Complexity and Smart Nudges with Inattentive Consumers

Stefania Sitzia
CBESS and School of Economics, University of East Anglia

Jiwei Zheng
CBESS, CCP and School of Economics, University of East Anglia

Daniel John Zizzo*
CBESS, CCP and School of Economics, University of East Anglia

CCP Working Paper 12-13

Version: November 2012

Abstract

In an experiment on markets for services, we find that consumers are likely to stick to defaults and achieve suboptimal outcomes. We unpack two key psychological reasons why they do this – complexity (in terms of non-linearity, number and bundling of tariffs) and consumer inattention -. The complexity induced by product bundling, non-linearity and number of tariffs has an important role, but this is overstated if the explanatory power of inattention is neglected. We show that a 'smart nudge' policy of automatically switching default tariffs can be used to exploit inattention-based consumer inertia to achieve better consumer outcomes.

Keywords: complexity, inattention, defaults, warnings, nudges, services, energy.

JEL Classification Codes: C91, D03, D04, L51.

Complexity and Smart Nudges
with Inattentive Consumers

Stefania Sitzia
CBESS and School of Economics, University of East Anglia

Jiwei Zheng
CBESS, CCP and School of Economics, University of East Anglia

Daniel John Zizzo*
CBESS, CCP and School of Economics, University of East Anglia

Version: November 2012

Abstract

In an experiment on markets for services, we find that consumers are likely to stick to defaults and achieve suboptimal outcomes. We unpack two key psychological reasons why they do this - complexity (in terms of non-linearity, number and bundling of tariffs) and consumer inattention -. The complexity induced by product bundling, non-linearity and number of tariffs has an important role, but this is overstated if the explanatory power of inattention is neglected. We show that a 'smart nudge' policy of automatically switching default tariffs can be used to exploit inattention-based consumer inertia to achieve better consumer outcomes.

Keywords: complexity, inattention, defaults, warnings, nudges, services, energy.

JEL Classification Codes: C91, D03, D04, L51.

1. Introduction

This paper presents a set of three experiments trying to identify *whether*, in markets for services, consumers are likely to stick to defaults and achieve suboptimal outcomes, *why* they do this and *what* can be done about it. In order for consumers to reap benefits from competition, they have to be actively engaged in spotting the best deal that is available to them. This is true both in the tautological sense that they are worse off if they go for a suboptimal choice and in the less obvious sense that firms may be under less competitive pressure if they do so (Giulietti et al., 2005). It is a stylized fact however that, in a number of services markets where choice is possible, consumers do not switch service providers even though the tariffs they are holding are suboptimal (Jamasb and Pollitt, 2005; OFT, 2008; DG Sanco, 2010; Lunn, 2011); furthermore, when choices are made, there is a question mark about whether they are necessarily optimal (Joskow, 2008; Wilson and Waddams Price, 2010). Relevant services markets include both ones that have always been in the hands of the private sector, such as bank account, mobile telephony and internet services, and ones that have been opened up to competition in many countries, such as consumer electricity and gas services, fixed telephony and multichannel TV services.[1] Market models that have introduced these and other deterrents to change supplier by postulating switching costs have shown the distortions that this produces for competition, for example in terms of market entry and prices (e.g., Klemperer, 1995; NERA, 2003).

Undoubtedly, financial switching costs can act as a partial deterrent to changing services supplier in some cases. Identifying the role of different kinds of switching costs can be hard with field data, though important attempts have been made with survey data (Wilson and Waddams Price, 2010) and very little switching, compared to the savings opportunities available, is observed even in markets, such as the U.K. retail electricity and gas markets, where financial switching costs are minimal. Attempts have been made to use survey data to infer non-financial reasons for not switching: the role of complexity in the tariffs employed and in the number of the tariffs employed has been highlighted (e.g., Lambrecht and Skiera,

---

[1] Vulnerable consumers (older, less educated and more disadvantaged consumers) are likely to be especially affected (see DG Sanco, 2009, for some across Europe evidence).

2006; OFT, 2008; Garrod et al., 2009; DG Sanco, 2009; Lunn, 2011) and driven policy recommendations (e.g., Joskow, 2008; Xavier and Ypsilanti, 2008; Ofgem, 2009, 2011; Independent Commission on Banking, 2011). For example, it has been brought as a good reason for why the drive for liberalization of consumer energy markets has halted in USA (Joskow, 2008) and for envisaging requiring tariffs to be simpler in the UK (Ofgem, 2011). Carlin (2009), Gabaix and Laibson (2006), Spiegler (2006) and Ellison and Ellison (2004) provide models explaining how complexity and confusion inducing strategies may be desirable for firms.

The potential role of inattention in explaining suboptimal consumer outcomes has been mentioned, but is, in comparison, somewhat understated.[2] Yet, we suspect that, as with the inattentive agents of Sims (2003) and Reis (2006), although perhaps not necessarily due to a rational allocation of cognitive effort, real life time constrained consumers may simply not pay attention to tasks regarding the choices of services. Putting it simply, it may not be in their minds in the way in which saving 20 cents at a supermarket buying groceries is.[3] A key contribution of this paper is to build on this intuition.

While survey data are insightful and important as they directly refer to real life choices, when it comes to understand the psychological motivations behind behavioral choices, they obviously suffer from potential limitations[4] such a difficulty to draw clear conclusions because of a range of alternative and undeclared factors,[5] forgetfulness and selectivity in

---

[2] Using survey data, Giuletti and Waddams Price (2005) claimed lack of awareness did not play a big role in lack of switching in U.K. energy markets, whereas, using more recent survey data, Wilson and Waddams Price (2010) could not reject that it did, though their evidence is not unequivocal on what did. Oftel (2000) noted inertia/lack of interest and lack of awareness of alternatives as two out of four reasons for not switching supplier in a survey on the telecoms market.

[3] As supermarket shopping becomes increasingly an online shopping experience with default consumer baskets from previous purchases, supermarket shopping might arguably itself become more sensitive to inattention problems. There are a number of other models of economic behavior incorporating inattention, such as Hong and Stein (1999), Hirshleifer and Teoh (2003), Gabaix (2011) and Woodford (2012); Della Vigna (2009) contains a review of some of the implications. The usual interpretation of inattention is in terms of lack of consideration of some features of a product. Inattention could however be in relation to a whole task.

[4] Different studies are obviously affected by specific limitations to different degrees, depending on how the surveys are devised.

[5] For example, Coombs and Shaharudin (2011) criticize survey studies on the suboptimality of banking services supplier choices because of their inability to control for enough alternative explanations. A key reason is that surveys simply do not control enough for the possibility that, given their preferences, consumers may be getting a good deal. In a contingent valuation study with US survey data on electricity supply, Hartman et al. (1991) find a significant status quo bias in terms of stated valuations.

recall,[6] the unconscious nature of many of the choices that people make,[7] and/or the need to self-justify past choices towards those conducting the survey or indeed to engage in self-deception to rationalize possibly suboptimal choices that one has done in the past.[8] A specific problem lies with the fact that, if a significant part of the suboptimality of consumer behavior is because consumers do not pay attention, drawing attention of survey responders to issues they have not thought themselves of before may not be the best way to identify the extent to which inattention is a problem, as ex post rationalizations may then be unavoidable and survey responses may underestimate the role of inattention. A further problem is that it is difficult to see in most surveys whether, as suggested in the context of the number of binary lotteries and of available US 401(k) retirement plans by Iyengar and Kamenica (2010), and more broadly by Beshears et al. (2008), it is the case that complexity may interact with a status quo bias, in the sense that consumers may be less likely to want to take a decision if faced with a more complex decision problem. More seriously, unless natural experiments are possible or ex ante and ex post surveys have been done when structural policy breaks have taken place, it is difficult to test the effects of policy changes with field data.

Our paper addresses these issues by using an experimental methodology. Our first goal is to verify *whether*, in the absence of financial switching costs and using the stylized environment of the U.K. electricity and gas markets as a benchmark, we can identify a lack of switching and suboptimal outcomes when switching does take place.[9] The second goal is to get a better understanding of *why* suboptimal outcomes take place. We test the role of complexity, which we decompose as complexity in the relationship between prices and quantity (linear vs. non-

---

[6] King Li (2011) reviews evidence on selectivity in memory recall and presents an experiment this taking place significantly within just six weeks, if in a different setup.

[7] That there is a split between conscious, explicit knowledge and subconscious, implicit knowledge is a well known stylized fact in psychology (e.g., Shanks and St. John, 1994). Sub-thalamic brain activity takes place when subjects stick with the default, whereas heightened pre-frontal activity takes place when such default is overridden (Fleming et al., 2010), suggesting that, whereas rejecting the default may require a conscious effort, sticking with the default does not. Zizzo (2003) show a dissociation between learning to provide optimal verbal responses and learning to make the optimal behavioral choices.

[8] Psychologists label the tendency for survey responders to provide the responses that they see as putting them in as good light as possible with the researchers the social desirability bias (e.g., Crowne and Marlowe, 1960; Stober, 2001), which can be connected to the desire to receive respect (e.g., Ellingsen and Johannesson, 2007) and both self-image and self-deception as two dimensions of it (Paulhus, 1984).

[9] Cason et al. (2003) describes a market experiment considering the implications of financial switching costs for market structure. Our experiment does not have financial switching costs but rather lets insufficient switching emerge endogenously from consumer decisions (or failures to decide).

linear tariffs), in the presence or absence of bundling (single product vs. dual products tariffs), and in the number of tariffs.[10] We also test the role of consumer inattention by suitably developing a methodology used by Lei et al. (2001) based on the presence of an alternative task, and we consider two possible alternative tasks across different treatments.

Learning why suboptimal choices take place helps us achieve the third goal, which is to test the effectiveness of policies designed to improve consumer outcomes. We are able to evaluate policies putting limits on the number and type of tariffs such as the regulatory constraints on complex tariffs recently proposed by the UK regulator Ofgem (2011). We also test two nudge manipulations that may help outcomes without reducing the consumer's freedom to choose. The first is a simple awareness raising manipulation by which subjects are advised of the existence of a better tariff when they have made a suboptimal choice. This can be connected to Ofgem's (2012) consideration to trial out a 'market cheapest deal' warning scheme, by which companies would be required warn consumers of the existence of a better deal in the market.

The second turns the power of defaults on its head by making it work to achieve better rather than worse consumer outcomes: a 'smart nudge' is employed which automatically identifies the best tariff and uses this as the default choice. This may be connected to the U.K. Prime Minister David Cameron's recent suggestion of forcing energy companies to offer the cheapest of their tariffs (BBC, 2012; Waddams, 2012), but, while these proposals are company specific, our nudge would work in terms of the best tariff *in the market* as opposed to *by a specific company*.[11]

Our key finding is that tariff complexity and the number of tariffs matter, but that inattention matters as well. Regulatory measures to reduce complexity are likely, therefore, to be of only

---

[10] Kalayci and Potters (2011) have an interesting experiment where sellers choose product complexity, in terms of number of attributes of an abstract product, and find some evidence of consumer exploitability, though subject to consumers having to make decisions within 15 seconds; in an experiment again on product complexity (with products modelled as abstract lotteries) but no time constraints, Sitzia and Zizzo (2011) find some qualified (though only qualified) evidence of consumer exploitability. Unlike these experiments, we consider tariff complexity, number of tariffs and product bundling, and we employ tariffs mapped up from a real world markets. Also, in treatments with time constraints subjects do have anyway plenty of time to decide (2 minutes), as verified against a control treatment without such time constraint.

[11] The relationship between our proposal and David Cameron's suggestion will be discussed further in Section 7.

partial value. By using smart nudges and making the power of default work for instead against consumer welfare, we can obtain optimal outcomes around 85% of the times.[12]

The rest of the paper is structured as follows. Section 2 provides some background on the specific markets that our experiments use as model and on the existing evidence on insufficient switching and suboptimal choices. Section 3 describes some common features of the experiment. Sections 4, 5 and 6 are on Experiment 1, 2 and 3 respectively. Section 7 provides a discussion and section 8 concludes.

2. The Institutional Background

The institutional setup on which we model our experiments primarily is the UK electricity and gas markets. These are mature markets which have been liberalized since 1996-1999 and which are comparatively simple in terms of the product they offer (energy). They are also comparatively transparent markets with a wide availability of online search and switching websites.[13] These websites enable both the identification of the best tariffs for any given level of consumption and easy switching of service provider at the click of a mouse. One element of complication is that tariffs can be either for electricity only, or for gas only, or they can be dual tariffs bundling together both electricity and gas; our experiment will focus on electricity only and dual tariffs.[14] As a further complication, the number of tariffs in the market is large: as an illustration, when we collected data for our experiment, we found as many as 72 electricity and 80 dual tariffs in the London, UK, energy market using an online website.[15] 70% of consumers found the number of available tariffs confusing in a UK Ofgem regulator survey (Ofgem, 2008).

Consumers tend to stick to their status quo in terms of energy supplier (NERA, 2003; Ofgem, 2009, 2011; Behavioural Insights Team, 2011): this acts as their default choice. For example, only 18% of all respondents to a Ofgem consumer survey switched electricity supplier in

---

[12] A growing literature is emerging testing and discussing nudging and optimal defaults in a number of other contexts: they include among others Choi et al. (2003), Thaler and Sunstein (2003, 2008), Carroll et al. (2009), O'Neill (2007), Kooreman and Prast (2007), Beshers et al. (2006) and Downs et al. (2009).

[13] Examples include http://www.which.co.uk/switch/, http://www.uswitch.com/, http://www.gocompare.com/gas-and-electricity/ and http://www.confused.com/gas-electricity.

[14] A positive correlation between switching electricity and gas has been found (Giuletti et al., 2005).

[15] Ofgem (2012) contains estimates, as do earlier Ofgem retail market reviews.

2009, and only 17% switched gas supplier (Ofgem, 2010). While 2009 was a year where average prices fell slightly (by 6%), even in a year such as 2008 where average prices went up considerably (by 37%) and so one would have imagined a push to shop around for better deals, only 19% and 20% of gas and electricity customers respectively switched supplier (Ofgem, 2010). In a recent U.K. energy market survey, DECC (2012) reports that only 5% of consumers planned to switch in the next 12 months with just 26% of consumers treating switching as a possibility. While internet penetration and so the access to search and switching engines has increased with time, Jamasb and Pollitt (2005) reported fairly comparable switching figures of just 22% from domestic and small commercial figures in 2002/2003.

Furthermore, when switching takes place, the best (cheapest) tariff is often not chosen. Using data from 2005 and 2000 surveys, Wilson and Waddams Price (2010) estimated that only 8 to 20% of consumers opted for the best tariff given their annual consumption levels. Only around 2/3 of consumers stated that they felt they got a better deal by switching suppliers in 2008 and 2009 (Ofgem, 2010).

While we model key stylized features from the UK electricity and gas markets, we believe that the key issues that we identify and help address are more general. Joskow (2008) discusses the US experience with electricity market liberalization where again there is evidence of consumers failing to switch suppliers with the assumption of the consumers' ability "to shop intelligently" being called into question (pp. 34-35) and being one source of the apparent failure of US consumers to reap much benefits from liberalization. Jamasb and Pollitt (2005) report and discuss limited consumer switching in all EU countries where switching is possible, with the UK actually being the lead with its 22%, and Nunez (2011) notes the failure of ensuring effective competition due to, among other things, the lack of transparency and problems with price setting. An EU DG Sanco study found that electricity switching rates are above 10% in only seven EU countries, with just 32% of EU consumers having compared offers from different suppliers and with average savings to be obtained € 100 per year (DG Sanco, 2010).

Going beyond the energy sector, other services sectors have similar recognized stylized features: the existence of a default option, the potential complexity of the associated tariffs and costs, the limited switching and a significant likelihood of suboptimal switching. Ofcom (2009) reports switching rates roughly equivalent or lower than for electricity and gas in relation to bank account, internet, fixed and mobile telephony services, and multichannel TV services. In a study on bank users, OFT (2008) found that 47% of surveyed consumers had not even considered switching bank account. International evidence on limited switching in telecommunications services is discussed by Xavier and Ypsilanti (2008) and Lunn (2011). This is notwithstanding the fact that online search engines may also be available for such services.[16]

3.   Experimental Design: Some Common Features

The experiments were run at the University of East Anglia in 2011 and 2012. Before the beginning of the experiment, subjects had to read instructions and complete a questionnaire with the purpose of checking they had understood what the tasks involved. If they had any doubts they could ask for clarification. Once questionnaires were collected and doubts clarified the experiment started. All experiments involved individual choices where subjects had repeated opportunities to choose among a set of tariffs. The experimental instructions and details on all tariff tasks are in an online appendix.

The tariffs. In February 2011, we collected all the electricity and gas tariffs available in the UK market as available to a London consumer using the "Which?" website. The tariffs ranged from simple ones with one tier (i.e., a single marginal price) to more complicated ones with two tiers and a ceiling (i.e. a marginal price and, once consumption exceeds a ceiling, a second and lower marginal prices) or a standing charge and one tier (i.e. a fixed price plus a single marginal price). The tariffs in our experiment were partly real tariffs collected in this way and partly derived by us using the same structure as the real ones (*derived* tariffs in what follows). The process of selection and derivation of all tariffs, as well as the full list, is described in detail in an online appendix. We employed 144 tariffs. Two thirds of the tariffs

---

[16] For example, in relation to internet services, http://broadband-comparison.net/ is a price comparator for USA, http://www.comparebroadband.com.au/ is one for Australia and http://www.uswitch.com/broadband/ is one for the UK.

were real and one third was derived. The real tariffs were half for a single service (these are electricity tariffs) and the other half were dual tariffs (both gas and electricity). The derived tariffs were all dual ones. Subjects were only told that the tariffs related either to one good or to two goods (labeled as good A and good B). Table 1 shows a sample of tariffs used.

The tariff tasks. The tasks were 36 overall. At the beginning of each task, in most treatments, subjects were assigned a default tariff, the details of which will be discussed later. They could either stick to the default or look at other available tariffs and then decide which one they wanted. After choosing the tariff, they selected how much they wanted to consume out of 5 possible consumption levels: 1000, 2000, 3000, 4000 and 5000.[17]

The consumption level determined the revenue while the cost depended on the tariff chosen and amount consumed. The revenue was maximized when the optimal consumption level was 4000.[18] After subjects chose their consumption, they were told their earnings, which was determined as revenue minus cost. At the end of the experiment, one of the 36 tasks was chosen randomly and subjects were paid according to the choice made in that task.[19] Average earnings were around 20 pounds.

The default tariff. The default tariff was always a derived tariff designed in such a way that it was never the best to maximize earnings (except in two treatments, discussed later). The difference between the default tariff and the best tariff was usually at least around 6 pounds.

Nature of the tariffs employed in each task. The order of the 36 tasks was randomized. Half of them involved 4 tariffs and the other half 24. For both sets of tasks, 1/3 of the tariffs were single tariffs (all real), 1/3 real dual ones and 1/3 derived dual ones. For each of this subset of tasks, 2 tasks involved a choice among all simple tariffs (only one tier), 2 tasks a choice among all complex tariffs (two tiers plus ceiling, or standing charge plus one tier), and 4 tasks

---

[17] Actual consumers of course do not have pre-defined possible levels of consumption. By having only 5 levels, we wished to keep things as simple as possible in this part of the experiment, however, bearing in mind that actual consumers do have past consumption as a guide to future consumption, and so the level of consumption is not that much of an issue.

[18] The average yearly household electricity consumption in the UK is around 4000 kwh. The gas consumption is approximately 4 times this amount; in the experiment, we scaled this down by a factor of 4 for simplicity.

[19] In all treatments, subjects could use a calculator on the computer screen to help them with their choices of tariffs and consumption levels. The calculator had 4 boxes for inputting consumption levels and the values of tier 1, tier 2, ceiling and standing charge of the tariff they wanted to check the cost of.

a choice among a mix of tariffs, with half of the tariffs being simple and the other half complex.[20]

Number of subjects. We aimed for at least 30 subjects per treatment in each of the three experiments. There were 460 subjects overall.[21]

4. Experiment 1 – Product Complexity

Experiment 1 tests for the effect of complexity – in terms of tariff complexity, number of tariffs and bundled nature of products - on consumer outcomes. It also acts as a control for key features present in the other experiments: specifically, the effect of having a search engine, of having two minutes to make a decision and of having a default tariff as we have implemented it. It has 5 treatments (DE, mDE, D, DF and F). Figure 1 shows the relationships between these treatments and those in the other experiments.

4.1 The Treatments

*Treatment DE (**D**efault tariff and search **E**ngine)*. This treatment had a default tariff and a search engine. In each task subjects were shown the default tariff in a first screen; from this screen subjects could either stick to the default tariff or go to a second screen where they could see all the tariffs involved with the default tariff highlighted (see Figure 3). For each task, if subjects did not make a choice within two minutes, they were assigned the default tariff.[22] When deciding the tariff to select, subjects could use a search engine, which was a stylized version of internet search engines: subjects provided the default tariff details and a consumption level; the search engine then gave the entire list of tariffs with the difference in earnings relative to the default tariff.

*Treatment mDE (**m**ild **D**efault tariff and search **E**ngine)*. This treatment also had the search engine, but the default tariff was implemented more mildly. Specifically, there was no first

---

[20] These four tasks differed depending on the combination of default tariff (simple or complex) and best tariff (again, simple or complex).
[21] There were 30 subjects in the mDE, DE, DEAI and the Experiment 3 treatments (described below), and 50 subjects in the others.
[22] Based on other treatments without this time cutoff, we knew that subjects take around 1 minute to make a choice. We then fixed the 2 minutes cutoff so that we are comfortable that any difference in behavior cannot be attributed to subjects simply not having the time to take a decision.

screen with just the default tariff; rather, in each task we simply showed subjects the screen with all tariffs, with the default one highlighted.

*Treatment D (**D**efault tariff).* This is the same as treatment DE but does not implement a search engine. It therefore allows us to control for the effect of having a search engine.

*Treatment DF (**D**efault tariff and **F**orced choice).* This treatment is the same as treatment D but removes another feature of the market services: namely, the idea that, if you do not make a decision, you are simply stuck with the default. In the previous treatments subjects that did not make a choice within 2 minutes would stick automatically to the default. Treatment DF does not allow for this: subjects do not go ahead until they choose a tariff.

*Treatment F (**F**orced choice).* This treatment makes a final step relative to the DF treatment: unlike the DF treatment, there is no default tariff.

4.2 Results

4.2.1 Overview

Subjects can end up in two ways with a less good a deal than if they had chosen the best tariff: either by sticking with the default tariff (in treatments other than F) or by switching to another suboptimal tariff. Define the *default rate* as the percentage of times subjects stick to the default tariff, and the *suboptimal switching rate* as the percentage of times subjects switch to a suboptimal tariff. In Experiment 1 (and 2, discussed later), the *suboptimal outcome rate* is then defined as the sum of the default rate and suboptimal switching rate.

Table 2 presents default rates, suboptimal switching rates and suboptimal choices for the 5 treatments. In treatment DE the default rate is around 14%. Have a milder default does not make a difference (Mann-Whitney p = 0.89).[23]

Result 1. The search engine is effective in reducing consumer inertia and this translates into more optimal consumer outcomes.

---

[23] All p values in this paper are two tailed. For all bivariate tests, unless specified otherwise, tests are run at the subject level to control for any within-subject non independence of observations. In the DE and mDE treatments, the search engine was used 25% and 32% of the times respectively.

Support. While making no difference in terms of suboptimal switching rate (p = 0.28), having a search engine (DE treatment) clearly reduces the default rate relative to the D treatment (p =0.004). Correspondingly, the suboptimal outcome rate goes down (p = 0.002).

Result 2. When subjects are forced to take action, inertia reduces substantially and this implies more optimal outcomes. Removing a tariff marked as default does not further change the proportion of optimal outcomes.

Support. The default rate in the DF treatment is significantly lower than in the D treatment (p < 0.001). With an unchanged suboptimal switching rate (p = 0.57), this leads to a lower suboptimal outcome rate (p = 0.0002). Conversely, Table 2 shows that this rate is basically identical between the DF and the F treatment (p = 0.68), though subjects who would have stuck with the default rate are replaced by subjects who simply make a wrong choice.[24]

4.2.2 The Role of Complexity

Table 2 also reports averages of our 3 key variables for different dimensions of complexity.[25]

Result 3. Product bundling has a statistically significant effect of about 8% on the suboptimal outcome rate, marginally affecting both default rate and suboptimal switching rate.

Support. In Wilcoxon tests, the suboptimal outcome rate is significantly higher for dual markets than for single product ones.[26] The effects on default and suboptimal switching rates are small enough not to be always significant by treatment, though they are so in the pooled data.[27]

---

[24] In the F treatment there is not, by definition, a default tariff as such. We can however check for the likelihood that the tariff that is the default tariff in the other treatments is chosen, so as to control for the idiosyncratic preference for such tariff, i.e. for the extent that, in the other treatments, the default tariff is chosen not because it is the default. This likelihood is listed in Table 2 as the 'default rate' for the F treatment; it turns out that it is only 3% on average. In Table 2 we classify any other suboptimal choice under the heading of 'suboptimal switching rate'; this is purely for comparability with the other treatments, as obviously there is no default tariff to switch from as such, and is higher than the suboptimal switching rate in the DF treatment (p = 0.016).

[25] For task simplicity, we constructed all tasks in such a way that the optimal tariff was such for whatever level of consumption. That said, we also checked the data on consumption level; on average subjects choose the optimal consumption level 75% of the times.

[26] Wilcoxon p = 0.004, 0.002, < 0.001, = 0.004, 0.01 and p<0.001 respectively for DE, mDE, D, DF, F and the pooled data.

[27] For the suboptimal switching rate, Wilcoxon p = 0.17, 0.008, 0.26, 0.03, 0.02 and < 0.001 for treatment DE, mDE, D, DF, F and pooled data respectively. For the default rate, Wilcoxon p = 0.01, 0.14, < 0.001, = 0.59, 0.06 and < 0.001 for treatment DE, mDE, D, DF, F and pooled data respectively.

Result 4. Having a mix of tariffs of different complexity is sufficient to increase suboptimal outcomes by about 14%, going up to 22% when all tariffs are complex. The result is driven by suboptimal switches rather than by default choices.

Support. Table 2 shows the average performance of subjects according to tariff complexity. Differences in suboptimal outcome rates are significant between the simple and mixed tariffs tasks on the one side, and the complex tariffs on the other side, in the pooled data and for most treatments individually.[28] While not unequivocal, the result is also supported by the regression analysis in an online appendix. It is largely driven by more switches being suboptimal[29] rather than by changes in default rates, though the latter are significant when comparing simple and complex tariffs in the pooled data (Wilcoxon p = 0.01), though not when comparing simple and mixed tariffs (Wilcoxon p = 0.25).[30]

Result 5. Having a higher number of tariffs increases suboptimal switching and consequently suboptimal outcomes by around 23%. This is not purely a random choice effect.

Support. From a glance at Table 2, the default rate is virtually unchanged (23%) between 4 tariffs and 24 tariffs tasks, whereas the suboptimal switching rate jumps up and consequently the suboptimal outcome rate goes up by around 23% (Wilcoxon p < 0.001 for all treatments and the pooled data). This pattern cannot be explained by observing that, if subjects choose at random, they should go for the best tariffs 1/6 less under the 24 tariffs than they do under 4 tariffs (a random choice effect). If this were the case, we would observe the default rate too should go down when there are 24 tariffs, which we do not observe. More broadly, we do not observe a random choice of tariffs.[31]

4.2.3 Other Results

---

[28] In relation to mixed vs. simple tariffs, Wilcoxon p = < 0.001, 0.02, < 0.001, < 0.001, < 0.001 and < 0.001 for treatment DE, mDE, D, DF, F and pooled data respectively. In relation to complex vs. simple tariffs, Wilcoxon p = 0.001, 0.07, < 0.001, < 0.001, < 0.001and < 0.001for treatment DE, mDE, D, DF, F and pooled data respectively.

[29] p <0.001 in all treatments and in the pooled data, for complex (mixed) tariffs vs. simple tariffs, except for treatment mDE, where p = 0.15 (0.004, respectively) and DE for mixed vs simple (p = 0.002).

[30] In relation to mixed vs. simple tariffs, Wilcoxon p = 0.07, 0.38, 0.68, 0.33, 0.96 and 0.25 for treatment DE, mDE, D, DF, F and pooled data respectively. In relation to complex vs. simple tariffs, Wilcoxon p = 0.08, 0.41, 0.22, 0.09, 0.79 and 0.01 for treatment DE, mDE, D, DF, F and pooled data respectively.

[31] We can test this at the level of individual choices by noting that, if subjects simply randomized their choices across outcomes, each tariff should be chosen ¼ of the times in 4 tariffs tasks and 1/24 of the times in 24 tariffs tasks. Both hypotheses are rejected (p < 0.001 in a test of proportions).

Earnings. Subjects sticking to the default tariff gained 56046 points on average. In contrast, on average switchers earned 78804 points (over 6 pounds more) if they got the best tariff and 65800 points (around 3 pounds more) if they did not. Switching was therefore, broadly speaking, a winning strategy, even if the best tariff was not chosen; 'super suboptimal' choices, i.e. switching choices to tariffs worse than the default, were few.[32] An online appendix contains a table with earnings by treatment for all three experiments.

Complexity of default tariff and best tariff. By looking at mixed tariffs tasks, we can isolate whether having a complex default tariff or best tariff as complex makes a difference. The only consistent effect we find is one of complexity of the best tariff on the suboptimal switching rates (and, consequently, the suboptimal outcome rate): if the best tariff is complex, it is harder to spot it and as a result the suboptimal switching rate more than doubles from 19% to 45% on average (Wilcoxon $p < 0.001$). As a large increase applies also to treatments with a search engine, this suggests that the search engine is insufficiently used.[33]

Learning and use of search engine. Table 2, panel (d), compares performance in the first half and the 2nd half of the experiment[34] Unsurprisingly given the picture on earnings, the default rate tends to fall in all treatments, but remains close to 36% in treatment D in the lack of a search engine, whereas for the DE treatment it goes below 10%. The fall in suboptimal switching rate is less pronounced, with a minimum of about 1 switch out of 4 remaining sub-optimal in all treatments. On average 25% and 32% of subjects used the search engine in treatment DE and mDE, respectively, with little variation between first half and second half of each experiment (see an online appendix for more details).

5. Experiment 2 – Inattention to the Task

Experiment 2 considers the impact of inattention to the task on consumer behavior. This is very difficult in an experimental setting because there is a natural bias that subjects have in coming to the lab to *do something*; this is different from households not paying attention to

---

[32] Only 3% of switches were 'super suboptimal' in 4 tariffs mixes, and only 4% in 24 tariffs mixes.
[33] In the DE (mDE) treatment, the suboptimal switching rates are 0.16 (0.27) and 0.46 (0.42) with simple and complex best tariffs, respectively (Wilcoxon $p < 0.001$ and $p = 0.03$ respectively).
[34] The regression analysis in the online appendix controls for experimental period. The online appendix also provides figures on how performance evolved on a period by period basis for all three experiments.

specific tasks, such as choices of services, because it is not part of their weekly or monthly or yearly routines. All treatments are identical to the treatment DE from Experiment 1 except for the following. Our first treatment (DEAI) adds an alternative task for subjects to engage in: specifically, a second computer screen where subjects can browse the internet, check email or Facebook, and the like. Our second treatment (DEA) also has an alternative (unincentivized) task, but now an unpleasant one employed in the real effort experiment literature (e.g., Abeler et al., 2011) to measure psychologically costly real effort, counting 1s in a matrix of 0s and 1s. Our third treatment (DEAD) primes subjects to pay attention to this alternative task, by having this alternative screen on the first screen of each task. Figure 1 shows the relationships between Experiment 2 and the others.

5.1 The Treatments

*DEAI (**D**efault with search **E**ngine and **A**lternative **I**nternet task) treatment*. In this control treatment, subjects had the choice either to freely surf the web or to pay attention to the tariff tasks on another screen. In one computer screen they could browse the web; in the other one they could perform the tariff tasks. If they did not make any active choice of tariff in any period within two minutes, as in the DE (or mDE) treatments the default tariff was selected for them; they were then still required to select their consumption level.

*DEA (**D**efault with search **E**ngine and **A**lternative task) treatment.* In this other control treatment, subjects again had two screens in front of them. In one they could perform the tariff tasks. In the other, they could perform a counting task consisting in counting the 1s in 0-1 matrices (see Figure 4). This is a task deemed unpleasant enough in the real effort literature (as in Abeler et al., 2011) as to be considered a good measure of real effort. In our experiment, and as made clear in the instructions, it was also entirely unincentivized, which means that subjects should have ignored the task and focused entirely on the tariff tasks, on which their earnings depended exclusively. By comparing performance in the DEAI and the DEA treatment, we can verify whether the nature of the alternative task – and the degree it may be pleasant – matters for our results.

*DEAD (**D**efault with search **E**ngine and **A**lternative **D**efault task) treatment.* This is the key treatment of experiment 2. We employed the same counting task as in DEA, but this was now placed on the first screen of each task (see Figure 5). On the same screen subjects also saw the default tariff and so, if they wished, they could choose this tariff in this screen and move straight to the consumption page. Alternatively they could opt to see all the tariffs involved in the task and select the tariff of their choice as usual.

Using the language of Zizzo (2010), our experimental manipulation deliberately employs a *purely cognitive* experimenter demand effect as an experimental tool to make subjects pay attention as a default to the counting task.[35] We would argue that, even with this purely cognitive experimenter demand, the tariffs task is likely to be more salient in the experiment than going to a switching website and changing energy tariffs can ever be for real world households. As a result, our inattention manipulation is likely to simply provide *lower bounds* on the kind of effects that inattention may produce in the real world. The comparison between performance in the DEA and the DEAD treatments will be especially useful in isolating this effect as the alternative task is the same in the two treatments. As a result, a preference for the alternative task would not be able to explain any differential performance between the two treatments.

In further treatments in Experiment 3, we address the potential criticism that subjects may not be inattentive as a result of a purely cognitive experimenter demand effect, but rather may simply be wanting to do what they see that the experimenter wants them to do. This would be a form of *social* experimenter demand effect (Zizzo, 2010), that would be incompatible with our inattention interpretation -. To anticipate our conclusion from section 6.2, Experiment 3 will allow us to reject this interpretation.[36]

---

[35] Again in the language of that paper, this is akin to a legitimate *magnifying glass* use of experimenter demand effects: namely, one that employs demand effects as an artificial tool to replicate in the laboratory real world conditions that would otherwise not be paralleled. A purely cognitive experimenter demand effect relates to the cognitive process by which subjects make sense of an unfamiliar laboratory decision environment, and in this case can be seen to underpin treating the counting task as a default. It does *not* involve a desire to do what is perceived as what an experimenter wishes them to do, which we discuss next.
[36] Additional evidence against this interpretation is also available. During Experiment 2 (and in Experiment 3) we found it useful to add a social desirability questionnaire at the end of the experiment, which can be interpreted as a measure of sensitivity to social pressure (Stöber, 2001) and has been found as capable of

5.2 Results

5.2.1 The Role of Inattention

Table 3 and Figure 6 present default rates, suboptimal switching rates and suboptimal choices for the 3 treatments of Experiment 2.

Result 6. The default rate and suboptimal switching rates, and consequently the suboptimal outcome rate, is not different between the DEAI and the DEA treatments. There is no support for the nature of the alternative task being different.

Support. Table 3 shows that in both cases suboptimal outcome rates are about 50% (Mann-Whitney p = 0.99), roughly equally split between default choice rate and suboptimal switching rate (respectively Mann-Whitney p = 0.48 and 0.49).

Result 7. The default choice rate is significantly higher in the DEAD than in the DEA treatment, and three times as large as in DE. Overall suboptimal outcomes go up by 20% in the DEAD treatment relative to DE. Inattention matters.

Support. Tables 2 and 3 show that that the default choice rate jumps up to 46% in the DEAD treatment (Mann-Whitney p < 0.001 and 0.03 relative to DE and DEA). The suboptimal switching rate is value is lower in DEAD though the effect is marginal or insignificant (Mann-Whitney p = 0.09 and 0.002 respectively relative to DEA and DE). Overall, 63% of outcomes were suboptimal in DEAD, against 43% in DE and 50% in DEA (Mann-Whitney p = 0.05 and 0.08 relative to DEA and DE). That the DEAD treatment has higher default rates and suboptimal outcome rates than DE is also confirmed by the regression analysis in the online appendix. The different default rate between DEAD and DEA can be interpreted in terms of inattention. Further support for the interpretation of the worse performance in DEAD in terms of inattention is provided by Result 8.

Result 8. In the DEAD treatment subjects pay less attention to the tariffs task than in the DEA treatment, and there is a strong correlation between engagement with the alternative task and higher default rate.

predicting behavior (Zizzo and Fleming, 2011). Regression analysis on the default rate described in the online appendix shows that all of the key effects described below survive controlling for social desirability.

Support. In the DEA and DEAD treatments we have a measure of engagement with the alternative task, specifically the percentage of times each subject has played the counting task. Subjects do 164 counting tasks on average in the DEAD treatment, vs. 51 in the DEA treatment (Mann-Whitney p < 0.001).[37] In the DEAD treatment the two variables are strongly correlated, unlike in the DEA treatment (Spearman ρ = 0.56, p < 0.001, in DEAD, vs. 0.17, p = 0.24 in DEA).

5.2.2 The Role of Complexity

Result 9. Product bundling no longer matters when inattention to the task is a problem.

Support. As shown by Table 3, the 8% difference in suboptimal outcome rate with product bundling found in Experiment 1 (Result 3) is almost unchanged with DEAI (Wilcoxon p = 0.10) but reduces to just 3% in DEA (Wilcoxon p = 0.001) and disappears entirely in the DEAD treatment (Wilcoxon p = 0.79).

Result 10. There is again a tariff complexity effect driven mainly by suboptimal switches, but this appears smaller when inattention is a problem.

Support. Table 3 shows that, under DEAD, the Result 4 effect sizes are roughly halved, with an increase in suboptimal outcomes of 7% with mixed tariffs and 10% with all complex tariffs, relative to having all simple tariffs (Wilcoxon p = 0.09 and 0.09, respectively). The effects in DEA and DEAI are of intermediate size.[38]

Result 11. There is again a tariffs number effect driven by suboptimal switching, and consequently the suboptimal outcome rates by around 18%.

Support. As per Table 3, Result 5 on the tariffs number effect is replicated quite robustly. Default rates are virtually unchanged with 24 tariffs, but suboptimal switching rates and

---

[37] The proportion of subjects that played the counting task at least once is also greater in the DEAD treatment (Mann-Whitney, p < 0.001).

[38] For the suboptimal switching rate, Wilcoxon p = 0.05 and 0.002 (all complex vs. all simple tariffs) and 0.005 and 0.001 (mixed vs. all simple tariffs) for treatments DEA and DEAI respectively; for the default rate, p = 0.09 and 0.38 (all complex vs. all simple tariffs) and 0.37 and 0.45 (mixed vs. all simple tariffs) for treatments DEA and DEAI respectively.

consequently suboptimal outcome rates go up substantially in all treatments, including DEAD (Wilcoxon $p < 0.001$ in all cases).[39]

5.2.3 Other Results

The results on earnings and on the complexity of default and best tariffs from Experiment 1 are replicated in Experiment 2.[40] Table 3, panel (d), compares performance in the first half and the 2nd half of the experiment. As in Experiment 1, the fall in default rate is more pronounced than the fall in suboptimal switching rate, which remains virtually unchanged in the DEAD treatment. However, even in the 2nd half of the experiment some 35% of the subjects stick to the default in the DEAD treatment, with the inattention problem basically offsetting the beneficial effect of the search engine in moving from the treatment without (D) to the treatment with search engine (DE; see section 4.2.3). Subjects used the search engine in 16%, 28% and 24% of the times in the DEAD, DEAI and DEA treatment respectively. Worryingly, even in the 2nd half of the experiment only 18% of the subjects used the search engine in the DEAD treatment. If many subjects do not pay attention to the tariffs task in the first place, the scope for the search engine is obviously more limited.[41]

6. Experiment 3 – The Nudges

Experiment 3 has two objectives, using the DEA and the DEAD treatments from Experiment 2 as the benchmark. First, it aims to test the effect of two nudges versions of which could be implemented by policy makers to obtain more consumer optimal outcomes. Second, as commented in section 5.1, it aims to test the interpretation of the DEAD treatment results in terms of subjects simply wanting to do what they see that the experimenter wants them to do, a form of *social* experimenter demand effect.

---

[39] An exercise to check the non-randomness of choices, as we did for Result 5, would yield the same outcome.

[40] Earnings were 55790, 66551 and 78838 points on average for subjects who stuck with the default tariff, switched suboptimally or switched optimally. Considering the mixed tariff tasks, the proportion of suboptimal switches was again roughly twice as large if the best tariff was complex: e.g. 0.238 vs. 0.120 in DEAD (Wilcoxon $p < 0.001$).

[41] The proportion of times subjects engage in the counting task did decrease as the experiment progressed, from 43% to 21in the DEAD. For DEA and DEAI, we do not have that information because subjects played on separate screens so the counting tasks and tariff periods cannot be linked.

A *warning nudge* is one where, at the end of each of the 36 tasks, subjects who achieved a suboptimal outcome were given a message that they could have earned more money had they chosen a different tariff. This nudge could potentially be useful for policy, since policy makers could place as a requirement for companies to specify in energy bills that better tariffs could be available, and indeed Ofgem (2012) is considering trialing this out in the U.K. energy market. The nudge also helps us test for the social experimenter demand effect interpretation: if subjects felt that they were supposed not to focus on the tariffs task because this was not what we wanted of them, by telling them that they could have earned more money by choosing another tariff, and indeed telling them repeatedly and insistently that this was the case if they kept ignoring the tariffs task, we made clear that this was not what we wanted at all. In this case in the DEAD treatment we should expect the default tariff rate to be comparable to that of the DE treatment.

A *default nudge* is one where the best tariff is selected automatically as the best tariff. Subjects are not told that this is the case and can still switch to a different tariff if they so wish. The inattention problem is solved by not requiring subjects to be attentive for achieving better outcomes. The intuition is that, outside the laboratory, if the policy makers were to identify a likely best tariff based on either personal or aggregate information (e.g. individual past consumption), and review this at intervals, this would enable better outcomes while enabling consumers who have a different preference to choose which tariff they actually prefer. Inside the laboratory, the welfare analysis is obviously more straightforward as the best tariff is always such for all subjects.

Figure 1 shows the relationships between Experiment 3 and the others.

6.1 The Treatments

*WDEAD treatment (**W**arning Nudge with **D**efault, search **E**ngine and **A**lternative **D**efault task)*. This treatment is identical to the DEAD treatment (and so with a prominent alternative task: Figure 5) with the only difference that we provide warning nudges whenever subjects make a suboptimal choice.

*WDEA treatment (**W**arning Nudge with **D**efault search **E**ngine and **A**lternative task)*. This treatment is identical to the DEA treatment (and so with an alternative task on a different screen: Figure 4) with the only difference that we provide warning nudges whenever subjects make a suboptimal choice.

*BDEAD treatment (**B**est **D**efault with search **E**ngine and **A**lternative **D**efault task)*. This treatment is identical to the DEAD treatment with the only difference that the default tariff is now the best one.

*BDEA treatment (**B**est **D**efault with search **E**ngine and **A**lternative task)*. This treatment is identical to the DEA treatment with the only difference that the default tariff is now the best one.

6.2 Results

6.2.1 Overview

Table 4 and Figure 7 describe the consumer performance in Experiment 3. In BDEA and BDEAD, the suboptimal outcome rate now coincides with the suboptimal switching rate as sticking to the default tariff is optimal.

Result 12. When inattention was a problem, providing a warning nudge did not help. There is no evidence for a social experimenter demand effect driving the worse performance by subjects in the DEAD treatment.

Support. Table 4 points to a fall in default rate (Mann-Whitney $p = 0.063$) and consequently suboptimal outcome rate in WDEA relative to DEA (Mann-Whitney $p = 0.01$).[42] WDEAD's default rate however is not statistically significantly different than DEAD's default rate (Mann-Whitney $p = 0.48$), and as the suboptimal switching rate is also unchanged (Mann-Whitney $p = 0.42$) this leads to six subjects out of ten still having suboptimal outcome rates in both treatments (Mann-Whitney $p = 0.73$). Default rate and suboptimal outcome rates are however clearly higher than in the DE treatment (Mann-Whitney $p = 0.001$ and 0.03, respectively), a fact confirmed by the online regression analysis. Subjects, therefore, do not

---

[42] The suboptimal switching rate is virtually unchanged (Mann-Whitney $p = 0.64$).

focus on the prominent counting task because they think they are told not to look at the tariffs task, for even when it is highlighted to them again and again that they are making wrong choices, they do not change their behavior.[43]

Result 13. A default nudge is effective in achieving better consumer outcomes.

Support. Table 4 shows that, with a default nudge, just around 15% of outcomes is suboptimal, with the large majority of subjects sticking to the default. The suboptimal outcome rate in BDEAD are significantly lower than in DEAD, WDEAD or DE (Mann-Whitney $p \leq 0.001$ in all cases), and those in BDEA are significantly lower than in DEA, WDEA or DE (Mann-Whitney $p < 0.001$ in all cases). This is confirmed by the online regression analysis. No other treatment across all three experiments comes as close to an optimal outcome as these treatments.

6.2.2 The Role of Complexity

We first consider the effect of complexity in WDEAD and WDEA, and then move on to BDEAD and BDEA.

Result 14. In the presence of a nudge warning, product bundling matters only marginally when inattention to the task is a problem.

Support. In WDEA, there is a 7% increase in suboptimal outcome rate with product bundling roughly in line with DEA and Experiment 1 (Wilcoxon $p = 0.02$). The WDEAD there is 5% increase, a small effect just enough to achieve a marginal significance level (Wilcoxon $p = 0.06$).

Result 15. In the presence of a nudge warning, there is again a tariff complexity effect driven mainly by suboptimal switches, and again this is smaller when inattention is a problem.

Support. This result simply replicates Result 10. In WDEAD, the increase in suboptimal outcomes is only 3% and 9% in tasks with mixed tariffs and in tasks with all complex tariffs

---

[43] One could argue that, in a social experimenter demand effect interpretation of the prominent alternative task, the existence of repeated and insistent warnings should be especially effective in the 2nd half of the experiment, when subjects not attending to the tariffs task will have received a number of them. However, as discussed below in section 6.2.3, there is no evidence of this.

respectively, in comparison to the case of all simple tariffs (Wilcoxon p = 0.42 and 0.06, respectively). The corresponding effects in WDEA are instead 6% and 16%.[44]

Result 16. In the presence of a nudge warning, there is again a tariffs number effect driven by suboptimal switching, and consequently an increase in the suboptimal outcome rate by 14% (in the WDEAD treatment) or 22% (WDEA treatment) as the number of tariffs goes from 4 to 24.

Support. As per Table 4, Results 5 and 11 on the tariffs number effect are again replicated fairly robustly. Default rates are similar with 24 tariffs, but suboptimal switching rates and consequently suboptimal outcome rates go up substantially in both WDEA and WDEAD (Wilcoxon p < 0.001 and p = 0.003 respectively).[45]

Result 17. With a nudge default, there is no effect of product bundling or of tariff complexity, and at most only a small effect of the number of tariffs.

Support. Table 4 shows that suboptimal switching/outcome rates are virtually identical with product bundling in both BDEAD and BDEA. There is a 3-5% effect in moving from all simple to all complex tariffs.[46] There is a statistically significant effect in BDEAD in moving from 4 to 24 tariffs (Wilcoxon p = 0.004), but a small one (6%) and not replicated in BDEA (2%, Wilcoxon p = 0.53).

6.2.3 Other Results

The results on earnings and on the irrelevance of the complexity of default tariffs from Experiments 1 and 2 are replicated in Experiment 3.[47] As in the earlier experiments, there is an effect by which, if the best tariff is simple in a mix of tariffs, this helped subjects, but, in

---

[44] For the suboptimal switching rate, Wilcoxon p = 0.04 (all complex vs. all simple tariffs) and 0.02 (mixed vs. all simple tariffs)   for WDEAD; p = 0.06 (all complex vs. all simple tariffs) and 0.12 (mixed vs. all simple tariffs) for WDEA. For the default rate Wilcoxon p = 0.58 (all complex vs. all simple tariffs) and 0.27 (mixed vs. all simple tariffs) for WDEAD; p = 0.24 (all complex vs. all simple tariffs) and 0.36 (mixed vs. all simple tariffs) for WDEA

[45] An exercise to check the non-randomness of choices, as we did for Result 5, would yield the same outcome.

[46] For the suboptimal switching rate, Wilcoxon p = 0.04 and 0.29 (all complex vs. all simple tariffs) and 0.13 and 0.16 (mixed vs. all simple tariffs) for treatments BDEAD and BDEA respectively; for the default rate, p = 0.07 and 0.29 (all complex vs. all simple tariffs) and 0.09 and 0.16 (mixed vs. all simple tariffs) for treatments BDEAD and BDEA respectively.

[47] Earnings were 55759, 65306 and 78613 points on average for subjects who stuck with a suboptimal default tariff, switched suboptimally or ended with the best tariff.

the light of the power of the nudge default, in this case the suboptimal outcome rate goes down by just 6-8%.[48]

Table 4, panel (d), compares performance in the first half and the 2nd half of the experiment. In WDEAD, the suboptimal outcome rate in the 2nd half of the experiment (57%) remains stuck at around the level which we found in the DEAD treatment (or the D treatment from Experiment 1). This confirms the ineffectiveness of the warning nudge and, since subjects sticking to the default tariffs will have received a large number of such repeated and insistent warnings by then, provides further evidence for the implausibility of the claim that subjects simply felt that the experimenter wanted them to ignore the tariffs task. The use of the search engine was also virtually indistinguishable (20%) from that in the DEAD treatment.[49]

Since the default tariff is the best tariff in BDEAD and BDEA, it is not surprising but is nevertheless reassuring that in these two treatments subjects tend to stick to the default tariff more with time.[50]

7. Discussion

7.1 Institutional Setup

We began our introduction by noting that our goals were to identify (1) *whether*, in markets for services, consumers are likely to stick to defaults and make suboptimal choices, (2) *why* they do this and (3) *what* can be done about it. To explore these questions, we chose the U.K. electricity and gas market as the stylized setup for our experiments, as this is a mature market in which financial switching costs are already minimal and search engines enable finding best deals virtually at the click of a mouse. In this sense, it represents a good model to answer the first question in a controlled laboratory environment in which alternative explanations – such

---

[48] Wilcoxon p = 0.27, 0.11, 0.05 and 0.04 for treatments WDEAD, WDEA, BDEAD and BDEA respectively.

[49] In the WDEA, BDEAD and BDEA treatments, the corresponding numbers were 39%, 12% and 18%. Note that subjects did an average of 120, 57, 159 and 41 counting tasks in the WDEAD, WDEA, BDEAD and BDEA treatments (Mann-Whitney p = 0.04 for WDEAD vs. WDEA and 0.03 for BDEAD vs. BDEA); the proportion of subjects that played the counting task at least once is greater in WDEAD than in WDEA (p = 0.01) and greater in BDEAD than in BDEA (p < 0.001). The divide between WDEAD and BDEAD on the one side and WDEA and BDEA on the other broadly replicates Result 8.

[50] Spearman ρ (period, default choice rate) = 0.32 (p = 0.06) and 0.54 (p < 0.001) for BDEAD and BDEA respectively.

as financial switching costs or whether consumers actually made the choice that is best for them – can be ruled out.

We used stylized real world tariffs, or stylized tariffs derived from real world tariffs, and we varied the number of tariffs, their linear or non-linear structure, and whether they are bundled up or not, in ways that are realistic and which present issues for tariffs for other services and countries, as is apparent for example from Joskow's (2008) analysis of the US consumer energy markets, Lunn's (2011) review for telecoms markets, European Commission (2007) for European wide evidence on banking services, or DG Sanco's (2009) study across a range of services across European markets. Undoubtedly, further research based on other real world institutional setups would be useful.

7.2 Are Outcomes Suboptimal?

The answer to our first question, emerging already from Experiment 1, is that a significant fraction of consumers does tend to make suboptimal choices, either because of sticking to a default or because of switching to a suboptimal choice. In our paradigmatic DE treatment where there is a default and a search engine but only one activity available, even with just 4 tariffs about 1/3 of the choices are suboptimal, rising to over a half when there are 24 tariffs (Table 2). Note though that subjects stick to the default in only around 15% of cases, which does not seem to fit with real world stylized facts regarding the percentage of consumers not switching (e.g, DECC, 2012). One key reason of difference is that real world consumers may simply not pay attention to saving money from switching energy supplier: their routine activities in their everyday life are more prominent. There is not a point in time in the day, the week, the month or even the year where, as a routine, subjects are required to pay attention to the task of choosing energy supplier, as there is anyway a default energy supplier; there is no equivalent of, say, the weekly major supermarket shopping trip that a household may do every Saturday morning in order not to run out of food. Conversely, subjects come to the laboratory with an expectation that they need to pay attention and engage in a task (see Lei et al., 2001) and it is no surprise that, given the availability of a search engine, they use it to get to much better outcomes, as we would expect with real world consumers as well. The question then becomes why, in the real world, consumers do not use search engines in an

equally effective way. Our intuition is that, because consumers do not pay attention, they often do not get to the stage where they are faced with a search engine: the problem may be made simple but this is not enough if it is simply not in their minds.

Experiment 2 enabled us to capture more precisely inattention-to-the-task issues by adding either a not prominent task (DEAI, DEA) or a prominent (DEAD) alternative task for subjects to engage in *while retaining the search engine*. The prominence of the alternative task is used as a tool to potentially induce inattention, if much less than what can be expected in the real world.[51] We tested the alternative interpretation that the prominent alternative task may make subjects believe that experimenters want them not to engage in the tariffs task. The WDEAD treatment provides direct evidence against this interpretation: this treatment differed from DEAD in that subjects sticking to the default task would be insistently and repeatedly told that they could have made more money by choosing an alternative task, and so, if anything, they should be led to believe that they should focus on the tariffs task. However, even in the second half of the experiment - after potentially 18+ such warnings -, there is no noticeable difference in consumer performance relative to the DEAD treatment (Result 12).[52]

In our DEAD treatment, as many as around 45% of choices stuck to the default, and even with just 4 tariffs over half of the choices are suboptimal. This is closely connected to the degree they pay attention to the alternative task (as per Result 8), which fits with the inattention interpretation of our manipulation. Given the presence of the search engine and the unincentivized nature of the alternative task, this is a surprising finding.

7.3 Why Are Outcomes Suboptimal?

We tackled the second question by aiming to identify the causal role of complexity and inattention, as well as their combination, in explaining suboptimal outcomes. We interpret inattention as inattention to a task, as opposed to inattention to some detail of a task. Of

---

[51] Given the experimental constraints, the instructions need to be clear about the nature of the tariffs task and this remains one of two possible tasks available to subjects. This, together with the natural bias of experimental subjects to do things, implies that what we have is just a lower bound to the type of real world inattention problem that we are trying to model in the laboratory.

[52] Additional evidence against this interpretation is referred to in a footnote in section 5.2.1, where there is also a further discussion of the relationship between our design and experimenter demand effects using the conceptual distinction between purely cognitive and social experimenter demand effects presented by Zizzo (2010).

course, complexity and inattention – even interpreted as inattention to a task - may be connected concepts: an environment which is complex is one where subjects may be more put off from bothering with cognitive effort. We could control for alternative explanations of suboptimal outcomes – e.g., financial switching costs or uncertainty of prospects in switching – by removing them from the laboratory environment.

In line with research emphasizing the potential of consumer exploitation due to complexity (e.g., DG Sanco, 2009, or Carlin, 2009), we found complexity effects involving product bundling, whether the tariff is linear or non-linear, and whether there are 4 or 24 tariffs. However, by neglecting the role of inattention, the role of complexity is overstated. The effect of product bundling virtually disappears once inattention is a problem (Results 9, 14). An effect of tariff complexity remains but is smaller when inattention is a problem (Results 10, 15): e.g., moving from a mixed tariffs decision environment like that currently present in UK energy markets to one with all simple tariffs – a greater simplification than may be possible or indeed desirable once one considers the potential costs that firms may incur by having less tariffs – would only reduce the suboptimal outcome rate by around 5% (DEAD and WDEAD in Tables 3 and 4). The number of tariffs effect is the most robust in affecting the choices of switchers, but might again require a complexity reduction (to 4 tariffs) beyond what may be realistic; furthermore, if a nudge default remedy is introduced to deal with the inattention problem, only a small effect remains (Result 17). Inattention is a problem because it implies that many subjects stick to the default, and this explains most of the suboptimal outcomes in treatments with an inattention problem such as DEAD and WDEAD, with default rates of the order of 40-45% and suboptimal switching rates of just around 15-20% (Tables 3 and 4). In essence, complexity matters, but economists and policy makers should pay more attention to the role of inattention[53] for tasks that do not fit in the usual household consumption routines.[54]

---

[53] Chetty et al. (2009) have noted the connection between salience and attention in terms of features of a task in an important paper on consumer responses to taxation.

[54] This last qualification is important. Our conclusions apply to the use of services with a default choice and which are not part of the regular routines of consumers in terms of what they pay attention to. Buying groceries at a (physical) supermarket would not qualify. Equally, buying a car would not qualify because, while not part of a usual routine, it does require an active choice (there is not a 'default car' which is bought automatically

7.4 What Can be Done about Suboptimal Outcomes?

We can answer this question in terms of direct effects on consumer welfare. Obviously any policy measure would also need to be looked at in terms of its effects on competition. By unpacking the psychological determinants of switching costs, however, the experimental methodology does allow to provide clear-cut messages on the potential effectiveness of measures either tackling them or not. Furthermore, the direction in which consumer welfare in our experiment is clear, unlike survey studies where, as per Coombs and Shaharudin's (2011) critique, it is not necessarily obvious whether consumers may be getting a good deal after all.

A radical and traditional paternalistic way of answering a large problem of consumer choice failure, with its direct and indirect welfare losses, is to remove the choice from the consumers by having a regulated monopolist. This is arguably one reason justifying the slowing down or halting of liberalization in consumer energy markets in USA (Joskow, 2008). This was not the approach considered in this paper.

We are able to evaluate milder, if still traditional, paternalistic policy measures restricting the number of tariffs, the linear vs. non-linear nature of tariffs, and the bundling of tariffs. The bundling of two tariffs together does not have much effect. Restricting the number of tariffs or forcing them to be linear improves consumer outcomes. However, even restrictive regulatory measures forcing tariffs to be linear and only four – with the potentially distorting effects on competition that such restrictions may have – would still only help partially, as consumers would need to pay attention to the choice of tariffs and many of them simply would not. In this sense, the scope of Ofgem's (2012) proposal to limit the number of tariffs provided *by each firm* to 4 per fuel, meter and payment type, will be only of partial help, as the number of tariffs in *the market as a whole* is likely to remain above our experimental upper number of 24 tariffs. Equally, while online search engines are useful and their use should be encouraged by policy makers in markets where they are not so common, their

---

unless an action is taken). The more active choices are taken, the more the potential role of complexity in affecting consumer outcomes.

usefulness is limited if consumers do not pay attention to the task and stick to the default as a result.

Experiment 3 tested two nudge remedies to try to tackle the inattention problem. The first was a warning that there exists a better tariff when a suboptimal tariff is chosen, of the kind that could be required of companies to put on the bills. However, we found no evidence of an improvement in consumer outcomes as a result of this nudge when inattention is a problem.[55] Undoubtedly the nature of the warning may matter, and future research needs to look at the effect of differently framed warnings, e.g. providing precise information on the amount of money that could have been earned if the best tariff was chosen..

The second is a 'smart nudge' which automatically identifies the best tariff and uses this as the default choice: rather than requiring consumers to be attentive, it provides them with an optimal default if they are inattentive, while leaving consumers with the freedom to choose an alternative tariff were they to so wish. This was sufficient to achieve optimal consumer outcomes 86% of the times.[56] The idea of a nudge default for consumer tariffs is in line with research on optimal defaults as developed in the context of retirement plans, such as Choi et al. (2003) and O'Neill (2007), and in our view can strike an appropriate compromise between limited regulator knowledge and need to achieve better consumer outcomes. Again, future research may be desirable, for example on the likely response of firms to the introduction of a policy based on this.

The results of our experiment can inform practical policy measures being proposed. For example, in the context of the U.K. consumer energy market, Ofgem (2011) proposed two types of tariffs, standard tariffs of which each company can have only three, and non-standard tariffs, which are of unlimited number but with automatic default to a standard tariff. According to these proposals, Ofgem would determine a standing charge plus require firms to

---

[55] Cronqvist and Thaler (2004) discuss the Swedish experience of encouraging active decisions for retirement plans, a policy that backfired and was abandoned in the light of suboptimal choices being made as a result.

[56] While giving the freedom to make suboptimal choices in our environment may not seem to make sense, it could be justified in a number of ways: (a) it avoids a traditional paternalistic solution of forcing consumers to make a choice; (b) it enables competition among multiple firms rather than creating a regulated monopolist; (c) real world consumers are heterogeneous in preferences and may, e.g., care about non price dimension of real world tariffs (e.g., 'green' energy based), and, if they do, in a 'smart nudge' environment, they would still be able to act on this preference from switching from the automatically determined default choice.

fix a single unit price on top of it. While attempting to tackle complexity, apart from requiring Ofgem to set prices, there is a question mark on the effectiveness of these measures as there would still be a large number of tariffs, including complex tariffs, that subjects would need to negotiate, and the defaulting to potentially expensive standard tariffs after a fixed term would not be picked up by online search engine. Ofgem's (2012) more recent proposals require all tariffs to be of the form of a standing charge and unit price (both determined by firms), which would make them complex from our viewpoint though Ofgem thinks of them as a simplification. However, even if we believed that these measures would be successful in tackling complexity issues, we would still expect significant suboptimal consumer outcomes as inattentive consumers stick to suboptimal defaults.

A crucial issue from a policy viewpoint is therefore to consider how to raise consumer attention or otherwise deal with consumer inattention. David Cameron's suggestion of giving customers the cheapest tariff offered by each firm (BBC, 2012) could be considered as a 'smart nudge' in the sense that it could be interpreted as each company being required to give their best tariff as default one to consumers. A key problem with it though is that it is not clear what incentive it would give to firms to provide lower tariffs, and in this sense it would risk soften the competition rather than making it tougher (Waddams, 2012). Our 'smart nudge' proposal is more radical and would rely on the concept of market's cheapest deal that is being developed by Ofgem (2012). While obviously further research is needed to evaluate its competitive effect and company costs that may be passed on to consumers, it would create a significant incentive for firms to undercut each other insofar as this would enable them to take business from competitors.[57]

7.5 Methodological issues

It is worth concluding this section by taking stock of three methodological issues that our experiments raise. First, they presented different methods of identifying experimental complexity and we saw that their effect is different, with bundling of two tariffs presenting

---

[57] Note that consumers would, of course, retain the freedom to switch to a different tariff. While this is unequivocally suboptimal in our BDEA and BDED treatments, it may not be so in the real world, where consumers have a preference for e.g. green tariffs.

very few problems. In evaluating models of complexity or identifying policy recommendations, the devil may lie in the detail of *what* complexity is being affected.

Second, what we provide is a methodology to study inattention in an experimental setup, based on more or less prominent alternative tasks, which in our view may be usefully applied in entirely different settings where inattention is considered a significant issue. For example, one may argue that the first task that a charity needs to overcome in encouraging giving is to make people pay attention to what it is about; or that, in real effort experiments, one needs to take into account the distractions that alternative task (e.g., browsing the internet) may have.

Third, we do not focus in this paper on whether inattention is rational in the sense that bounded-rational agents would optimally choose to allocate effort in that way (e.g., Sims, 2003; Gabaix, 2011; Woodford, 2012).[58] However, we note that subjects lost over 6 pounds on average (equivalent to over 9 U.S. dollars) by sticking to the default tariff, which is a significant amount by the standard of the university students that made our sample. Conversely, they knew they could not earn any money from the alternative task, and, while in the case of internet browsing (our control alternative task) it is possible to argue that subjects may do it as a pastime, this does not look equally plausible with the counting task, which has been used in experiments as a tool to measure psychologically costly real effort (e.g., Abeler et al., 2011); nor do we find any significant difference in performance depending on the nature of the alternative task (Result 6). It is hard to avoid the conclusion that subjects should have paid attention only to the tariffs task, given the sizeable incentives offered in that task and that task only. The fact that the alternative task still induced more suboptimal choices in the way it did at least suggests that the allocation of attention might not *always* be optimal, which reinforces the policy message of the need to take inattention seriously. Obviously, further research is needed.

8. Conclusions

---

[58] Cherumkin et al. (2011) try to fit Sims' (2003) rational inattention model to a dataset on risky choice. Goecke et al. (2011) find that, the more expensive information is, the less subjects collect it in a macroeconomic forecasting experiment.

We found that, in markets for services and even in the presence of a search engine, consumers are likely to stick to defaults and achieve suboptimal outcomes. The experiment aimed to unpack two key psychological reasons why they do this – complexity (in terms of non-linearity, number and bundling of tariffs) and consumer inattention. By employing an experimental methodology, we are in a position not only to identify the causal role of different psychological dimensions, but we are also able to test the effectiveness of policies designed to improve consumer outcomes. Our experiment, and our tariffs, are inspired by stylized features of UK electricity and gas markets, but the lessons we draw are likely to be more general, as both underlying features (such as non-linear tariffs and the presence of defaults) and psychological mechanisms are obviously more general.

Task complexity matters. In our experiment, this is a function of product bundling, tariff complexity and of number of tariffs. However, in the presence of a default tariff and of consumer inattention, markets are affected by large amounts of consumer inertia. Similarly, providing a warning on the existence of a better tariff does not improve outcomes when inattention is a problem, though further research is needed to look at alternative warning messages that one can provide. The reason why reducing complexity of the task solves the consumer inertia problem only partially is because subjects do not pay enough attention to the task in the first place and as a result just stick to the default.

A 'smart nudge' policy of automatically switching default tariffs was a pragmatic and effective policy solution in our experiment to obtain better consumer outcomes. This policy solution is in the same spirit, but more radical than, the one suggested by David Cameron in the context of the U.K. energy market, and, while further research is needed, it is likely to create a greater competitive pressure than the latter. It automatically changes default tariffs to the optimal one in a given time period, and by doing so it exploits inattention-based consumer inertia to achieve better consumer outcomes while leaving consumers free to choose an alternative tariff if so they wish.

References

Abeler, Johannes, Falk, Armin, Goette, Lorenz and David Huffman. 2011. Reference Points and Effort Provision. American Economic Review 101(2), 470-492.

BBC, 2012. David Cameron Stands by 'Lowest' Energy Tariff Vow. www.bbc.co.uk/news/ uk-19986929

Behavioural Insights Team, 2011. Better Choices: Better Deals. London: BIS and Cabinet Office.

Beshears John, Choi, James J., Laibson, David L. and Brigitte C. Madrian, 2006. Simplification and savings. NBER Working Paper 12659.

Beshears, John, Choi, James J., Laibson, David and Brigitte C. Madrian, 2008. How are preferences revealed? Journal of Public Economics 92(8-9), 1787-1794.

Carlin, Bruce I., 2009. Strategic Price Complexity in Retail Financial Markets. Journal of Financial Economics 91(3), 278-287.

Carroll, Gabriel D., James J. Choi, David Laibson, Brigitte Madrian and Andrew Metrick, 2009. Optimal Defaults and Active Decisions, Quarterly Journal of Economics 124(4), 1639-1674.

Cason, Timothy N., Daniel Friedman and Garrett H. Milam, 2003. Bargaining versus Posted Price Competition in Customer Markets. International Journal of Industrial Organization 21(2), 223-251.

Cheremukhin, Anton, Anna Popova and Antonella Tutino, 2011. Experimental Evidence on Rational Inattention. Federal Reserve Bank of Dallas working paper.

Chetty, Raj, Adam Looney and Kory Kroft, 2009. Salience and Taxation: Theory and Evidence. American Economic Review 99(4): 1145-77.

Choi, James J., David L. Laibson, Brigitte C. Madrian and Andrew Metrick. Optimal Defaults. American Economic Review Papers and Proceedings 93(2), 180-195.

Coombs, Justin and David Shaharudin, 2012. EU Substantive Areas: Economics. European Antitrust Review 4(1), http://www.globalcompetitionreview.com/reviews/37/sections/ 130/chapters/1378/economics/

Cronqvist, Henrik and Richard H. Thaler, 2004. Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience. American Economic Review Papers and Proceedings 93(2), 424-428.

Crowne, Douglas P. and David Marlowe, D., 1964. The Approval Motive. New York: Wiley.

DECC, 2012. DECC Public Attitudes Tracker – Wave 2. Summary of Key Findings. London: U.K. Government.

DellaVigna, Stefano, 2009. Psychology and Economics: Evidence from the Field. Journal of Economic Literature 47(2), 315-372.

DG Sanco, 2009. Consumers' Views on Switching Service Providers. Brussels: European Commission.

DG Sanco, 2010. The Functioning of Retail Electricity Markets for Consumers in Europe. Brussels: European Commission.

Downs, Julie S., George Loewenstein and Jessica Wisdom, 2009. Strategic for Promoting Healthier Food Choices, American Economic Review Papers and Proceedings 99(2), 1-10.

Ellingsen, Tore and Magnus Johannesson, 2007. Paying Respect. *Journal of Economic Perspectives* 21(4), 135-150.

Ellison, Glenn and Sara Fisher Ellison, 2004. Search, Obfuscation, and Price Elasticities on the Internet. Discussion Paper, MIT and NBER.

European Commission, 2007. Report on the Retail Banking Sector Inquiry. Brussels: European Commission Directorate-General for Competition.

Fleming, Stephen M., Charlotte L. Thomas and Raymond J. Dolan, 2010. Overcoming Status Quo Bias in the Human Brain. Proceedings of the National Academy of the USA 107(13), 6005-6009.

Gabaix, Xavier and David Laibson. 2006. Shrouded Attributes, Consumer Myopia, And Information Suppression In Competitive Markets. Quarterly Journal of Economics 121(2), 505-540.

Gabaix, Xavier, 2011. A Sparsity-Based Model of Bounded Rationality. NBER Working Paper 16911.

Garrod, Luke, Hviid, Morten, Loomes, Graham and Catherine Waddams Price, 2009. Competition Remedies in Consumer Markets. Loyola Consumer Law Review 21(4), 439-495.

Giulietti, Monica, Catherine Waddams Price and Michael Waterson, 2005. Consumer Choice and Competition Policy: a Study of UK Energy Markets. Economic Journal 115(506), 949-968.

Goecke, Henry, Wolfgang J. Luhan and Michael W.M. Roos, 2011. Rational Inattentiveness in the Lab: The Effect of Information Costs on Forecasting. TU Dortmund and Ruhr Universitat Bochum working paper.

Hartman, Raymond S., Michael J. Doane and Chi-Keung Woo, 1991. Consumer Rationality and the Status Quo. *Quarterly Journal of Economics* 106(1), 141-162.

Hirschleifer, David and Siew Hong Teoh, 2003. Limited Attention, Information Disclosure, and Financial Reporting. Journal of Accounting and Economics 36(1), 337-386.

Hong, Harrison and Jeremy C. Stein. 2007. Disagreement and the stock market. Journal of Economic Perspectives, 21(6), 109-128.

Independent Commission on Banking, 2011. Final Report Recommendations. London: U.K. Government.

Jamasb,Tooraj and Pollitt, Michael, 2005. Electricity Market Reform in the European Union: Review of Progress toward Liberalization & Integration. Energy Journal 26 (Special Issue), 11-41.

James J. Choi, David Laibson, Brigitte Madrian and Andrew Metrick, 2003. Optimal Defaults. American Economic Review Papers and Proceedings 93(2), 180-185.

Joskow, Paul L., 2008. Lessons Learned From Electricity Market Liberalization. Energy Journal 29(Special Issue 2), 9-42.

Kalayci, Kenan and Potters, Jan, 2011. Buyer Confusion and Market Prices. International Journal of Industrial Organization 29(1), 14-22.

King Li, King, 2011. Asymmetric Memory Recall of Positive and Negative Events in Social Interactions. Max Planck Institute of Economics working paper.

Klemperer, Paul, 1995. Competitition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade 62(4), 515-539.

Kooreman, Peter and Henriette Prast, 2007. What Does Behavioral Economics Mean for Policy? Challenges to Savings and Health Policies in the Netherlands. Tilburg University working paper.

Lambrecht, Anya and Skiera, Bernd, 2006. Paying Too Much and Being Happy about It: Existence, Causes and Consequences of Tariff Choice Biases. Journal of Marketing Research 43(2), 212-223.

Lei, Vivian Charles N. Noussair and Charles R. Plott, 2001. Non-Speculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality. Econometrica 69(4), 831-885.

Lunn, Peter, 2011. Telecommunications Consumers: A Behavioural Economic Analysis. ESRI Working Paper No. 417, December.

NERA,2003. Switching Costs. Part One: Economic Models and Policy Implications. London: OFT and DTI.

Nunez, Alberto, 2011. Liberalisation of the Electricity Sector in the European Union: Present State and Some Open Questions. Barcelona: Gas Natural SDG working paper.

O'Neill, Barbara, 2007. Overcoming Inertia: Do Automated Saving and Investing Strategies Work? Journal of Family and Economic Issues 28 (2) 321-335.

Ofcom, 2009. The Consumer Experience. London: U.K. Government.

Ofgem, 2009. Energy Supply Probe – Proposed Retail Market Remedies. London: U.K. Government.

Ofgem, 2010. Update on Probe Monitoring: Tariff Differentials and Consumer Switching. London: U.K. Government.

Ofgem, 2011. Simpler Energy Tariffs. London: U.K. Government.

Ofgem, 2012. The Retail Market Review – Updated Domestic Proposals. London: U.K. Government.

OFT, 2008. Personal Current Accounts in the UK. London: U.K. Government.

Oftel, 2000. Consumer Switching Behaviour in the Telecoms Market. London: U.K. Government.

Paulhus, Delroy L., 1984. Two-Component Models of Socially Desirable Responding. Journal of Personality and Social Psychology 46(3), 598-609.

Reis, Ricardo, 2006. Inattentive Consumers. Journal of Monetary Economics, 53(8), 1761-1800.

Shanks, David R. and Mark F. St John, 1994. Characteristics of Dissociable Human Learning Systems. Behavioral and Brain Sciences 17(3), 367-447

Sheena S. Iyengar and Emir Kamenica, 2010. Choice Proliferation, Simplicity Seeking, and Asset Allocation. Journal of Public Economics 94(7-8), 530-539.

Sims, Christopher A., 2003. Implications of Rational Inattention. Journal of Monetary Economics 50(3): 665-690.

Sitzia, Stefania and Daniel J. Zizzo. Does Product Complexity Matter for Competition in Experimental Retail Markets? Theory and Decision 70(1), 65-82.

Spiegler, Rani, 2006. Competition over Agents with Boundedly Rational Expectations. Theoretical Economics 4(1), 207–231.

Stöber, Joachim (2001). The Social Desirability Scale-17 (SDS17): Convergent Validity, Discriminant Validity, and Relationship with Age. European Journal of Psychological Assessment 17(3), 222-232.

Thaler, Richard H., Sunstein Cass R., 2003. Libertarian Paternalism. American Economic Review Papers and Proceedings 93(2), 175-179.

Thaler, Richard H., Sunstein Cass R., 2008. Nudge: Improving Decisions About Health, Wealth, and Happiness, New Haven and London: Yale University Press.

Waddams, C. (2012). The Likely Effects of Compelling Energy Firms to Give Customers 'The Lowest Tariffs'. Norwich: CCP Law and Policy Blog.

Wilson, Chris M. and Waddams Price, Catherine, 2010. Do Consumers Switch to the Best Supplier? Oxford Economic Papers 62(4), 647-668.

Woodford, Michael, 2012. Inattentive Valuation and Reference-Dependent Choice. Columbia University unpublished paper.

Xavier, Patrick and Dimitri Ypsilanti, 2008. Switching Cost and Consumer Behaviour: Implications for Telecommunications Regulation. Info 10(4), 13-29.

Zizzo, Daniel J., 2003. Verbal and Behavioral Learning in a Probability Compounding Task. Theory and Decision 54(4), 287-314.

Zizzo, Daniel J. 2010. Experimenter Demand Effects in Economic Experiments. Experimental Economics, 139(1), 75-98.

Zizzo, D. J., and Fleming, Piers. 2011. Can Experimental Measures of Sensitivity to Social Pressure Predict Public good Contribution? Economic Letters, 111(3), 239-242.

Figure 1. Map of Experimental Treatments
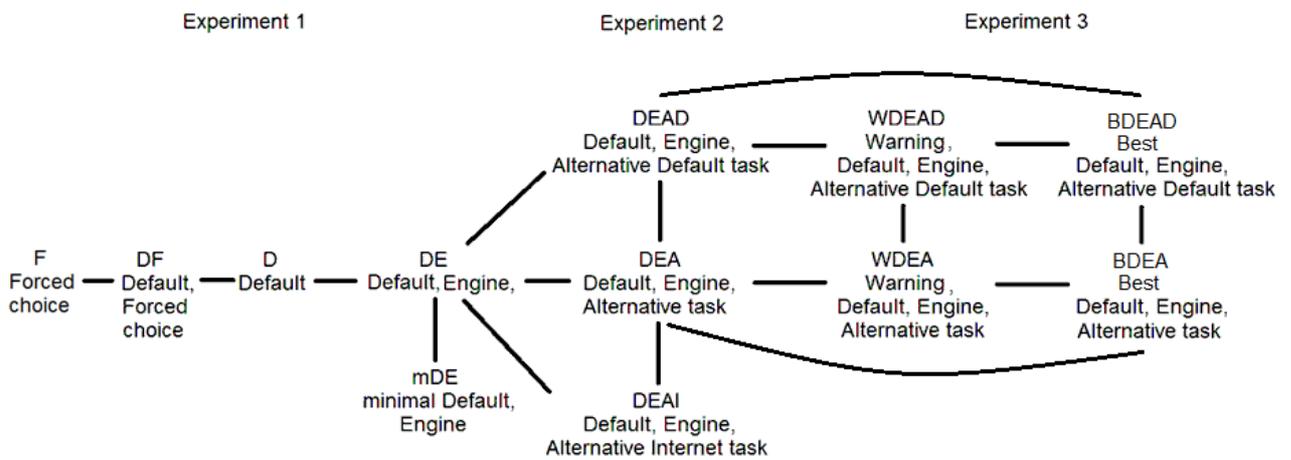
**Figure 2. Suboptimal Outcome Rates in Experiment 1**
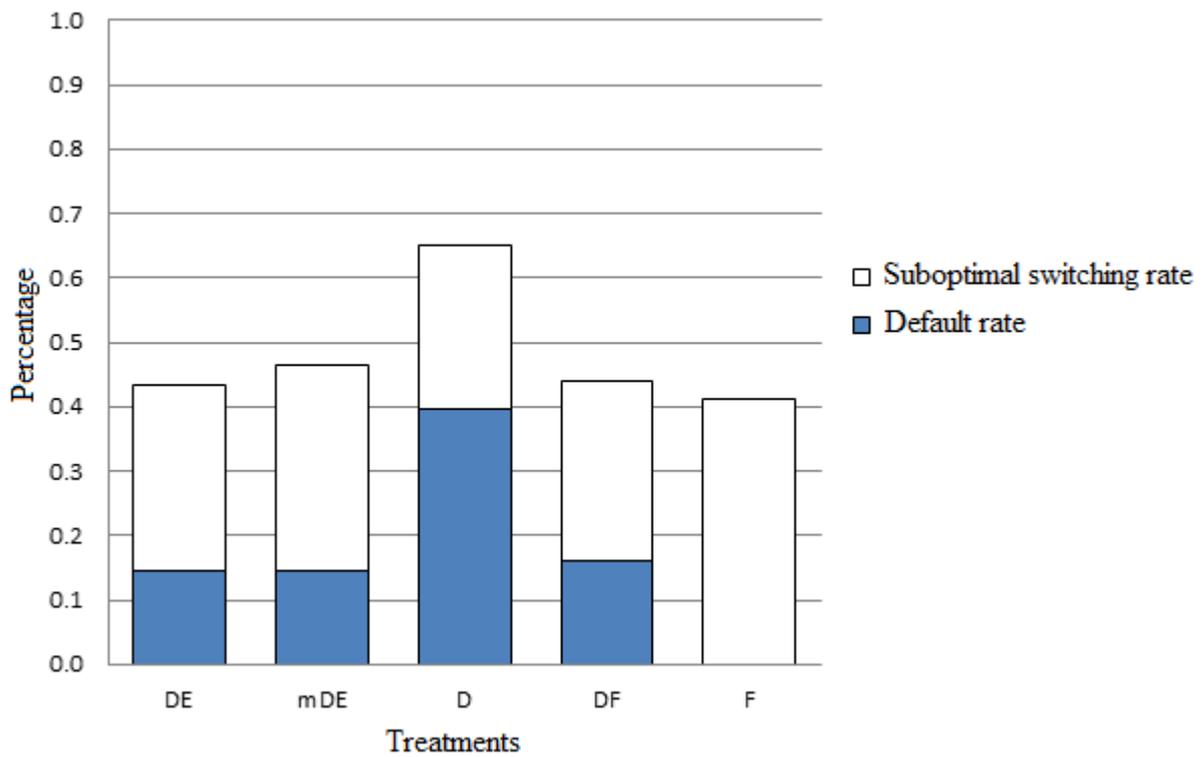
Figure 3. The Tariff Task in Treatment DE (screenshot)

Figure 4. Alternative Counting Task in DEA, BDEA and WDEA Treatments
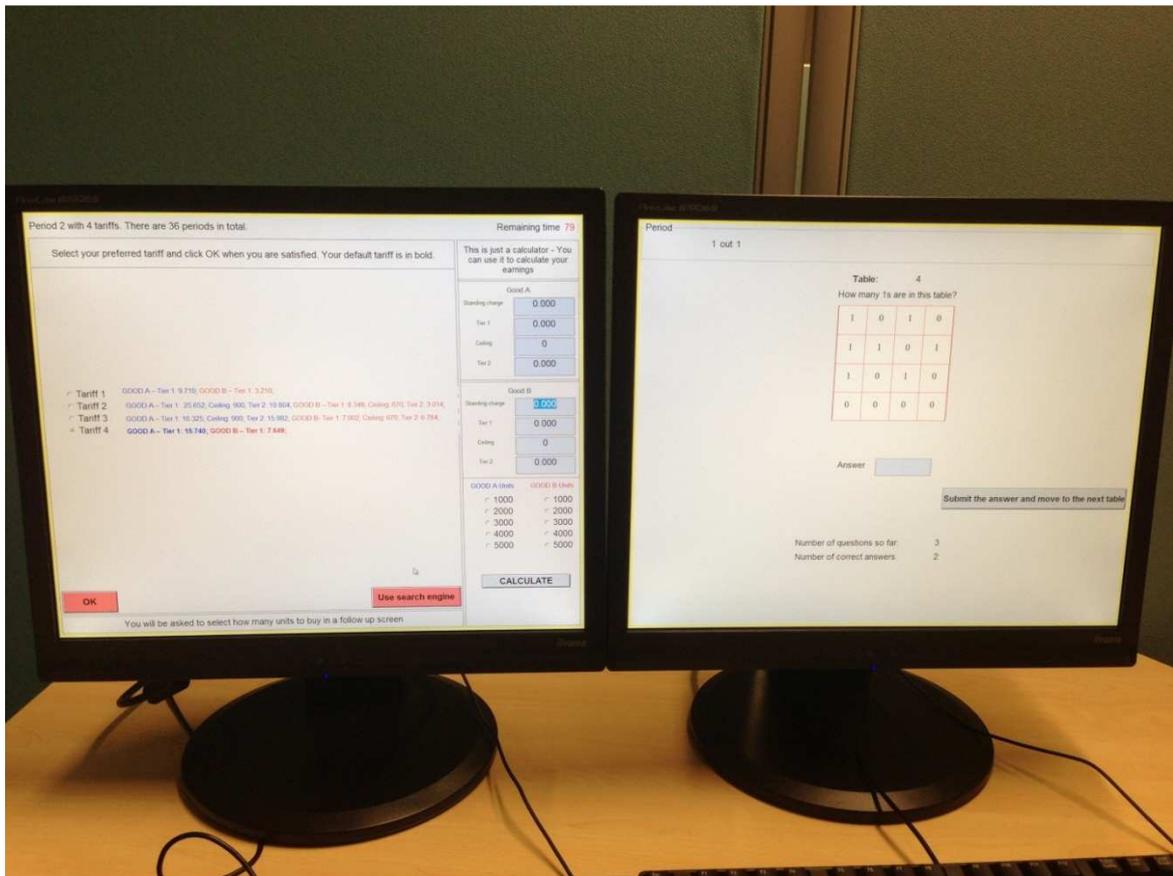
Figure 5. Counting Task Screen in DEAD, BDEAD and WDEAD Treatments

Figure 6. Suboptimal Outcome Rates in Experiment 2
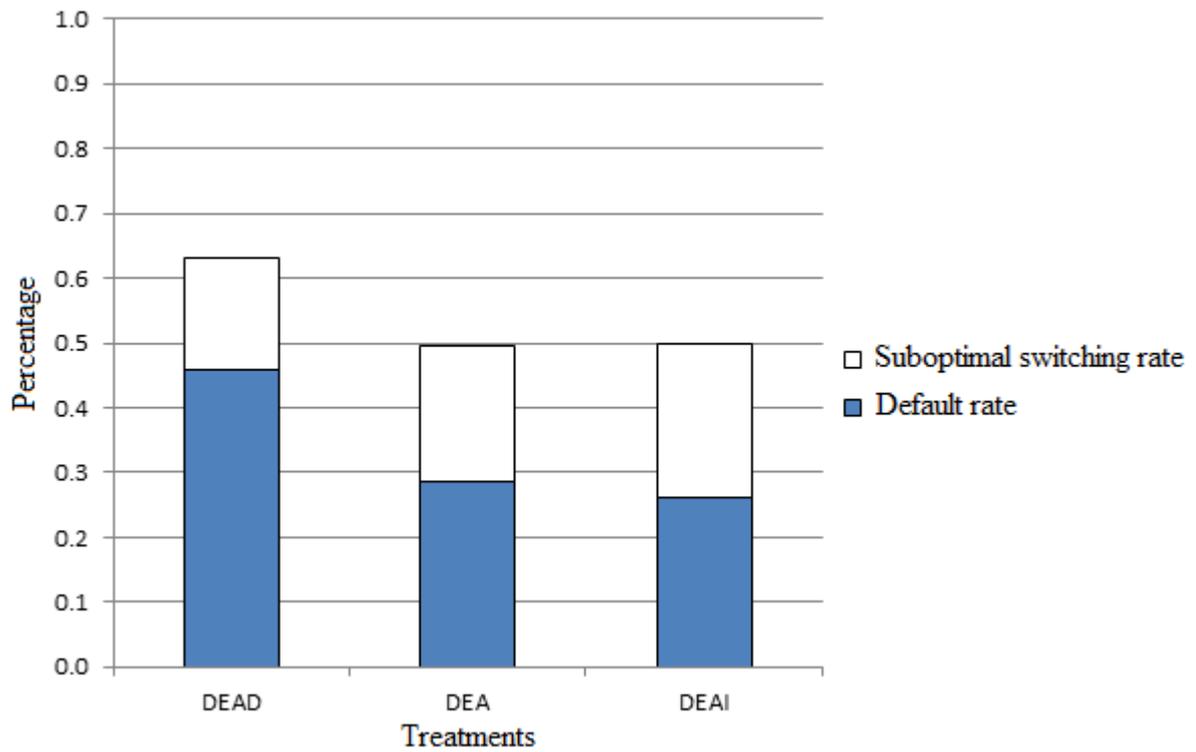
Figure 7. Suboptimal Outcome Rates in Experiment 3
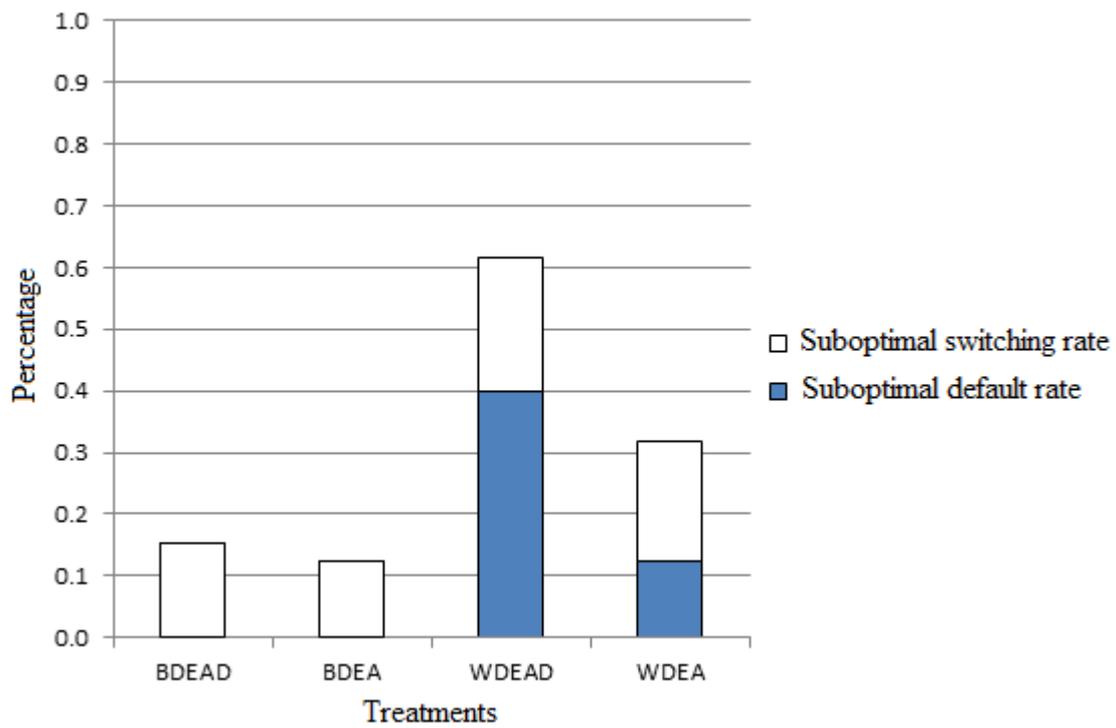
Table 1. Sample of Tariffs

Tariff	Good A				Good B			
	Standing charge	Tier 1	Ceiling	Tier 2	Standing charge	Tier 1	Ceiling	Tier 2
1	-	16.259	-	-	-	-	-	-
2	-	19.467	728	8.432	-	-	-	-
3	12.957	12.128	-	-	-	-	-	-
4	-	13.398	-	-	4.516			
5	-	19.992	900	12.19		5.878	670	5.521
6	13.23	12.338	-	-	9.569	7.269		

Table 2: Experiment 1 – Average Performance

Treatment	Type of Market		Complexity of Task			Number of Tariffs per Task		Overall
	Single Market	Dual Market	All Simple	All Complex	Mixed	4 tariffs	24 Tariffs	
Panel a - Default Rates								
DE	0.114	0.160	0.117	0.178	0.143	0.146	0.143	0.144
mDE	0.128	0.154	0.144	0.161	0.142	0.180	0.111	0.145
D	0.347	0.422	0.400	0.423	0.389	0.386	0.408	0.397
DF	0.155	0.166	0.157	0.207	0.153	0.161	0.163	0.162
F	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Average	0.202	0.242	0.223	0.260	0.223	0.232	0.226	0.229
Panel b - Suboptimal Switching Rates								
DE	0.275	0.297	0.167	0.339	0.308	0.185	0.394	0.290
mDE	0.281	0.338	0.217	0.306	0.347	0.204	0.433	0.319
D	0.237	0.263	0.153	0.337	0.259	0.151	0.358	0.254
DF	0.253	0.289	0.153	0.360	0.288	0.158	0.397	0.277
F	0.368	0.432	0.253	0.517	0.423	0.274	0.547	0.411
Average	0.283	0.324	0.189	0.372	0.325	0.194	0.426	0.310
Panel c - Suboptimal Outcome Rates								
DE	0.389	0.457	0.283	0.517	0.451	0.331	0.537	0.434
mDE	0.408	0.492	0.361	0.467	0.489	0.383	0.544	0.464
D	0.583	0.685	0.553	0.760	0.648	0.537	0.766	0.651
DF	0.408	0.455	0.310	0.567	0.440	0.319	0.560	0.439
F	0.368	0.432	0.253	0.517	0.423	0.274	0.547	0.411
Average	0.438	0.510	0.412	0.632	0.548	0.426	0.652	0.539

Panel d - First half and Second half						
	Default Choice Rates		Suboptimal Switching Rates		Suboptimal Choice Rates	
	First half	Second Half	First half	Second Half	First half	Second Half
DE	0.204	0.085	0.306	0.274	0.509	0.359
mDE	0.180	0.111	0.300	0.337	0.480	0.448
D	0.432	0.361	0.249	0.260	0.681	0.621
DF	0.193	0.131	0.293	0.261	0.487	0.392
F	0.036	0.028	0.398	0.360	0.433	0.388
Average	0.212	0.152	0.310	0.297	0.522	0.449

Table 3: Experiment 2 – Average Performance

Treatment	Type of Market		Complexity of Task			Number of Tariffs per task		Overall
	Single Market	Dual Market	All Simple	All Complex	Mixed	4 tariffs	24 Tariffs	
Panel a - Default Rates								
DEAD	0.462	0.456	0.447	0.473	0.457	0.438	0.478	0.458
DEA	0.262	0.298	0.267	0.313	0.284	0.274	0.298	0.286
DEAI	0.211	0.286	0.261	0.272	0.258	0.274	0.248	0.261
Average	0.327	0.356	0.335	0.365	0.345	0.337	0.356	0.346
Panel b - Suboptimal Switching Rates								
DEAD	0.167	0.176	0.127	0.193	0.179	0.100	0.246	0.173
DEA	0.215	0.207	0.140	0.237	0.220	0.118	0.301	0.209
DEAI	0.236	0.239	0.133	0.272	0.256	0.154	0.322	0.238
Average	0.201	0.202	0.133	0.228	0.213	0.119	0.285	0.202
Panel c - Suboptimal Outcome Rates								
DEAD	0.628	0.632	0.573	0.667	0.636	0.538	0.723	0.631
DEA	0.477	0.505	0.407	0.550	0.504	0.392	0.599	0.496
DEAI	0.447	0.525	0.394	0.544	0.514	0.428	0.570	0.499
Average	0.528	0.558	0.468	0.594	0.557	0.456	0.640	0.548

Panel d - First half and Second half

	Default Choice Rates		Suboptimal Switching Rates		Suboptimal Choice Rates	
	First half	Second Half	First half	Second Half	First half	Second Half
DEAD	0.533	0.382	0.170	0.176	0.703	0.558
DEA	0.340	0.232	0.226	0.193	0.566	0.426
DEAI	0.289	0.233	0.283	0.193	0.572	0.426
Average	0.403	0.290	0.218	0.186	0.620	0.476

Table 4: Experiment 3 – Average Performance

Treatment	Type of Market		Complexity of Task			Number of Tariffs per task		Overall
	Single Market	Dual Market	All Simple	All Complex	Mixed	4 tariffs	24 Tariffs	
Panel a - Default Rates								
BDEAD	0.850	0.850	0.878	0.828	0.849	0.878	0.822	0.850
BDEA	0.861	0.888	0.900	0.872	0.875	0.880	0.878	0.879
WDEAD	0.367	0.414	0.433	0.439	0.379	0.400	0.396	0.398
WDEA	0.100	0.136	0.100	0.156	0.122	0.102	0.146	0.124
Average	0.544	0.572	0.578	0.574	0.556	0.565	0.561	0.563
Panel b - Suboptimal Switching Rates								
BDEAD	0.150	0.150	0.122	0.172	0.151	0.122	0.178	0.150
BDEA	0.139	0.113	0.100	0.128	0.125	0.120	0.122	0.121
WDEAD	0.214	0.218	0.144	0.233	0.231	0.143	0.291	0.217
WDEA	0.172	0.207	0.150	0.256	0.192	0.106	0.285	0.195
Average	0.169	0.172	0.129	0.197	0.175	0.123	0.219	0.171
Panel c - Suboptimal Outcome Rates								
BDEAD	0.150	0.150	0.122	0.172	0.151	0.122	0.178	0.150
BDEA	0.139	0.113	0.100	0.128	0.125	0.120	0.122	0.121
WDEAD	0.581	0.632	0.578	0.672	0.610	0.543	0.687	0.615
WDEA	0.272	0.343	0.250	0.411	0.314	0.207	0.431	0.319
Average	0.285	0.309	0.263	0.346	0.300	0.248	0.355	0.301

Panel d - First half and Second half

	Default Choice Rates		Suboptimal Switching Rates		Suboptimal Choice Rates	
	First half	Second Half	First half	Second Half	First half	Second Half
BDEAD	0.833	0.867	0.167	0.133	0.167	0.133
BDEA	0.833	0.924	0.167	0.076	0.167	0.076
WDEAD	0.448	0.348	0.209	0.224	0.657	0.572
WDEA	0.180	0.069	0.226	0.165	0.406	0.233
Average	0.574	0.552	0.192	0.150	0.349	0.254