academic setting and in a policy setting.

Typically, the main focus of academic experimental economics is to establish the basic underlying principles of behaviour that underpin all market interactions (for example, how much people take wellbeing of others into account when making economic decisions), and generalise very widely. These experiments often involve university students as participants and feature a highly abstract environment with minimal labelling (for example, "product 1", "shop 1", prices in abstract currencies). Academic debate on external validity of experiments often focuses on whether findings from these very abstract laboratory experiments hold true across a wide range (or all) of real world markets.[8]

In contrast, policy research normally focuses on narrower questions: how consumers act in a particular market, and how specific changes in the context might change their behaviour. For experimental results to be externally valid in answering those questions, they need to be as close as possible to how consumers behave in the specific markets of interest. The possibility that those results might not hold elsewhere is of little consequence. Moreover, the narrower focus in policy research may make it possible to reflect more features of the particular market of interest (for example, use more product labelling or a more representative sample). This means that some academic criticisms of traditional laboratory experiments (such as high abstraction or the unrepresentativeness of the student population) would be far less relevant.

8 That is, most academic experiments fall under the definition of laboratory experiments in the framework proposed by Harrison and List (2004). This specific type of experimental research is the primary focus of most prominent critics of the methodology (such as Levitt and List, 2007a and 2007b), and of its defenders (for example, Camerer (2011) or Falk and Heckman(2009)).

2. There are ways to strengthen external validity through design

For experimental findings to be relevant in the real world, the design of the experiment needs to capture the features of the situation it investigates that are fundamental to the decision. For instance, the task of shopping around needs to reflect the basic attributes of the real-life decision: spending effort to look for alternative providers, comparing product prices and attributes, and experiencing gains from choosing better options.

However, it is impractical to exactly mirror the problem in the experiment. In fact, it is often not even desirable because additional detail would reduce control and make causal identification more difficult. When stylising the experimental environment while preserving external validity of the findings a useful guiding principle is to keep to simplifications or abstractions that, when corrected in the real world, would not reverse the direction of the treatment effects being observed in the experiment. That is, the departure from reality itself could not be driving the effects of the treatments.[9]

In contrast, experimental simplifications that could influence behaviour in the opposite direction from the treatments are less likely to be a problem for external validity. If treatments have an effect in the experiment, despite features of experimental design that should weaken this effect relative to real life, it is reasonable to expect that the identified treatment effects would be present beyond the experiment. For example, the experimental setting may be deliberately made simpler than the choices in the real world by making the products consumers need to compare differ only in price. If participants still make errors in identifying the cheapest option for them, it is reasonable to conclude that they will struggle even more when comparing products which vary in both price and quality.

When testing for the presence of a detrimental behavioural effect in consumer choice, having a known direction of bias against this effect relative to the real world ensures reasonable external validity if the experiment finds that a particular treatment makes consumer choices significantly worse. However, failing to find an effect will generally produce a lower external validity than in a more realistic design. This is because participants could still struggle when faced with a more complex real-life version of the products even if they did not have problems in a simplified experimental setting. Introducing more detail, however, can also threaten internal validity because it may become more difficult to establish that changes in consumer behaviour are caused by the treatments being tested rather than the specific details of design.

3. The external validity of qualitative findings of well-designed experiments is generally not in doubt

There is an important difference between the external validity of the presence of the effects detected in an experiment and their relative ranking, and the external validity of their precise size. Even critics of experimental methods generally do not dispute that the qualitative findings from stylised experiments hold in real-world settings. The main focus of the debate is around the generalizability of the sizes of those effects.[10]

In an earlier analysis of the use of experiments for policy carried out for Ofcom, Duke, Huck and Wallace (2009) also observe that even the more abstract laboratory experiments provide reliable predictions about the real-world differences between treatments—for example, identifying whether price format has an effect on consumer decisions, and determining how different price formats rank in terms of consumer outcomes.

9 Kessler and Vesterlund (2012) and Zizzo (2011) explore what deviations from the real world are permissible while preserving external validity of the findings in more detail. In general, departures from reality are most likely to be a problem if those omissions (confounds), by themselves, could have an effect on the outcomes of interest in the experiment, and that effect goes in the same direction as the hypothesised effects of treatments that are being tested. In this case it is possible that if confound were removed, treatments could have no effect on their own.

10 See Kessler and Vesterlund (2012) for a comprehensive survey of arguments around external validity of qualitative findings from experiments.

However, the validity of the precise quantitative effects observed in the experiment decreases as the experimental environment becomes more abstract. More realistic environments and larger samples of participants are generally needed to extrapolate the estimated magnitude of treatment effects found in the experiment into real-world settings, and even then caution will often be appropriate.

4. Predictions from experiments are generally aligned with findings from more 'realistic' research methods.

There are surprisingly few studies that explore the same hypotheses in experimental and in real-life (field) settings to directly test generalizability of behavioural experiments. However, a systematic review of such studies by Camerer (2011) shows that across a wide range of settings, including collectable card markets, fishing, donations and exam grading tasks, experimental results generalise well to the real world in all but one case (cooperation in fishing). Moreover, even after extending the comparison between laboratory experiments and field studies to research on similar, but not exactly matched, questions, Camerer still finds that predictions of the two methodologies are almost always comparable in the direction of the effects (and sometimes in magnitude as well).[11] Regrettably, no similar comparative analysis has been done to date on predictions of experiments and field trials in policy contexts specifically.

With these general points on real-life reliability of experimental findings in mind, this section concludes by identifying areas of regulatory analysis where experimental research is likely to make most useful and reliable contributions. These conclusions also draw on material in Annex 1, which contains a more detailed assessment of the factors that can improve the external validity of experimental findings, and of some limitations that arise from the nature of the experimental approach. The additional factors discussed in the annex include but are not limited to: designing incentives, choosing the participant sample effectively, and managing the challenges arising from participants' knowledge that their decisions are being monitored.

When are experiments most valuable?

As with any other methodology, the suitability of experiments needs to be assessed on a case by case basis. However, the advantages and limitations outlined in this paper point towards **five general principles** that can help decide when experiments are likely to be useful in answering regulatory questions:

1. The key elements of **the consumer decision can be represented relatively well** within the limitations of experimental design, with no obvious differences between the experimental setting and the real world that would to reverse the effects being tested.

In particular, it is possible to:

a. design outcome measures that can be observed in experimental settings and mirror effects of interest in the real world

b. use treatments that are sufficiently representative of the real-life effects being investigated.

11 In particular, in cases where lab-field methodologies were not perfectly matched Camerer (2011) observes good comparability between predictions in 20 studies and poor comparability in 2 of them.

2. **Policy objectives require an understanding of general mechanisms** that drive consumer behaviour across different products, **or of subconscious behavioural effects which consumers cannot reliably report.**

3. It is valuable to **observe directly how consumer behaviour changes** in response to treatments, as well as monitoring the resulting outcomes.

4. Insights into the **nature, direction and relative size of the effects are more important than measuring their absolute magnitude precisely.**

5. **Practical constraints and timelines allow scope** to commission, design and undertake experimental research. Experiments normally take between four and six months to design, implement and analyse. This makes them a relatively quick way of testing consumer behaviour robustly, but still potentially difficult to accommodate urgently.

What types of questions are experiments most suitable for?

The OFT's framework for thinking about consumer behaviour (OFT, 2010b) suggests that, to drive effective competition and protect their interests in the market, consumers need to be able to:

- access information about product offers
- assess the available alternatives in a reasonable way to identify the option that offers best value and best suits their needs
- act on this information by buying the best-value option they have identified.

All of these behaviours can be affected by behavioural biases, potentially leading to poor market outcomes for consumers. Experimental methods can support regulatory analysis of these three elements to a different extent.

Experiments can often be particularly useful in exploring consumers' ability to **assess** offers effectively and make the right decisions about product value. Since the experimenter controls the content and the format of the information that is presented to participants, it is possible to isolate the effects of both the information format and the inherent product complexity on consumers' decisions. With representative samples of participants it is also possible to explore which types of consumers are most at risk of making errors. Furthermore, the external validity of findings on barriers to assessing information is quite high. Experimental settings are simplified compared to real-life equivalents, so if participants have difficulties in an experiment, it is generally easy to predict that problems would also arise in the real world. It is harder to investigate the additional barriers to assessing information which arise from persuasion and trust in direct sales. Nevertheless, experiments that investigate face-to-face interactions are possible in some cases (as the example of PPI research discussed earlier suggests).

Using experiments to explore barriers to consumers **accessing** information – such as shopping around or reading terms and conditions – can be more challenging. If it is thought that a particular barrier to accessing information arises from the structure of a transaction, experimental evidence can be very useful because it can help observe how consumers shop around and what information they look at in different transaction formats. The add-on insurance experiment, for instance, was specifically designed to explore whether point-of-sale provision of add-on prices would stop shoppers accessing price information from other sources.

Experimental settings are generally simplified so, again, finding that even small barriers to accessing information about offers on the market have a large effect is likely to have reasonably high validity in the real world (provided that search effort costs and incentives are represented adequately).

There are, however, many other types of practical and behavioural barriers to consumers accessing information that do not lend themselves easily to experimental testing. These can arise for a number of reasons, such as lack of consumer engagement with financial products, not knowing where to look for information, or procrastinating in shopping around. Uncertainty about sources of information or lack of interest in accessing it can be difficult to reflect in experiments, in which participants voluntarily choose to participate and where clear instructions need to be given about the basic nature of the task and the available information to make the results tractable.

Finally, experiments appear least likely to be applicable in exploring barriers to **acting** on information. This is because important obstacles to acting on a decision (or failing to make a decision at all) generally cannot be captured in typical experimental tasks. Some of the most important barriers to acting are about consumer inertia: inattention to the purchase (or to the need to take action), over-confidence about the likelihood of taking action (for example, cancelling an auto-renewed contract in the future) and procrastination.[12] All these relate to consumers' expectations and actions over time, with a consumer having a choice on when to engage with the product (and generally failing to do so). As experimental settings generally directly focus participants on the task and take place over a short and fixed time period, it is very difficult to represent and explore these effects outside of the natural settings. Therefore, field trials generally appear to be preferable here. There is, however, a small number of recent experimental studies that successfully investigate how consumer inertia is affected by inattention (Sitzia et al, 2012) and over-confidence (Ericson, 2011; Letzler and Tasoff, 2013). These suggest that there may be some potential in using experiments for exploring barriers to acting with sufficiently creative experimental design.

What stage of regulatory analysis might experiments best support?

The FCA's OP1 concluded that behavioural economics can be valuable at all three stages of regulatory analysis: (i) identifying risks, (ii) diagnosing problems and (iii) developing remedies. This also applies to behavioural experiments.

Experiments are likely to be **particularly valuable at the problem diagnosis stage,** when it is necessary to understand the underlying drivers of consumers' behaviours or mistakes. Firstly, as highlighted in OP1, it can be very difficult to test which of the different hypotheses about a potential behavioural bias at play is true – yet this understanding is often essential for developing appropriate remedies. Behavioural experiments can often help discriminate between the different alternatives effectively. Secondly, regulatory interventions often affect a number of product types or markets at once, which makes it necessary to determine whether the underlying behavioural problems are common across those different settings. Behavioural experiments can explore the scope for an intervention to have positive effects across a range of individual products. Both of these factors made the add-on insurance experiment a crucial component of assessing problems in add-on insurance markets as part of the GIAO competition study.

12 Online Appendix D in Handel (2011) provides a good overview of the wide range of possible drivers of consumer inaction. The important role that over-confidence and procrastination play in consumer behaviour has been widely documented, for example in work by DellaVigna and Malmendier (2006), and Letzler and Tasoff (2013).

Perhaps less obviously, experimental research can also be genuinely value-adding for identifying *potential* problems. The application here does not come from doing experimental research when analysing a particular risk that has been detected; instead, by helping to understand what factors make it hard for consumers to access or assess information, experiments can **help identify risk triggers for future use.** For example, the OFT's work on the advertising of prices has found certain price formats, such as drip-pricing or time-limited offers, to be the most confusing (OFT, 2010a). Although not conclusive on their own, these experimental findings can help focus attention on higher-risk sales practices of specific firms to investigate further.

The area where perhaps most caution needs to be exercised in the use of experiments is developing and assessing remedies, even though regulators have used experiments for this purpose most often. Here it is important to draw a distinction between the early stage of developing and narrowing down a list of potential remedies, and the later stages of thoroughly assessing the impacts of the preferred options.

We know experiments can establish the presence of effects and their relative sizes but are much less effective in predicting the absolute size of the effects (for example, that a treatment leading to a 50% reduction in mistakes in an experiment would achieve the same outcome in the real world). Because of their simplified settings, experiments also have much more external validity in finding a problem rather than proving its absence – for example, a finding that a certain form of information provision eliminates errors may not entirely hold in the added complexity in the real world. Therefore it is suggested that experiments **may be most useful at earlier stages of remedy development to weed out options that clearly work less well,** or to identify unexpected outcomes relatively quickly. They may be less suitable for the later stages of remedy development. For example, an experiment on its own is not very likely to provide sufficient evidence to undertake a cost-benefit analysis of a remedy, because a reasonable indication of the magnitude of the effects is needed to weigh benefits against the costs.

Closer to implementation, once the preferred options have been identified, it is more important to establish whether an intervention will have the desired effect in the intended real-world setting, and to what extent. Testing remedies with real consumers and firms using field trials is likely to be preferable at that stage. However, it may sometimes be impossible to run a field trial, for example due to logistical constraints or ethical concerns that treatments might have potentially large adverse effects on some of the participating consumers. In those cases, experiments can offer a useful alternative way for exploring the effects of the proposed intervention, even at late stages.

Part II: Using experiments in practice: the add-on insurance competition study

4. Design and objectives of the experiment

The next two sections focus on the FCA's first behavioural experiment, which was jointly developed by the FCA, London Economics and academics from UCL as part of the FCA's GIAO competition study. The aim of this research was to explore how the specific features of the add-on transaction affect the way in which consumers shop around for and choose optional contracts of insurance.

The full technical report on this experimental research, containing a detailed discussion of methodology and a statistical analysis of the results, was published as part of the evidence for the market study and is available as an online appendix to this paper.[13]

The purpose of the current paper is to put this research on insurance add-ons into a broader regulatory context, by:

- explaining how the experiment was designed to increase its value as a part of a competition investigation (Section 4)
- highlighting the key findings, many of which were used in the GIAO study but are also of wider regulatory relevance (Section 5).

The role of behavioural research in the study

Past regulatory experience, both of the FCA and other regulators, has suggested that general insurance products that are sold as an add-on to other goods and services are often associated with poor consumer outcomes and, as in the case of payment protection insurance (PPI) or extended warranties, ineffective competition.[14] The aim of the FCA's GIAO study was to explore whether there were general problems across add-on markets that might lead to competition not working effectively and consumers purchasing poor value or unsuitable products. And, if there were common problems, whether they were driven by common features of the add-on markets that could be mitigated by general regulatory remedies.

What many of the past problematic add-on products had in common was the form of the transaction itself—how these products were introduced to the consumer during the purchase of a more salient primary product—rather than similarities of the nature of insurance cover

13 The full technical report by London Economics (2014) on the results from the experiment is available at: http://www.fca.org.uk/static/documents/market-studies/gi-add-ons-experimental-consumer-research-report.pdf

14 Examples of past competition investigations of add-on insurance products include Office of Fair Trading (2012) Extended Warranties on Domestic Electrical Goods: An OFT market study and notice of the OFT's intention to accept Undertakings in Lieu of a Market Investigation Reference; and Competition Commission (2009) Market investigation into payment protection insurance. Final Report. Additional examples of poor outcomes were identified in FCA's *thematic reviews on motor legal expenses insurance* (TR13/1) and *mobile phone insurance* (TR12/2).

or the types of firms in the market. There is also growing academic evidence on consumer difficulties associated with add-ons, and the overall importance of the form of the transaction for consumers' ability to exercise a competitive constraint on firms by shopping around and choosing products well.[15] For example, earlier research by the OFT showed that making the price format more complex and non-transparent changed the quality of consumer choices in the market for the worse. For this reason, understanding the effect of the add-on mechanism on consumer behaviour and how that in turn shapes market outcomes played a key role in the market study. This analysis was an integral part of diagnosing the scale and underlying causes of potential problems, and for pointing towards remedies.

Hypotheses about how the add-on mechanism can affect consumers' ability to drive effective competition – for example, by increasing complexity of choice or reducing attention paid to the insurance purchase because of the focus on the primary product – informed the evidence-gathering in the study, including data requests to firms and traditional consumer research (focus groups and surveys). However, sometimes discriminating between different ways in which the add-on mechanism could impede effective choices also required additional experimental evidence. This was for three reasons:

- **Generalizable findings:** For practical reasons most of the firm and consumer evidence gathered in the study was restricted to five add-on products: travel, home emergency, guaranteed asset protection (GAP), personal accident and gadget insurance. These products were chosen to cover a wide range of consumer groups and distribution channels, which helped explore the general nature of problems and the mechanisms that generated them. The experimental research made a significant contribution to this analysis by providing an opportunity to test for generalizable effects of the add-on mechanism itself directly, in a simplified setting that captured the essential structure shared by most add-on products.

- **Identifying specific effects:** Add-ons could be associated with several barriers to consumers engaging with the market effectively—for example, more complex multi-part pricing, reduced transparency of add-on prices, and reluctance to consider alternatives once the add-on is offered at the point of sale. In the real world, these factors often occur together, making it difficult to determine which aspects of an add-on transaction genuinely cause difficulties for consumers, and which are unproblematic. Furthermore, in real life, consumers in add-on and stand-alone markets could differ in their ability or preferences, making comparisons of outcomes between the two groups harder. As discussed in Part I, the high degree of control afforded by experimental environments helps to isolate these effects individually.

- **Overcoming self-reporting bias:** Given the hypothesis that consumers may not pay much attention to add-on products due to focusing on the primary purchase, their recall of their shopping around and decisions about add-ons could be limited. Moreover, consumers would not consciously recognise and report many important behavioural influences (such as how much price format may have affected their decision). All traditional consumer research was designed to be as robust to this constraint, but in the experimental environment we could also directly observe the effects on consumer behaviour without relying on self-reporting at all.

15 For a summary of the academic literature on the add-on mechanism see Huysentruyt and Read (2010) or Baker and Siegelman (2012). OFT research on price frames, and of drip-pricing in particular, is an example of the wider empirical evidence that is indicative of the potential anti-competitive effects of non-transparent optional extras (OFT, 2010a and 2013)

How did the add-on insurance experiment work?

The detailed explanation of the experimental design that was used in add-on insurance experiment is out of the scope of this paper. The rest of this section outlines the key elements, and how they helped align the research with the objectives of the GIAO study as a whole.

The experimental environment

The experiment was set up as a simplified online experience of shopping around for and 'purchasing' a primary product and an associated optional insurance product. The shopping around task was done five times with five different primary products, described in more detail below.

Each primary product had an associated risk of an adverse event occurring, against which the participants had an option to insure—for example, the risk of a newly-purchased boiler breaking down and requiring an emergency repairman call-out. By purchasing insurance the consumer was fully insured against the cost of the adverse event. To avoid choices being driven by unobservable personal beliefs about the different risks, participants were informed up-front of the probability of the adverse effect occurring, and the amount of money they would lose if it happened. The loss probabilities and sizes, as well as primary product prices, were set at realistic levels using firms' data submitted in the GIAO study or external sources. The prices of insurance varied between 0.9 times and five times the actuarially fair price, to explore consumer behaviour with products that offer a wide range of value for money (as it also varies considerably among real add-on products).

In all treatments, prices for add-on insurance and for stand-alone insurance were drawn from the same distribution, meaning that neither type of insurance was inherently cheaper or more expensive than the other. This was done to distinguish the effects of the mechanism of insurance being provided through different channels (point-of-sale, as an add-on more generally, or by independent stand-alone providers) on shopping around and the pure financial reasons that could drive consumers to focus on particular sales channels more (for example, to ignore stand-alone alternatives because the likelihood of getting a good deal there is lower.)

Representing shopping around

The participants started with seeing an offer from one primary provider (with the associated insurance), and could choose to search for more alternatives. Even in a simplified experimental environment it was important to reflect the inconvenience of looking for alternatives, which is an essential feature of shopping around in real life. Therefore, searching for more offers involved a time delay before the information about the next option was revealed. The delay started at 5 seconds, but increased as the participant undertook more searches, representing the increasing effort of finding deals beyond headline firms.

Once the search cost had been incurred for revealing a new offer, however, going back to previously revealed offers was instantaneous. Previously revealed offers did not remain on the screen, but could quickly be revealed by clicking on a link which said "Go to shop X". This is broadly similar to switching to an already open tab in an internet browser.

For each product, consumers could search for offers from up to eleven different primary product providers. In treatments where there was an option to view additional stand-alone insurance offers, there were also eleven stand-alone insurance providers. There was no difference in the quality or risk of the primary product or the insurance offered by the different providers within each task. Providers only differed in the prices offered for the primary product and insurance, which were randomly generated to provide enough variation for there to be an incentive to shop around.

Because there was no variation in product quality, respondents faced a simple task of minimising the cost of the product bundle (primary and, optionally, insurance) they wanted to purchase. The simplified setting means that the quality of decisions observed in the experiment is likely to be higher than in real life. This was a deliberate design choice to increase the confidence with which any errors found in the experiment could be extrapolated into the real world, where additional complexity would make mistakes even more common. This is in line with the approach to enhancing the external validity of experimental findings described in Section 3.

Incentives

Incentives for participants were also designed to reflect the main elements of shopping around for products and buying insurance. They also involved real-money payoffs which were linked to performance in the tasks, to motivate the participants to reduce the prices they paid and make good choices. Due to obvious financial constraints, the actual payoffs were much smaller than what would be involved, for example, in buying and potentially losing a real tablet computer. However, to improve the plausibility of the experimental tasks and increase participants' engagement, the tasks used realistic product prices and loss sizes in GBP terms, which were converted into smaller pay-outs at the end of experiment (at a conversion rate participants were told about in advance).

Participants were told what the value of the primary product was to them, and that they could 'trade-in' the product at the end of the experiment for this value. This gave them a reason to purchase the cheapest primary product (their earnings at the end of the experiment were trade-in value minus price paid for the product). The cost of insurance, if participants chose to buy it, was also subtracted from those earnings, creating incentives to minimise it. The payments were strongly linked to performance, with actual 'earnings' for participants in the experiment ranging from 8 pence to almost £5.

For the analysis to be generalised to real-life markets, it was also important for incentives to reflect exposure to risk associated with buying insurance. If the participant had decided not to purchase insurance, and the adverse event occurred, the cost of the adverse event was deducted from the participant's payment when they traded in the primary product. Incentives to purchase insurance in the experiment were therefore aligned with the incentives of purchasing insurance in the real world: by paying a smaller fee up-front, the participant could avoid a higher loss which occurred with a given probability.

Products

As previously outlined, participants in the experiment did the shopping task in the context of five different primary products. These products were: a home boiler, a tablet computer, a laptop computer, a luxury holiday for two and 12 days' car hire. Each product had a corresponding adverse event which exposed the participant to a loss, against which they could insure. For example, the adverse event for the home boiler was a breakdown requiring an emergency repairman call-out, which, participants were told, happened to '4 out of 100 buyers'.

The products differed not only in their description, but also in their parameters. For some products, insurance prices were very small relative to the price of the primary product, and for others this ratio was much larger. For some products the cost in case of the adverse event was very high and for others this cost was lower. Similarly, the likelihood of the adverse event occurring varied across products.

Differences in behaviour for different products could therefore be due to two separate effects:

- the parameters of each product (for example, relative product prices or likelihood of risk)

- personal biases participants might have towards different products such as their perception that some products require insurance more than others

To disentangle these two effects, the experiment also included unlabelled versions of the products (for example, 'product 1') – so some respondents did not see the labels 'home boiler', 'tablet computer' etc when they performed the tasks. The objective characteristics of the products, however, were unchanged. Exploring the effect of different formats of the add-on transaction with different products (in objective and subjective terms), and in the abstract framework, was important to establish whether any observed problems were genuinely driven by a common add-on mechanism rather than product-specific characteristics.

Participants

The research was carried out with 1,514 UK-based members of the YouGov survey panel, who represented a broad mix of ages, education levels, household incomes and levels of financial and problem-solving ability.[16]

While it is common and often more efficient to carry out experimental research with students, the objectives of the GIAO study required a demographically mixed sample for two reasons. First, it was important for the study to understand whether there are certain types of consumers that are disproportionally affected by the add-on mechanism. If many vulnerable consumers struggled to make the right choices, this could still raise concerns about how well the market is working for consumers, even if problems are not universal. Secondly, it was not apparent in advance whether students' behaviour in relation to insurance decisions would be representative of (or at least better than) the population of consumers in general insurance markets, who typically have more practical experience with these products.

What effects did the experiment test?

The purpose of the experiment was to explore how the structure of the add-on transaction affected a number of types of consumer behaviour which help drive effective competition. The relevant outcome variables that were tracked in the experimental environment for each treatment broadly fell into three categories:

- **Shopping around:** as measured, for example, by the number of alternatives viewed or the likelihood of making a purchase without shopping around at all.
- **Purchasing behaviour:** including take-up rates, price sensitivity and average prices paid.
- **Quality of choices:** likelihood and size of errors in choosing the best deal available[17] as well as indicators of the underlying drivers of those errors, such as limited attention (for example, focusing only on the primary price in choosing a supplier) or confusion (for example, surprise at total cost).

The discussion of the findings from the experiment in Section 5 provides more detail on the measures used in each of the three categories.

16 Participants were recruited to be representative of the UK population in terms of age, gender, social grade, education and geography. Where there were differences, the analysis of outcomes controlled for demographic characteristics.

17 Although the experiment monitored the effects of insurance take-up rates, those were not judged to be indicative of errors on their own because of the potential rational explanation based on high risk aversion. Instead, errors were defined as objective failures to minimise costs, taking insurance preferences as given.

Building up the different factors that can contribute to a typical add-on transaction incrementally (for example, separating complexity from the time when the add-on price was revealed), and observing how shopping around and choices changed in response, made it possible to isolate the effects of the individual factors on the outcome variables above in a way that would not have been possible in the real world where all factors normally occur together. The experiment investigated these incremental effects of different elements of the add-on mechanism by randomly allocating each participant to one of the following environments:

i. **'Insurance only':** The simplest scenario where the only task was to shop around for stand-alone insurance.

ii. **'Up-front add-on':** Shopping around both for the primary product and the optional insurance. Each provider of a primary product also offered an associated add-on, and prices of both products were transparently displayed up front; this is similar to how prices of some add-ons are listed in search results on price comparison websites for home or motor insurance.

iii. **'Add-on revealed at the point of sale (POS)':** The treatment was identical to (ii) but for the fact that the existence and price of the add-on offered by each primary provider were not shown transparently, but instead was revealed only after the consumer selected the primary product and added it to their shopping basket (but before they confirmed the purchase, allowing scope for changing one's mind). This was a simplification of a common practice of drip-feeding add-ons during the primary purchase.

iv. **'Add-on revealed at POS & easy SA alternatives':** As in (iii) but with an option to search for and view alternative stand-alone (SA) insurance offers directly at the point of sale after the add-on offer was revealed by the primary provider.

v. **'Add-on revealed at POS & harder SA alternatives':** As in (iv) but with a minimal barrier to accessing stand-alone insurance offers at the point of sale. Add-on and stand-alone offers could no longer be viewed simultaneously, instead requiring one click to switch between them (similar to switching between two tabs in a web browser).

These treatments are interesting in their own right as they represent different current or potential ways in which consumers can encounter optional add-ons. More importantly, the treatments were designed so that pairwise comparisons between them show the incremental effects of the different elements of the add-on mechanism.

- **Complexity of comparing bundles with multiple prices without a clear total:** Comparing consumer behaviours in (i) and (ii) shows the effects of having to compare different firms' offers that include two prices, primary product and insurance, even if both are fully transparent, relative to a straightforward insurance purchase.

- **Reduced transparency from add-on price being revealed only at the POS:** The only difference between (ii) and (iii) is the timing of the insurance price being shown to the consumer, so any changes in behaviour can be attributed to transparency.

- **The overall add-on mechanism for POS insurance offers:** Comparing POS add-on (iii) and a simple insurance purchase (i) identifies the overall effect of the traditional practice of offering insurance as an add-on during the primary purchase, compared to buying it separately.

- **Improvements (if any) from immediate access to stand-alone alternatives alongside a POS add-on offer:** Comparing (iv) and (iii) shows the effects of the simplest possible way of prominently introducing stand-alone alternatives at POS. Although not itself a part of the add-on mechanism, this treatment helps identify whether inertia at POS deters willingness to engage with alternatives.

- **Minimal realistic barriers to accessing stand-alone alternatives at POS:** Having assessed the best-case scenario in terms of availability of alternatives at POS in (iv) as a benchmark, it is useful to explore how quickly the beneficial effects of access to stand-alone offers disappear in the face of more realistic constraints to shopping around immediately. Comparing (iv) and (v) shows the detrimental effects (if any) of even small impediments to accessing stand-alone insurance at POS.

In addition to the five treatments that changed the format of the transaction, the experiment also explored the effects of how information was presented. The experimental environments which each participant could be allocated to differed in these two ways for all five experimental tasks:

vi. **Yearly vs monthly insurance price:** All insurance contracts in the experiment were annual. However, the participants were randomly allocated to either seeing insurance offers quoted in monthly terms or being given a full price for a year. In both cases, once the participant selected an insurance offer and placed it in their shopping basket, the final total annual price was provided and they could still easily remove the insurance policy at this point. Comparing behaviours of consumers in these two price scenarios allowed exploration of whether the provision of smaller monthly prices changes incentives for shopping around or purchasing insurance and whether the additional calculations required for monthly offers impede effective comparisons across alternatives.

vii. **Product labelling:** As discussed above, participants either carried out the tasks with five different 'realistic' primary and insurance products (car hire, boiler, etc.) or were allocated to an environment where the five products had the same objective characteristics (for example, price, exposure to loss event) but were labelled in abstract terms ("product 1", "insurance 1", etc.). Comparisons among product types and between their labelled and unlabelled formats suggested how general or context-specific any effects of the add-on mechanism were. They also investigated whether participants' choices in the experiment were sensitive to the types of products – for instance, whether the likelihood of buying insurance affected by the price of insurance relative to the cost of the primary product, regardless of the actuarial value.

Table 2 summarises the seven types of treatments: their mechanics in the experiment, any real-life practices they reflected, and which effects of elements of the add-on transaction were tested by pairwise comparisons between different treatments. The next section outlines the key findings about the effects of these treatments above on outcome metrics.

Table 2: Summary of the main treatments

Treatment	Description in the experiment	What it represents in real life	Questions explored
i. Insurance only	Participants searched for and had an option of buying a separate insurance product for a primary product they had already owned.	Structure of a pure stand-alone insurance purchase.	This was the simplest and most transparent treatment, so it **served as a benchmark** for insurance search quality.
ii. Up-front add-on	Participants searched for and bought a primary product (e.g. a home boiler or a tablet computer) and also had the option to purchase add-on insurance. Each provider offered a primary product and an optional add-on, and the price of the add-on was displayed next to the price of the primary product.	Transparent add-on pricing (e.g. add-on prices given in price comparison website search results for home or motor insurance). And extent to which choice of the bundle is driven by the higher (more salient) primary price rather than total cost.	Comparing outcomes (ii) with (i) showed effect of additional **complexity of comparing bundles that involve multiple prices.**
iii. Add-on revealed at POS	Identical to the 'up-front add-on' scenario except that the add-on insurance offer (i.e. the existence of the offer and its price) was only revealed once the respondent had clicked on the 'Buy' button to add the primary product to the shopping basket. Participants could return and view insurance offers for primary products they had already searched and check their add-on prices without any additional time delay or cost.	Standard point-of-sale add-on presentation in many online purchases of primary products. Add-on existence and price are only revealed once the purchase of the primary product from a particular provider starts.	**Comparing (iii) with (ii) showed the effect of the reduced price transparency** often associated with POS selling, and the extent of POS advantage. The difference between (iii) and (i) is overall effect of a **typical add-on vs stand-alone.**
iv. Add-on revealed at POS & easy SA alternatives	Identical to the 'Add-on revealed at POS' treatment with the exception that Participants could also search for alternative stand-alone insurance. The stand-alone insurance offers were displayed next to the primary and POS add-on, so the consumer could easily compare the POS add-on and the alternative stand-alone insurance offers.	Adding immediate information about SA insurance alternatives at POS with ability to buy insurance independently without additional effort.	Observing behaviour in (iv) explored the extent to which inertia at POS might prevent engagement with alternatives. Comparing outcomes of (iii) and (iv) explored welfare effects in the best possible scenario of requiring provision of alternative insurance offers at POS.
v. Add-on revealed at POS & hard SA alternatives	Identical to the 'Add-on at the POS & easy SA alternatives' treatment, with the exception that when searching for an alternative stand-alone insurance, the participant could no longer see the offers for the primary product with the add-on insurance. This made it more difficult to compare the POS add-on and the alternative stand-alone offers. However, barriers were minimal – it required one click to switch between the add-on and the stand-alone, similar to the effort required to switch between tabs in a web browser.	An option to search for alternatives at the point of sale but with minimal barriers (e.g. having to click on a link and switch to a new tab to access a price comparison website of stand-alone providers).	(v) is a more realistic representation of POS provision of alternative insurance offers and participants' willingness to engage with them. Comparing (iv) and (v) explored the **effect of even minimal barriers to accessing SA** information on the way in which it was used.

Treatment	Description in the experiment	What it represents in real life	Questions explored
vi. Monthly vs annual pricing	Participants either saw their (annual) insurance offers quoted in monthly terms or as a full price for a year. After they selected an insurance offer and placed it in their shopping basket, the total annual price was provided in both cases, and participants could easily remove the insurance policy from their 'basket' and continue shopping around.	Different pricing approaches observed for add-on products (and beyond): giving the price of a long-term contract in smaller monthly terms (with the total cost not available or less salient), compared to showing the full price.	Comparing monthly and annual scenarios tests **whether monthly pricing itself changes consumer behaviour** – for example, reducing willingness to shop around or increasing the likelihood of buying insurance (because it looks cheaper).
vii. Product labels	The tasks involved five products, which differed in their characteristics (e.g. price) and descriptions. For some participants the same products were labelled in abstract terms.	The natural variation in insurance products that can be sold as add-ons.	The extent to which any effects identified in the experiment are driven by a c**ommon add-on mechanism** rather than being product-specific.

5. Key findings from the add-on insurance experiment

This section outlines and explains the main findings of the experiment with minimal technical language and detail, leaving the detailed statistical analyses for the full research report.[18] However, all results reported here are statistically significant unless stated otherwise.

Overview of the results

Due to the large number and variety of factors tested in the experiment, this section starts with a high-level qualitative overview of the key findings and themes that have emerged from the research, summarised in **Figure 3** below. Overall, the research found strong evidence that the structure of the add-on transaction materially affects a range of important consumer behaviours. These include willingness to shop around, deciding whether to buy insurance at a particular price, and identifying best deals available. Through this, the add-on mechanism can significantly affect consumer outcomes. Some of the key effects of individual elements, and their combinations, are outlined below.

The complexity of comparing offers that involve multiple prices (of the primary product and insurance) without a total cost, rather than buying insurance separately, did not have a material effect on shopping around for insurance. However, dealing with multiple pieces of information rather than considering them one at a time (primary first, then the add-on at POS) reduced consumers' willingness to shop around for the primary product. Greater complexity of comparing offers from different providers, even when all prices were transparent, also led to significant deterioration in quality of consumers' decisions – they were much less likely to identify the best available deals, lost more money as a result, and found calculating total costs challenging.

The reduced transparency of revealing the insurance only at POS, after the primary product provider was initially chosen, acted as a powerful barrier to shopping around, even when comparing treatments which both involved bundles with multiple separate prices to isolate the effects of changing transparency alone. It led more than half the participants to buy the first insurance offer they saw without any further search or comparison, even though the costs of shopping around were much smaller than in real life. This element of the add-on mechanism also made consumers much less sensitive to the price of insurance when buying it, leading to many very expensive insurance offers being accepted. Less transparent insurance prices also significantly increased the likelihood of consumers not being able to identify the cheapest deal available to them, and to lose more money on average as a result.

Given the detrimental effects of both complexity and delays in revealing information about the add-on, it is perhaps not surprising that the scenario which combines both – **offering insurance**

18 London Economics (2014), available at: http://www.fca.org.uk/static/documents/market-studies/gi-add-ons-experimental-consumer-research-report.pdf

as an add-on at point of sale—resulted in significantly worse consumer outcomes in insurance purchases on most metrics, compared to buying insurance separately.

Although revealing the add-on insurance offer only at the point of sale clearly led to fewer consumers going back to check other add-on alternatives, **providing an option to search for stand-alone alternatives immediately at the point of sale significantly improved consumer outcomes on most metrics.** Consumers generally used the option to shop around for insurance much more, paid lower prices (in part because they had access to more offers) and were more willing to reject very expensive offers. Interestingly, adding stand-alone alternatives only caused a small increase in the likelihood of consumers being able to identify the best available option, suggesting that the complexity of the transaction overall and low transparency of add-on alternatives continued to play a role. However, losses from ineffective shopping around were smaller in this scenario.

Even **minimal barriers to accessing stand-alone alternatives at the point of sale,** however, reduced the positive effects described above quite quickly. With stand-alone alternatives slightly less prominent and convenient, consumers: became likely to do less shopping around (or none at all); became slightly less sensitive to insurance prices; paid more for insurance; and lost more money on average due to not identifying best available offers.

These outcomes were still much better than offering add-ons at the point of sale without any alternatives. However, the significant impact on consumer behaviour of even minimal barriers to accessing alternatives raises questions about the challenges of introducing stand-alone alternatives more prominently at the point of sale in the real world, where barriers to taking action are likely to be much higher. The significant effects of small barriers also help to demonstrate that the low likelihood of consumer engagement with stand-alone alternatives (which are available from other sources in principle) when faced with a point-of-sale offer is part of the general add-on mechanism.

Finally, **presenting prices of annual contracts in monthly terms rather than as the full cost for the year** also had some adverse effects on consumer behaviour. The fact that monthly prices 'looked lower' did not change willingness to buy insurance, but it significantly reduced the likelihood of consumers shopping around and the extent to which they did so. This also increased the prices consumers paid on average and the amount of money they lost as a result of being less able to identify best available offers correctly. It was clear from the participants' behaviour in the experiment that they found the monthly format more confusing: they were far more likely to change their minds about buying a particular insurance product when shown the total cost just before confirming their purchase, than when faced with full yearly prices from the start.

The findings described above held true across different product labels, suggesting that they reveal the genuine underlying effects of the elements of the add-on mechanism rather than being driven by product-specific features.

The rest of the section discusses the outcome measures and results in more detail. It first presents the effects of the five treatments that represent different structures for insurance transaction on the three broad categories of outcomes – shopping around, purchasing behaviour, and quality of decisions observed in the experiment. Comparisons between these treatments, which allow the impacts of different elements of the add-on mechanism to be isolated, are also discussed. The last sub-section focuses on other notable impacts, such as the effects of presentation of information (price formats or product labels) or of the socio-demographic characteristics of consumers who participated in the experiment.

Figure 3: Key effects of the elements of the add-on transaction on behaviour and outcomes

	Complexity of multiple prices	Lower transparency from delaying add-on offer until POS	Overall effect of add-on at POS vs. stand-alone	Access to stand-alone alternatives at POS	Small barrier to accessing stand-alone offers	Monthly pricing for an annual policy
Change compared to:	Insurance only	Add-on prices up-front with the primary product	Insurance only	Only add-on revealed at POS	Immediate access to alternatives at POS	Annual pricing
Shopping around						
Number of insurance offers seen	↑	↓	↓	↑	↓	↓
Frequency of buying the first insurance product seen without looking further	=	↑	↑	↓	↑	↑
Number of primary product offers seen		↑		=	=	↓
Purchasing behaviour						
Likelihood of buying insurance	↑	↓	=	↑	↓	=
Likelihood of buying an add-on		↓		↓	↑	=
Average price paid for insurance	↑	↑	↑	↓	↑	↑
Proportion of very expensive insurance offers accepted (price over 4.5x expected payout)	↑	↑	↑	↓	↑	↑
Quality of decisions						
Likelihood of not choosing the best deal available	↑	↑	↑	↓	=	=
Average amount of money per consumer because of not identifying best deals	↑	↑	↑	↓	↑	↑
Consumers changing their mind after seeing total cost of selected products (indicates confusion)	↑	↑	↑	↓	=	↑

Legend:

↑/↓ A material adverse effect on consumers (the element increases/decreases the metric in the direction that is likely to be associated with worse outcomes by 25% or more).

↑/↓ A small adverse effect on consumers (the change in the metric is towards worse outcomes for consumers, but by less than 25%.)

↑/↓ A material positive effect on consumers (the element increases/decreases the metric in the direction that is likely to be associated with better outcomes by 25% or more).

↑/↓ A small positive effect on consumers (leading to change in the metric of less than 25%).

= No statistically significant effect identified.

↑/↓ Indicate effects where consumer outcomes are uncertain. Larger arrows show stronger effects of 25% or more.

Empty cell: Metric is not applicable or comparable for the given element (e.g. number of primary offers seen relative to stand-alone insurance).

Impacts of the transaction structure on shopping around

Participants in the experiment could 'search', or shop around, for additional offers in the following ways:

- Search for additional insurance offers, either by revealing add-on insurance offers associated with the primary products (in the treatments that involved add-ons at point of sale), or by searching for alternative stand-alone insurance offers, if available.
- Search for additional primary products – for example, additional offers for tablet computers.

This section analyses how the search behaviour for both primary products and insurance offers was affected by the various treatments.

Search for insurance

The increased complexity of buying insurance as part of a multi-price package (alongside the primary product) rather than separately did not by itself significantly change how much participants shopped around for insurance. How much participants searched for insurance products was similar between the two tasks where they searched for the insurance product only and the 'Up-front add-on' situation. In the latter case, however, the search for insurance was passive, as participants would see insurance offers by default as they shopped around for the primary product.

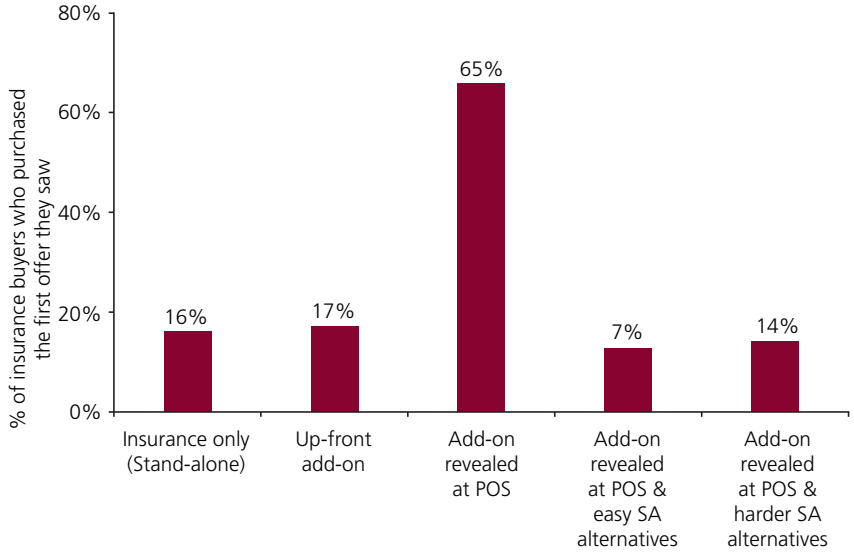
Reducing the transparency of the add-on product, however, significantly deterred search. When the insurance offer was only revealed after selecting the primary product but before confirming the purchase (the 'Add-on revealed at POS' treatment), participants chose to view significantly fewer insurance offers than when add-on prices were given up-front (almost three fewer offers viewed). Over 70% of participants only viewed the insurance offer associated with the primary product they purchased. This compares to fewer than 20% who did so when the insurance offer was shown up-front.

The reduced transparency associated with only revealing the add-on price after the consumer selected a primary product dramatically increased the number of people who bought insurance without shopping around at all. As can be seen in **Figure 4** below, in the 'Add-on revealed at POS' scenario 65% of insurance buyers purchased the first insurance offer they viewed without searching any further. This contrasts starkly with only 17% of participants who did so when the insurance offers were presented side by side with the primary product.

The timing of when the add-on was introduced in the transaction was the only difference between 'Up-front add-on' and 'Add-on revealed at POS' treatments, suggesting that the marked reduction in shopping around is genuinely driven by the lack of transparency about the add-on existence and price when it was only revealed at POS. It is also very difficult to attribute the change in behaviour purely to rational search costs, such as shopping around becoming more difficult or time-consuming. Participants still needed to confirm their purchase after they selected the primary product offer and had the add-on revealed to them, so they could easily go back to any other primary product seller they had already viewed and reveal their add-on prices in a matter of two clicks without any extra cost or delay. This suggests that the reduced transparency associated with delaying add-on offers until the point of sale triggers behavioural inertia. In the real world this inertia can amplify rational reluctance to shop around when faced with a POS add-on offer due to the inconvenience of going elsewhere at that point in the primary product purchase.

In principle, behavioural inertia could also reduce willingness to consider alternative insurance offers even if those are made available at the point of sale. Comparing search behaviour in the treatments 'Add-on revealed at POS' and 'Add-on revealed at POS & easy SA alternatives' tests whether this is the case.

Figure 4: Proportion of insurance buyers who bought the first insurance offer they saw without searching further



This comparison shows that **participants were willing to engage with alternatives when those were provided directly at POS, and shopped around more for insurance as a result.** Those who could search for stand-alone insurance offers directly at the point of sale made use of this option despite the additional search costs (time delays) involved in revealing independent offers. On average, they viewed 3.2 more offers than those who did not have access to alternatives at POS. This amounts to more than doubling their search effort. They were also significantly less likely to simply purchase the first insurance offer they came across when the option to search for stand-alone alternatives was available alongside the add-on insurance, as can be seen by comparing the third and fourth bar in Figure 4 above.

To measure the openness to, and effects of, considering alternative insurance offers at POS in principle, the 'Add-on at the POS & easy SA alternatives' treatment introduces alternative offers at POS in the simplest and most prominent way possible. Comparing this scenario with 'Add-on at the POS & harder SA alternatives' investigates how important the ease of accessing alternative insurance offers is for search.

Making it only slightly more difficult to search for alternative insurance—similar to requiring consumers to switch to another tab in a web browser—led participants to view roughly 10% fewer offers (5.00 as opposed to 5.48) and made them significantly more likely to purchase the first offer they came across. This finding suggests that despite the fact that POS access to alternative insurance offers appeared to be very beneficial to consumers, even very small differences in how easy it is to access this alternative insurance can have large effects.

Search for the primary product

In a direct reversal of the findings for insurance search, the only material effect for the primary product was that **reduced transparency of the add-on product (delaying it until the point of sale) led participants to shop around for the primary product slightly more** than in the case where both primary and add-on prices were presented side-by-side. Compared to the 'Up-front add-on' scenario, participants in the treatments in which the add-on was introduced only at point of sale viewed almost one primary offer more on average, increasing their search effort by 19%. A likely explanation for this behaviour is that willingness to shop around when both prices were revealed up-front was affected by the unavoidable complexity of being confronted with multiple prices at once without a clear total, as this treatment was the only case where participants saw both primary and add-on prices immediately. Delaying add-ons until POS, in contrast, allowed consumers to think about different elements of the price one by one.

There were no material differences in search for the primary product among the three different treatments where the add-on was revealed at the point of sale. Perhaps unsurprisingly, **the existence of stand-alone insurance alternatives at POS, or barriers to accessing them, did not affect how much participants shopped around for the primary product.**

These findings are important for two reasons. Firstly, they bring out the some divergences between choices that consumers may perceive as more difficult and those where they struggle most to make good decisions. The perceived complexity of considering both add-on and primary product prices transparently side-by-side was not associated with more errors than in cases where the add-on was revealed only at the point of sale. In fact, the opposite was true. Secondly, the findings highlight that reduced transparency can be viewed by consumers as beneficial in complex situations, as it allows them to shop around for the primary product more easily. This effect could be even more pronounced in the real world where there are multiple add-ons to keep track of. Viewing them all alongside the primary product without clear total prices could be sufficiently confusing for consumers to significantly impede shopping around for the primary product.

Impacts of the transaction structure on purchasing behaviour

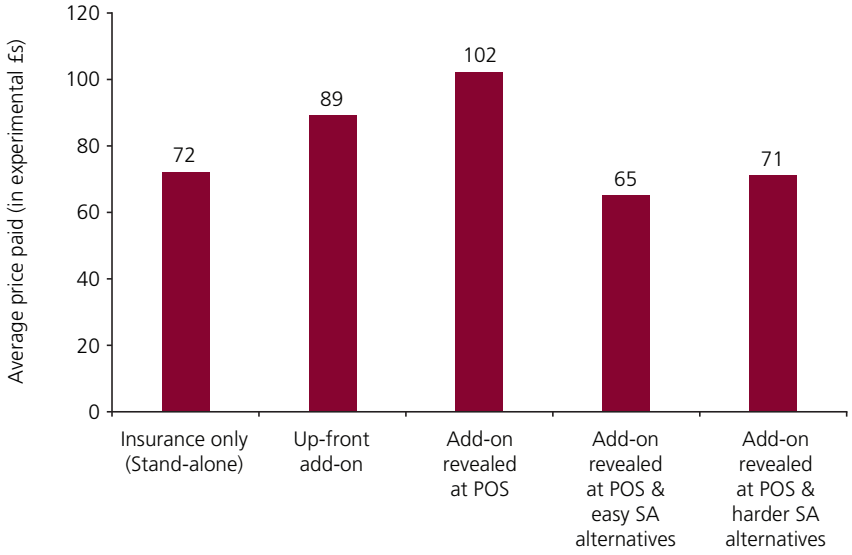
The discretionary nature of the add-on insurance purchase (and any other optional extras) could make similar search behaviour have very different effects on consumer decisions and outcomes in the cases of add-ons compared to multi-part pricing of single products, which has been investigated in previous research by OFT. In principle, even if certain transaction formats like revealing the price only at POS deter shopping around, consumers are still free to avoid high-price add-ons by not buying them after they see the offer, in contrast with drip-pricing that involves compulsory charges. It is therefore important to investigate how the effects of different add-on structures on shopping around translate into prices paid and insurance take-up rates.

Average price paid

Simple comparisons of the price participants paid for insurance are not very meaningful between some of the treatments, due to the different constraints on the consumer choice. For example, the maximum number of insurance products participants could see in treatments that involved stand-alone alternatives was double that of simple add-on treatments, meaning that participants had a better chance of finding a good deal in those scenarios. Furthermore, in treatments without stand-alone alternatives, participants were constrained by having to consider the insurance and primary product costs together, meaning that sometimes accepting a higher price on the insurance offer could have been necessary to get the cheap primary product that would minimise costs overall.

Nonetheless, two notable findings in comparable settings show that the way in which insurance was presented had a material effect on how much participants paid for it.

Figure 5: The average price paid for insurance



First, as can be seen in **Figure 5** above, **delaying the insurance add-on offer to the POS resulted in consumers paying £13, or 15%, more for their add-on insurance compared to a more transparent format where insurance was presented up-front.**[19] For comparison, the 'actuarially fair' premium (i.e. the size of the insurable loss times the likelihood of the loss happening) for insurance was about £41, meaning the reduced transparency of add-on offers resulted in participants paying an extra one third of the actuarial cost of insurance on average, compared to those in who were in an environment where add-on prices were given up-front.

The second meaningful comparison can be made between the two treatments that involved both add-on and stand-alone alternatives. Doing this shows that **even small barriers to accessing stand-alone alternative offers at the point of sale increased prices paid for insurance overall, particularly for add-on products.**

While **Figure 5** suggests that the average price paid is only moderately increased by £6, or about 10%, when stand-alone alternatives became harder to access, this disguises a significant difference between the prices paid for stand-alone and add-on insurance in the two treatments. The price for the alternative stand-alone insurance was almost identical with an average of £62 and £63 paid per product. However, the average price for add-on insurance increased more significantly as a result of small barriers to alternative offers, rising to £86 from the £74 paid when access to alternatives at the point of sale was immediate.

This result also suggests that when the search for alternative stand-alone insurance is more difficult, the price of the add-on insurance revealed at the point of sale needs to be higher to induce participants to search for stand-alone alternatives.

19 Here, and throughout the discussion of the experimental results, pound values reported refer to the experimental units (framed as £s) in which prices and costs were presented to the participants, rather than the actual remuneration participants received at the end of the experiment.

Purely rational motivation—effort of needing to make a single additional click to reveal a stand-alone alternative—is unlikely to be material enough to explain this outcome. However, once participants do search for alternatives, they do so with equal effectiveness in both treatments.

Take-up of insurance

Because of the optional nature of add-ons, any effects on prices paid need to be considered alongside the changes in insurance take-up rates.[20]

The lower transparency of the point-of-sale add-on prices and the associated higher add-on prices led to only a small change in willingness to buy insurance. In particular, the proportion of consumers who chose to purchase insurance in the 'Add-on revealed at POS' scenario was 5 percentage points lower than when add-ons were transparently shown up-front (69% compared to 74%). This effect is likely to reflect responses to average prices rather than the inherent effect of the mechanism on willingness to buy insurance, because the difference between the two treatments disappears after controlling for the number of insurance offers viewed.

Although the change in take-up between the two treatments appears to suggest some response to price in the demand for insurance, the experiment has also found that in general **the reduced transparency of the 'Add-on revealed at POS' format made consumers much less sensitive to price when deciding whether to buy insurance from a particular provider.** For example, when faced with an add-on offer at the point of sale, participants accepted 23% of the offers to purchase significantly overpriced insurance (those with prices over 4.5 times the actuarially fair value). Acceptance rates of such offers in other treatments were just 5% for the 'Up-front add-on' (a directly comparable setting) and even lower, between 1% and 3%, for treatments with stand-alone alternatives.

In the other case where insurance prices increased between two comparable treatments – **when small barriers were introduced to searching for stand-alone alternatives – there was also only a small decrease in insurance demand.** Insurance take-up rate fell from 85% for 'Add-on at POS & easy alternatives' to 81% 'Add-on at POS & harder alternatives'. Again, this difference is not statistically significant after controlling for the number of offers seen. A more interesting finding for this pair of treatments, however, is the presence of additional behavioural effects from even small barriers to accessing alternatives, which are hard to reconcile with rational responses to prices. When a one-click barrier to searching for stand-alone alternatives was introduced, participants became more likely to buy insurance as an add-on (30% of participants buying add-ons, up from 25% in the treatment with easier access to alternatives) and considerably less likely to buy it as a stand-alone (51%, down from 60%). The shift of demand towards add-ons happened even though, as described above, the add-on price was significantly higher in the treatment with harder alternatives.

In general, however, participants were more likely to purchase insurance when they had access to stand-alone offers at the point of sale as well as the add-on. This effect remained even after controlling for the number of offers they actually viewed, which suggests that higher take-up rates aren't just driven by the fact that with stand-alone alternatives participants had access to more offers and were therefore more likely to find a good deal. A possible explanation could be that seeing a prominent option to search for insurance at the point of sale focused their attention on the insurance purchase.

20 It is the difference in behaviour between treatments, however, that is of interest in exploring the effects of the add-on transaction structure. The absolute value of the take-up rates in the experiment would not necessarily correspond to the real world because of the simplified nature of cover, provision of explicit loss event probabilities and weaker, if any, emotional attachment to the product being insured. However, these factors are much less relevant for external validity of comparative findings.

Overall, insurance take-up rates showed little or no response in treatments where consumers paid higher prices. This suggests that reducing add-on transparency and introducing barriers to searching for alternatives were profitable changes from the point of view of insurance sellers (assuming the costs stayed constant). There was no scope for firms responding to consumer behaviours in the experiment, but in a dynamic environment one would expect more firms adopting those opaque sales tactics to maximise profits.

Impact of the transaction structure on quality of choices

Definitions of errors and loss measures

The findings above suggest that some elements of the structure of the add-on transaction may lead consumers to shop around less, but there is another important question: could the add-on mechanism lead people to shop around less well?

Given the nature of the study, it is not possible to directly measure and compare welfare across different treatments because consumer preferences are unobservable. More risk-averse people may make a mistake when they do not purchase insurance and risk-neutral consumers may make a mistake if they buy insurance at a high price. Similarly participants who value time a lot make a mistake when they search too much and others may make a mistake when not shopping around enough. As discussed in OP1, separating genuine preferences from mistakes is a major challenge for behavioural analysis.

One way to identify poor choices in this context is to look for clear inconsistencies with previous decisions, or with uncontroversial principles of choice such as preferring more money to less, other things being equal. The experiment uses this principle by focusing on an error measure which only requires the choices made by each participant to be consistent while accepting their revealed preferences on how many offers to search for (i.e. tolerance for delays) and whether to buy insurance (i.e. their risk aversion).

The objective for participants in each treatment was to minimise costs and maximise earnings. The error measure is therefore defined as the proportion of participants that did not choose the cheapest combination of the products they chose to buy among the offers from alternative providers they had searched for (and already incurred associated the cost of the time delay). It is important to note that this error measure is based on all insurance offers available to the consumer, including those add-on insurance offers that participants did not choose to reveal without extra cost when they had viewed the respective primary product offer.[21]

For example, consider a participant who viewed one tablet for £455, and a second one for £450. After revealing the second tablet, the participant decided to reveal the add-on insurance offer associated with it, which was for £90, and purchased this combination (tablet plus insurance). If the add-on insurance offer of the first tablet had been less than £85, this participant would count as having made an error regardless of whether she revealed the add-on offer associated with the first tablet (since £455+£85=£450+£90).

The loss measure to estimate the consequences of worse choices is calculated as the actual amount lost in earnings (in experimental pounds) in the experiment if the participant made an error compared to choosing the best deal available. For example, if the add-on insurance offer

21 As previously stated, search and viewing insurance offers was easy in this experiment, with only one click of the mouse required to reveal an add-on offer associated with a primary product. Therefore extreme assumptions about the value of time to the participant are needed to justify not revealing add-ons, in rational terms.

on the first tablet had been £65, the loss to the participant from forgoing that option would be £20 as they paid £20 more than they could have done if they went for the offer from the first provider. The loss was zero for participants who did not make an error.

Likelihood and size of observed errors

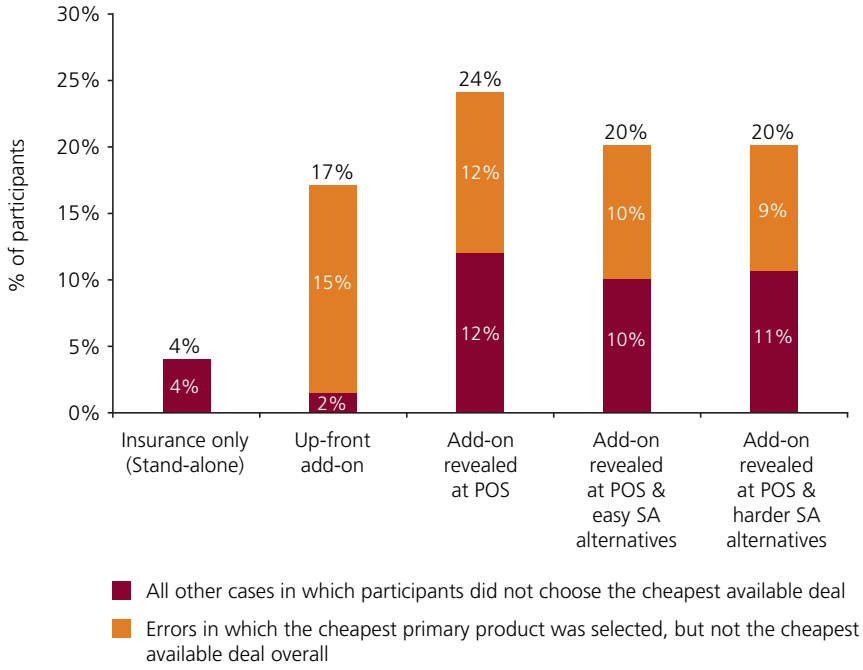
As expected, the proportion of consumers making errors was minimal for insurance-only search, which had been deliberately designed to be the simplest possible case of shopping around to serve as a benchmark for quality of choice. Only 4% of participants did not choose the best possible deal in this setting, most likely because of forgetting about earlier offers or not paying attention to the task.

However, as can be seen from **Figure 6** below, **introducing the additional complexity of comparing prices of both primary and add-on prices at the same time increased the proportion of participants who made a errors by 13 percentage points – more than fourfold** – compared to the single-price benchmark. The breakdown of the average losses per participant in **Figure 7**, between those who just bought the primary product and those who bought insurance as well, also reveals insights into the scale of the effect. Firstly, it demonstrates that not only did participants struggle to identify the cheapest deals in the 'Up-front add-on' far more than when shopping around for insurance alone, but that the associated losses were larger as well. Secondly, losses were more than £7 larger for insurance buyers than non-buyers, suggesting keeping track of the additional price dimension was a challenge for participants. This suggests that the increased complexity of dealing with packages of products with multiple prices made it much more difficult to compare offers from different sellers effectively.

The ability of participants to shop around effectively even in this simple setting was further worsened by the reduced transparency associated with add-ons being revealed only at the point of sale. Almost a quarter of participants did not identify the cheapest deal available to them in this treatment, with the error rate increasing by 7 percentage points compared to the scenario where add-ons were shown up-front. The error rate increased even more among insurance buyers, to 29% from 17% in the scenario where add-on prices were advertised transparently. The significant further increase in errors in the 'Add-on revealed at POS' treatment was largely driven by the fact that participants chose not to reveal the majority of add-on insurance offers which were available to them. Having incurred the search cost for the primary product, the associated add-on insurance offer could be revealed immediately and without cost; yet the majority of participants chose not to do so and, as a result, made losses.

The size of average losses by consumers also increased significantly as a result of reduced add-on transparency in the 'Add-on revealed at POS' treatment. The effect is particularly pronounced among insurance buyers, for whom the average loss almost doubled compared to the scenario where add-on prices were transparent, increasing from £13 to £22. Average losses also increased among those who only bought the primary product, suggesting that the lack of transparency around some of the elements of the bundle made it harder for consumers to shop around effectively in general. There is an interesting contrast between the suggestion that consumers seemed to perceive shopping around for a primary product to be easier when add-ons were delayed until POS compared to transparent up-front prices, and the finding that neither the likelihood nor size of errors in primary product purchases improved as a result of this change.

Figure 6: Frequencies and types of errors by participants

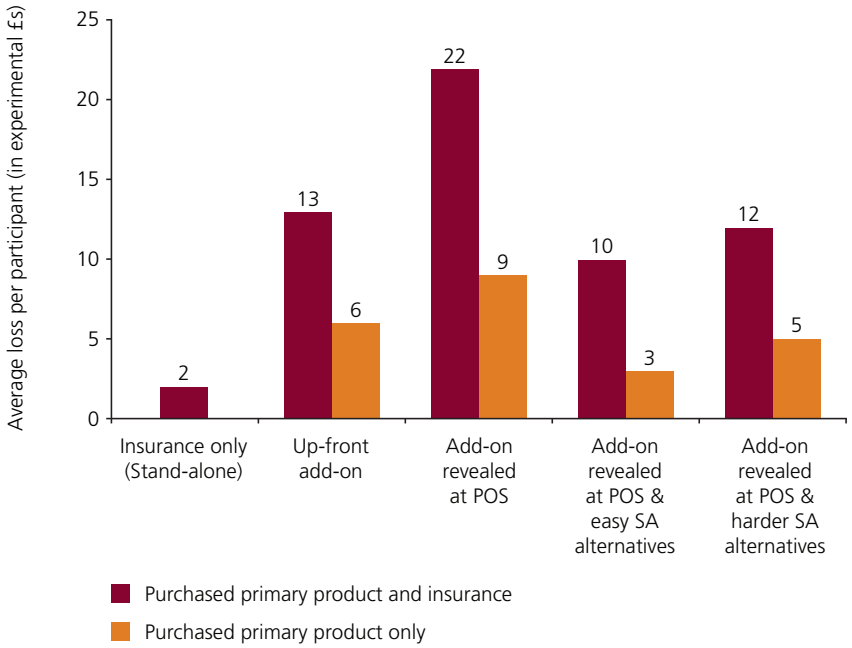


Note. The percentages shown above each bar are the proportion of participants that did not identify the cheapest available deal overall in each of the treatments. The figures shown in the dark red portion of the bars are proportion of participants that made did not choose the best deal overall but instead accepted the offer that had the cheapest primary product.

In the **'Add-on revealed at the POS', half of all errors were associated with the consumer purchasing the cheapest primary product instead of the cheapest primary and insurance combination.** This offers some insights into the underlying drivers of errors: consumers in this scenario seem to have focused their search efforts more on the primary product and devoted less of their attention to the add-on insurance product. Notably, focusing too much on the primary product price in choosing the provider appears to be associated specifically with the reduced transparency of the add-on not being revealed until POS. In contrast, in the 'up-front add-on' scenario, only just over 10% of the mistakes are associated with choosing the cheapest primary product. This suggests that, at least in the experimental setting with all prices presented with equal prominence, the primary product was not necessarily the most salient driver of choice among providers.

There are some suggestions that provision of alternatives at the point of sale is beneficial, although the improvements are not as material as for many other outcomes discussed above. The likelihood of consumers not choosing the best deal decreased slightly from 24% in the 'Add-on revealed at the POS' treatment to 20% in the 'Add-on revealed at the POS & easy SA alternatives' treatment. The size of the losses was affected more materially, however, almost halving from the 'Add-on revealed at POS' levels.

Figure 7: Average amount of money lost due to not purchasing the cheapest option seen per consumer



There is no material difference in the quality of decisions between the two treatments in which stand-alone alternatives were provided at the point of sale. Specifically, the barriers to accessing stand-alone alternatives do not make consumers any less likely to identify the best deals out of those available.

Impacts of information presentation and demographics

Monthly versus yearly framing of prices

As discussed previously, the experiment also tested whether or not the format of the insurance price itself affected consumer behaviour. Some participants saw the prices for insurance products (both add-on and stand-alone) framed as a lump sum cost for a year, or they saw the cost of insurance broken down into smaller monthly components. In both cases the participants were clearly told they were buying an annual policy.

In the monthly price framing, the total cost was displayed as soon as the participant placed the insurance product in their shopping basket. They could still easily remove the insurance policy from their basket at this point in the experiment. This intervention therefore was only a minor modification that is likely to understate the effects of computational complexity involved in converting prices. Nonetheless, monthly framing of annual offers produced the following effects on consumer behaviour:

Consumers shopped around less under monthly framing, viewing significantly fewer insurance offers. The average number of insurance offers viewed under the yearly lump sum framing was 4.68, but in the monthly framing this number was reduced to 3.94, a decrease of 16% in the number of offers viewed. Monthly price also made participants more likely to buy the first insurance offer they viewed. 19% of participants did so under the annual price

framing, compared to 26% of buyers who saw monthly prices, suggesting that the smaller value reported for monthly insurance encouraged more people to search less.

Similarly, the **average price at which insurance was bought was higher under monthly framing.** Insurance policies priced as yearly were bought for £76 on average while monthly priced policies were bought for £81 on average. This was entirely driven by a higher cost paid for add-on insurance, as there were no differences in stand-alone insurance prices between the two price frames. This result is in line with the observation above that the main effect of low prices is to discourage starting to shop around when faced with an insurance offer. The research did not find a statistically significant effect of the price frame on the take-up of insurance.

Some of the findings in the experiment also suggest that monthly pricing reduced participants' ability to shop around effectively. Average loss to consumers was higher under the monthly framing than it was under the yearly framing. **The average loss participants made under the monthly framing was £12 per product, while it was only £8 in the yearly framing. This represents a 50% increase in the average loss** made per purchase, indicating that consumers were finding it more difficult to make the correct decisions. The likelihood of making a mistake, however, increased only slightly, from 16% to 17%, between annual and monthly price formats.

Another measure of decision-making quality monitored in the experiment was the tendency of individuals to remove the insurance product from their shopping basket after seeing the total insurance cost. The experimental results clearly show that **those who saw insurance prices quoted in monthly terms were much more likely to be surprised and change their mind about purchasing a policy when they saw the total price given on an annual basis.** Across all treatments, the likelihood of changing one's mind in response to the total cost was 2-3 times higher for those in the monthly price frames than in yearly format, reaching over 30% for 'Up-front add-ons' and 'Add-on at POS'.

This result confirms that the adverse effects of the monthly framing identified above are likely to be an underestimation of the effects of this practice in the real world where prices are often not immediately converted to the equivalent yearly cost, giving consumers fewer opportunities to realize their misunderstanding of the price. It also demonstrates that a significant proportion of participants in the monthly framing made choices they considered to be sub-optimal when given better (annual) information about costs, suggesting that monthly framing led them to initially under-estimate the costs of insurance.

The effects of product labelling

The results of the experiment show that labelling products generally increased the take-up of insurance. This is likely due to the fact that participants found it easier to relate to purchasing a 'tablet computer' than simply 'product 1'. Similarly, participants were also somewhat more likely to purchase the first offer seen when the product was unlabelled. Both of these findings could be indicative of a lesser engagement with the task in an abstract setting. There was no systematic difference between labelled and unlabelled products on other important dimensions.

The analysis also revealed that take-up of insurance was higher for products for which the cost of the primary product was much larger than the cost of the insurance. This finding suggests that adding insurance with a significantly smaller cost than the primary product itself was less painful, possibly because consumers' assessment of what was and was not expensive was 'anchored' on the high price they were already prepared to pay for the primary product.

Similarly, as the cost of the primary product rose relative to the cost of the insurance policy, participants were less likely to search and more likely to simply purchase the first insurance offer seen. Once again, this could suggest that participants anchored on the expensive primary product and then did not put much search effort into the insurance product. Unlike the relative price framing effects on take-up, however, investing less search effort could be rational when insurance add-on is very cheap relative to the primary product. For example, in some cases consumers could observe that insurance price would need to fall dramatically (for example, more than halve) to compensate them for the difference between the primary products they chose and the next-best primary offer. They could then conclude that the value of looking at other insurance offers is not very high because gains from a better insurance price were unlikely to outweigh the cost savings on the primary product they were currently getting.

Although some of the patterns varied across products overall as described above, **the effects of the structural treatments tested in the experiment were still present and had the same ranking regardless of the product label.** This suggests that the differences in behaviour in response to treatments were driven by a common underlying add-on mechanism.

The effects of socio-demographic characteristics

Some traditional demographic variables, such as effects of age and education, generally had no or only minor and unsystematic effects on the observed outcomes, suggesting that the effects of the add-on mechanism discussed above hold across a variety of consumer groups. There were some more interesting findings, however, on the effects of other consumer characteristics.

Participants who could correctly answer a financial literacy question also paid far less for insurance (nearly £8). They were likely to view more offers, although at the same time were slightly less likely to purchase insurance at all. However, performance on the financial literacy question was not a predictor of the size of the loss due to errors in the experiment.

The post-experimental survey measured the number of insurance products currently owned by each participant, including both main insurance products and a variety of add-ons. As this measure could be effectively a proxy for risk aversion, perhaps not surprisingly, those who owned more insurance were more likely to buy it in the experiment. Furthermore, participants who owned more insurance products tended to pay less for insurance in the experiment and tended to view more insurance offers. It is possible that this is due to the greater amount of experience these participants had with the insurance market; however, it could also be a consequence of enjoying the task of buying insurance more (for example, due to higher risk aversion) and so investing more effort. Perhaps surprisingly, a more traditional measure of risk aversion that was elicited through an additional gamble after the experiment did not have statistically significant effects on the outcomes measured in the experiment.

The most important predictor of experimental performance was cognitive ability, as tested by a short, simple question that is commonly used to measure ability to suppress intuitive responses and give a deliberative answer to a problem. The participants were incentivised for answering this question correctly. Those who scored higher on this cognitive ability measure did significantly better in most areas of interest in the experiment. They paid less for insurance (nearly £5 less), viewed more insurance offers and had lower losses. Even more importantly, participants who did not do well on this question were particularly likely to have higher losses in treatments that included non-transparent add-on presentation at POS (including where alternatives were available).

6. Conclusions

Behavioural experiments played a key role in the GIAO study. This methodology allowed the FCA to gather valuable evidence about the effects of the different elements of a typical add-on transaction on consumer behaviour, and explore how general those effects were.

The stylised environment of the FCA's add-on insurance experiment captured the essential features and incentives of shopping around for an add-on product alongside a primary purchase. It also generally made shopping around and comparing different providers simpler for consumers than it often is in real life. This approach to experimental design allowed much greater confidence that barriers to making effective choices observed in the experiment wouldl be relevant for choices in real-world markets. In fact, the predictions of the experiment were strongly in line with survey findings about consumer behaviours in specific markets investigated in the GIAO study.[22]

The experimental findings uncovered material ways in which different elements of buying insurance as an add-on can contribute to poor consumer outcomes. They also shed light on important interactions between these elements. For example, isolating the effects of complexity due to multiple prices from the reduced transparency of the add-on offer revealed only at the point of sale showed that simply presenting prices of add-ons transparently up-front without giving clear total costs may not address all the difficulties consumers face in shopping around. In fact, it could also make it more difficult to compare primary products. This and other insights offered by the experiment are very useful in designing appropriate behaviourally-informed remedies.

Many of the findings could also have wider applications. Even though the experiment involved insurance contracts, the setting in which consumer behaviour was explored is representative of many other cases where optional extras are offered during the purchase of a primary product. Combined with the OFT's work on the detrimental effects of partitioned pricing, the FCA's research on add-ons demonstrates that the structure of the transaction has very strong, and generalizable, effects on consumers' ability to drive effective competition. It also highlights the importance of considering how and when information is presented to consumers in designing regulatory interventions, as well as thinking about its content.

There seems to be great potential for regulators to use experiments across many areas of regulatory analysis to improve understanding of how consumers make decisions. There are also many other potentially valuable methods of studying consumer behaviour, however – field trials, consumer surveys, and identification from observational data, to name a few. It is important to continue learning about the relative merits of these methodologies in policy and competition contexts, as applying the right tools to the right questions should help regulators design effective policy that is grounded in better understanding of consumers.

22 See Chapter 4 of the *GIAO study report* (FCA, 2014) for more detail on consumer research for the study and the alignment between findings from experimental, quantitative and qualitative consumer research methods.

References

Adams, P., and Hunt, S. (2013). Encouraging consumers to claim redress: evidence from a field trial. *FCA Occasional Paper 2*. Available at: http://www.fca.org.uk/static/documents/occasional-papers/occasional-paper-2.pdf

Agarwal, S., Driscoll, J. C., Gabaix, X.,and Laibson, D. (2013). Learning in the credit card market. Working paper. Available at: http://ssrn.com/abstract=1091623

Baker, T., and Siegelman, P. (2013). Protecting Consumers from Add-On Insurance Products: New Lessons for Insurance Regulation from Behavioral Economics. *Connecticut Insurance Law Journal*, 20, 13-1.

Bertrand, M., and Morse, A. (2011). Information disclosure, cognitive biases, and payday borrowing. *The Journal of Finance*, 66(6), 1865-1893.

Camerer, C. (2011). The promise and success of lab-field generalizability in experimental economics: a critical reply to Levitt and List. Working Paper. Available at: http://ssrn.com/abstract=1977749

Chater, N., Huck, S. Inderst, R., (2010). Consumer decision-making in retail investment services. *Report by Decision Technology for the European Commissio*n. Available at: http://ec.europa.eu/consumers/strategy/docs/final_report_en.pdf

DellaVigna, S. and Malmendier, U. (2006). Paying Not to Go to the Gym. *American Economic Review*, 96 (3), 694–719.

de Meza, D., Irlenbusch, B., and Reyniers, D. (2010). Disclosure, trust and persuasion in insurance markets, IZA Discussion Paper Series No. 5060.

Duke, C., Huck, S. and Wallace, B. (2010). Experimental Economic Research. A London Economics report for Ofcom, Available at: http://stakeholders.ofcom.org.uk/binaries/research/telecoms-research/experiments.pdf

Ericson, K.M. (2011). Forgetting We Forget: Overconfidence and Prospective Memory. *Journal of the European Economic Association*, 9 (1), 43–60.

Erta, K., Iscenko, Z., Hunt, S. and Brambley, W.(2013). Appling behavioural economics at the Financial Conduct Authority, *FCA Occasional Paper 1*. Available at: http://www.fca.org.uk/static/documents/occasional-papers/occasional-paper-1.pdf

European Commission (2012). Proposal for a Regulation of the European Parliament and of the Council on key information documents for investment products. *Commission Staff Working Document: Impact Assessment*. Available at: http://ec.europa.eu/internal_market/finservices-retail/docs/investment_products/20120703-impact-assessment_en.pdf

Falk, A. and Heckman, J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, 326, 535–538.

Financial Conduct Authority (2014). General insurance add-ons: Provisional findings of market study and proposed remedies. *Market Study* MS14/1. Available at: http://www.fca.org.uk/static/documents/market-studies/ms14-01.pdf

Handel, B. (2011). Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. *American Economic Review*, 103(7), 2643-2682. Online Appendix. Available at: http://emlab.berkeley.edu/~bhandel/wp/Handel_ASIN_OnlineAppendix_2013.pdf

Harrison, G. W., and List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 1009-1055.

Huysentruyt, M., & Read, D. (2010). How do people value extended warranties? Evidence from two field surveys. *Journal of risk and uncertainty*, 40(3), 197-218.

Kessler, J.B., and Vesterlund, L. (2012). "The external validity of laboratory experiments: Qualitative rather than quantitative effects" in Fréchette, Guillaume and Andrew Schotter (editors) *The Methods of Modern Experimental Economics*, Oxford University Press.

Letzler, R., and Tasoff, J. (2013). Everyone Believes in Redemption: Nudges and Overoptimism in Costly Task Completion. Working Paper. Available at: http://ssrn.com/abstract=2066930

Levitt, S.D., and List, J. A. (2007a). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives* 21, 153–174.

Levitt, S.D., and List, J. A. (2007b). Viewpoint: On the generalizability of lab behaviour to the field. *Canadian Journal of Economics* 40, 347–370.

London Economics (2014). Study into the sales of Add-on General Insurance Products: Experimental consumer research. A report for the Financial Conduct Authority. Available at: http://www.fca.org.uk/static/documents/market-studies/gi-add-ons-experimental-consumer-research-report.pdf

Loewenstein, G. (1999). Experimental economics from the vantage-point of behavioural economics. *The Economic Journal*, 109(453), 25-34.

Office of Fair Trading (2010a). The impact of price frames on consumer decision making, *a report by London Economics*. Available at: www.oft.gov.uk/shared_oft/economic_research/OFT1226.pdf

Office of Fair Trading (2010b). What does behavioural economics mean for competition policy? Available: www.oft.gov.uk/shared_oft/economic_research/oft1224.pdf

Office of Fair Trading (2012a). Airlines/Payment Surcharges: Investigation into pricing practices by 14 airlines. Available at: http://www.oft.gov.uk/OFTwork/consumer-enforcement/consumer-enforcement-completed/card-surcharges/

Office of Fair Trading (2012b). Eight supermarkets sign up to OFT principles on special offers and promotions. Press release. Available at: http://www.oft.gov.uk/news-and-updates/press/2012/116-12/#.UvhU9ZWPNaQ

Office of Fair Trading (2013). Partitioned pricing research: A behavioural experiment, *a report by London Economics*. Available: www.oft.gov.uk/shared_oft/economic_research/OFT1501.pdf

Office of Fair Trading and Competition Commission (2009). Road Testing of Consumer Remedies. *A report by London Economic*s. Available at: http://www.oft.gov.uk/shared_oft/economic_research/oft1099.pdf

Oxera (2013). Behavioural economics and its impact on competition policy. A practical assessment with illustrative examples from financial services. Prepared for *The Netherlands Authority for Consumers and Markets*. Available at: https://www.acm.nl/nl/download/bijlage/?id=11161

Sitzia, S., Zheng, J., and Zizzo, D. J. (2012). Complexity and Smart Nudges with Inattentive Consumers. Working Paper. Available at: http://ssrn.com/abstract=2180956

Zizzo, D. J. (2011). Confounds and objectives in economic experiments. Working Paper. Available at: http://ssrn.com/abstract=1909739

Annex 1: Detailed considerations on the validity and limitations of experiments

This annex expands on the discussion in Section 3 by providing further detail on factors that can contribute to the real-world applicability of findings from experiments, especially, in regulatory contexts. It also provides a more thorough analysis of some of the main limitations of experiments as a methodology.

Improving validity of experimental findings

The right balance between detail and abstraction, as well as the key features to capture, will vary across policy contexts. A few specific elements of experimental design may be particularly relevant for improving qualitative external validity and obtaining more realistic quantitative predictions in policy-focused experiments in financial services.

Labelling: For policy purposes it may often be appropriate to frame the tasks in terms of specific product types (albeit in a stylised form). In contrast with academic studies of general effects, policy research is generally concerned with behaviour in a particular market, so the fact that labels may trigger consumers' additional pre-conceptions about the product type of interest can actually be positive factor for external validity. For example, in testing the effects of disclosure on mortgage choice it is helpful if participants think about the task in similar way to a real mortgage decision. Furthermore, financial products are already very abstract and complex for consumers, so increasing abstraction may make it more difficult for participants to engage with the task.

Sometimes, more general patterns in consumer behaviour that apply to more than one product will be of policy interest. In those circumstances it may still be appropriate to use an abstract laboratory setting, as was done in the OFT's advertising of prices study (OFT, 2010a). Alternatively, it may be possible to explore general patterns without increased abstraction by using multiple specific product labels in an experiment and testing for similarities across them. This was part of the rationale for using several different product types, as well as an unlabelled choice, in the FCA's add-on insurance experiment.

Participant demographics: A common rationale for using university students as participants in experimental research is that their generally high cognitive skills and adaptability to abstract experimental conditions helps set a lower bound on the likely size of the errors made in the real world by the general public: if students make mistakes in an experiment, then the errors among a more diverse population of consumers are likely to be at least as large. These arguments can hold for policy-focused research as well, although in some cases an alternative mix of participants can be more desirable for the following reasons:

- **Experience:** Students may often lack experience with financial products, and can therefore be expected to perform worse in experimental tasks than a typical consumer. For example, when dealing with complex investment products, even finance students may lack the practical background knowledge about products that can be assumed for an investor.

- **Vulnerability:** For policy purposes, it may be particularly important to understand whether more vulnerable consumer groups make certain mistakes (or benefit from the remedies being tested), even if the errors are not universal across all consumers. In fact, part of the value of the experiment can be in identifying which types of consumers are particularly at risk of errors - something which requires a wide demographic range of participants.

- **Effects on the population:** Whether the experiment is used to diagnose the root causes of a problem or to test the effectiveness of a remedy, the policy objective may be to obtain a quantitative estimate of the effects of the experimental treatment on the population. For example, for a test of a remedy, the next step in the analysis might be to analyse its benefits if it were rolled out to the market at large. In this case it is important to understand whether the effects are similar across consumers and, if not, what the relative magnitude of the effects might be for different affected groups.

Incentives: To ensure that experimental results are not distorted by the participants' lack of effort, it is important to provide clear performance-based incentives for engaging in experimental tasks. Even with monetary incentives, however, it may be impossible to expose participants to the large-scale consequences of a major financial decision for practical or ethical reasons (for example, participants cannot lose money as a result of bad choices in an experiment). In this context, behavioural biases could help in framing tasks to achieve effort levels closer to those in the real world – for example 'losses' in an experimental environment could be simulated by giving the participants some money at the start of the experiment, and then making deductions from it. Because of reference-dependent loss aversion, it is likely that those payments will be felt as a genuine loss by participants. Similarly, framing experimental rewards in terms of large monetary amounts (which are then converted to more realistic payouts) may help induce higher effort even if the conversion rates are transparent to participants.

Limitations of experiments

The key general limitation of experiments is the extent to which the magnitude of the treatment effects can be extrapolated into the real world, although this may sometimes be mitigated through more realistic design. In some cases, however, the nature of the experimental setting itself may also make it difficult to capture important elements of the consumer decision, which might limit the extent to which experimental methods can be applied to explore consumer behaviour reliably even in qualitative terms.

This section sets out some of the factors that are often said to be difficult to capture experimentally and, where possible, assesses whether these limits can be circumvented. It is important to remember, however, that these constraints apply to the experimental setting as a whole. As discussed in Section 3, the fact that some of the factors listed below may be difficult to reflect in an experiment would only pose a threat to external validity of the results if this departure from reality (confound) could by itself influence behaviour in the same direction as hypothesised effects of the treatments being tested. In this case it is difficult to assert that in the real world, where this confound is not present, the treatments by themselves will still have

the effects observed in the experiment. If the departure from reality is unlikely to influence the effects of the treatments in the experiment, or is likely to weaken them relative to real life, it will rarely be a relevant concern for external validity of the findings.[23]

The typical **time frame** for experiments is perhaps one of the biggest constraints on using this methodology to explore financial decisions. Almost invariably, an experiment involves performing some activities over a single continuous time period (usually around 30 minutes), which is arranged and initiated by the experiment organiser. As a result, it is very difficult to capture decision-making over time, especially where the consumer is responsible for initiating the activity (for example, remembering to check a bank account balance regularly to prevent overdraft use). Some experiments, however, do go beyond the laboratory constraints and can therefore provide more meaningful insights into barriers to consumers taking action over time. For example, experiments by Ericson (2011) and Letzler and Tasoff (2013) explore over-confidence and procrastination by giving consumers choices of different types of mail-in rebates and observing the extent to which consumers follow through with their stated decisions over time.

The relatively short time over which the experiment takes place, and the repetitions of the tasks coming very soon after each other, mean experiments are also often not well-suited to exploring **learning** by consumers. It is common to monitor whether decisions by experimental participants change over the course of the experiments as they become more familiar with the stylised task. However, even if this acquired familiarity translates into better decisions in the experiment, it does not mean that consumers would learn in similar ways from repeated purchases in real-life markets. Firstly, the feedback on the quality of decisions participants receive between tasks can be reasonably transparent, whereas a real consumer may often not get much information about whether they have made an error. Secondly, field research suggests that even though consumers do change their behaviour in financial markets when they get salient feedback (for example, after incurring a fee on their credit card), the effects of this learning decay quite quickly as the memory of the consequences of the mistake fades (Agarwal et al., 2013). There is no scope for this learning decay to happen in an experiment.

As discussed above, the **stakes** of some of the real-world consumer choices also can be difficult to replicate through experimental incentives, even after behavioural framing. At the very least there may remain some uncertainty about the extent to which the effort induced by lower experimental incentives approximates real-world behaviour. Interestingly, however, existing evidence suggests that the size of incentives in experiments does not have a large effect on behaviour in economic experiments. For example, Camerer and Hogarth (1999) find that the main change in behaviour comes from the difference between some incentives and none at all; beyond that the study finds little effect of increasing stakes from moderate to high levels.

Even leaving stakes aside, the **level of engagement** in the task is also something that could differ considerably between experimental settings and real-world decisions. Typically, a participant in an experiment has no other activities competing for their attention which helps ensure they are focused on the task. This may considerably improve the quality of the decisions relative to real-life settings, where consumer attention can wander to irrelevant aspects of the choice or, indeed, to other activities entirely.[24]

23 See Zizzo (2011) for more information about this approach to assessing the materiality of confounds in experimental design, or Kessler and Vesterlund (2012) for another the similar approach.

24 This constraint can, in principle, be mitigated through creative experimental design. A recent experiment by Sitzia et al. (2012), for example, tries to control participants' level of engagement and increase inattention to more realistic levels by giving them more than one task to carry out simultaneously.

Moreover, participants choose to take part in the experiment (even though the details are generally not advertised at the time), which in itself can reveal a greater than average interest in the topic. Generally being able to make a case for stronger engagement of participants than there would be in reality is a positive contribution to external validity (as this suggests identified errors would be no smaller in real-world settings). It can be more problematic in testing remedies, such as some instances of testing disclosure, however, because the level of attention participants might pay to information provided might be too high.

Potential **emotional influences** on choice can be difficult to capture in an experimental settings; for example, the extent to which a decision on buying insurance may be driven by aversion to future regret. Framing the experiment in terms of real products can help elicit these emotions to some extent, but not fully. In some cases, however, this constraint might not undermine external validity if the experiment is designed well – for example, if the products are framed in the same way throughout the experiment, the differences in consumer choices between different treatments can be attributed to those treatments. In other cases, it may be possible to argue that the omitted emotional factors are very likely to strengthen the experimental effects—for example, because the identified errors in decision-making are likely to be worsened when consumer is stressed or is being pressured by a salesperson.

There are concerns about the extent to which the very **fact that decisions are monitored** in an experimental environment might distort them in comparison to natural, unobserved settings – for example, whether people are more likely to act in ways that they think the experimenter expects them to. (Zizzo, 2011; Camerer, 2011). It is worth noting, however, that this criticism applies to other methods where consumers know they are being observed, such as surveys and focus groups.

How large an effect this 'experimenter demand' has is disputed. It is likely to be less strong in circumstances where observation is less obvious – for example, in online experiments rather than a laboratory. The strength of this effect is also likely to vary depending on how easily consumers can second-guess the experimenter's expectations about their behaviour. The potential for this distortion needs to be recognised, however, as this is one of the reasons why traditional experiments are good at measuring relative effects across treatments but not as good at measuring absolute effects.

For example, if the experiment entails searching for and purchasing insurance offers, participants may be more inclined to purchase insurance in the experiment environment than they would be in the real world. The experiment environment itself creates a setting in which buying insurance appears to be the 'expected' or 'appropriate' behaviour. Therefore, the fact 80% of participants purchase insurance in the experiment does not mean that 80% of consumers in the market place will do the same. However, comparing between treatments is still valid: if 80% buy insurance in treatment 1 and 40% buy insurance in treatment 2 then one can expect this relative difference to persist in the real world.

The **format of the decision** in experiments is often limited to computer-based tasks, which makes it difficult to capture some of the ways in which consumers search for or buy products in real life. However, in many cases the online setting can be a useful simplification for problems exploring shopping around and consumer choice, as it may be the most transparent and least time consuming way to engage with the market. The omitted features of other purchase channels would be likely to strengthen the effects in other contexts. Moreover, for many retail markets this computer-focused constraint may not be very problematic, as a large and increasing proportion of consumer activity happens through online channels.

Experiments are not necessarily constrained to computer environments. Where it is necessary to capture an essential part of the decision, it is also possible to invite participants to take part in face-to-face or telephone tasks, as was done, for example, in the FSA's experiment on testing point-of-sale PPI disclosure (de Meza et al, 2010). Experimental tests involving face-to-face interactions can be costly because of a need to remunerate both 'sellers' and 'buyers' as well as funding a venue for their interaction. Non-computer settings also involve relinquishing more control because it is much more difficult to ensure the consistency of the environment between different participants (for example, how information is delivered). In fact, one of the main findings of de Meza et al. was that the difference in skills across the salespersons in the experiment had a much stronger effect on consumer decisions than any of the disclosure treatments that were tested.

Annex 2: Summary table of the results

Table 3: Summary of the results of the add-on insurance experiment discussed in this paper

	Insurance only	Up-front add-on alongside primary product	Add-on revealed at POS	Add-on revealed at POS & easy access to alternatives	Add-on revealed at POS & harder access to alternatives	Yearly price format	Monthly price format
Search							
Number of primary product offers seen		4.9	**5.8**	5.6	5.4	5.6	**5.2**
Number of insurance offers seen	3.8	**4.9**	**2.2**	**5.4**	**5**	4.7	**4**
Proportion of participants buying the first insurance offer they saw	16%	17%	**65%**	**7%**	**15%**	19%	**26%**
Purchasing behaviour							
Proportion of participants who:							
bought insurance	69%	**74%**	**69%**	**85%**	**81%**	77%	76%
bought insurance as an add-on to their primary product		**74%**	**69%**	**25%**	**30%**	38%	39%
Average price paid for all insurance	72	**89**	**102**	**65**	**71**	76	**81**
Average price paid for add-ons		**89**	**102**	**74**	**86**	87	**95**
Price sensitivity: % of available offers above 4.5x actuarial value accepted	3%	**5%**	**23%**	**1.2%**	**2.5%**	3.5%	**6.3%**
Quality of decisions							
Proportion of participants who did not identify the best offer available correctly	4%	**17%**	**24%**	**20%**	20%	16%	17%
Proportion of participants who selected the cheapest primary product instead of the best offer for the overall bundle they bought		2%	**12%**	10%	11%	6%	**8%**
Average amount of money lost from not choosing the best available offer (£):							
for those who bought the primary product and insurance	2.1	**13.2**	**22.4**	**10.2**	11.9	9.8	**14.0**
for those who bought the primary product only		5.8	8.9	**2.6**	5.2	3.4	**5.4**
overall	2.1	**11.3**	**18.2**	**9.1**	10.6	8.3	11.9
Average overpayment by consumers on insurance relative to its actuarially fair value (£)	21	**35**	**42**	**21**	**25**	27	**30**
Proportion of participants changing their mind after placing a bundle into the "shopping basket" and seeing the total cost	4%	**17%**	**21%**	**12%**	12%	7%	**20%**

Note: Cells highlighted in **bold** indicate that the difference between results for the current treatment and the one in the column immediately to the left (where comparable) is statistically significant at 5% level or higher. Table 2 provides guidance on interpreting pairwise comparisons between treatments. More detail about the results of the experiment and the statistical analysis undertaken is provided in the technical experimental report. (London Economics, 2014).

Financial Conduct Authority

PUB REF: 004876

25 The North Colonnade Canary Wharf
London E14 5HS
Telephone: +44 (0)20 7066 1000
Website: www.fca.org.uk