AER Energy Price Fact Sheet Study

PRE-ANALYSIS PLAN

23 November 2017

1. Introduction

1.1 Abstract and motivation

The Behavioural Economics Team of the Australian Government (BETA) will prepare advice for the Australian Energy Regulator (AER). The advice concerns energy price fact sheets (EPFSs) and responds to the AER's Consumer Price Information Issues Paper. BETA recommendations will follow from this mixed method study, constituted of a framed field experiment of the current EPFS design and five behaviourally informed EPFS designs as well as qualitative focus groups assessing EPFS content and design.

1.2 Experimental groups

- **Status quo** – current EPFS sample design provided by AER

- **Bar graph** – behaviourally informed EPFS design incorporating a household consumption bar graph to aid respondents in making comparisons

- **Base rate + houses** – behaviourally informed EPFS design incorporating the graphical representation of a house as well as the base annual bill rate to aid respondents in making comparisons

- **Base rate + household items** – behaviourally informed EPFS design incorporating the graphical representation of household items as well as the base annual bill rate to aid respondents in making comparisons

- **Base rate / discount rate + houses** – behaviourally informed EPFS design incorporating the graphical representation of a house as well as the base annual and discounted annual bill rates to aid respondents in making comparisons

- **Base rate / discount rate + household items** – behaviourally informed EPFS design incorporating the graphical representation of household items as well as the base annual and discounted annual bill rates to aid respondents in making comparisons

1.3 Research questions

1. Of available options, what is the preferred design and content for a one-page EPFS?

2. Among behaviourally informed EPFS designs, will study participants prefer to examine both base annual bill rate as well as discounted annual bill rate or just base annual bill rate alone?

3. Among behaviourally informed EPFS designs, will study participants prefer the graphical representations of houses, household items, or a bar graph?

1.4 Limitations

There are several limitations associated with this mixed methods design. First, interpreting a primary outcome (engagement with and preference for an EPFS design) will be complex and nuanced because there is no single primary outcome measure: 'engagement' is difficult to measure in this experimental context. Second, some framed field questions will be purely hypothetical (e.g. "how likely is it you might change energy plans if you observed a fact sheet like this one that indicated you would be better off"). Thus, from the perspective of revealed preferences it will be difficult to tell how much of an effect the treatment fact sheets would have on 'actual' energy literacy or energy-related decision-making. Third, a very small subset of framed field questions is differentiated by experimental group. This is because we wanted to ask one bespoke survey question about each EPFS design.

2. Methods

2.1 Framed field experiment

Our study is interested in testing EPFS designs across a representative sample of Australia that is 21 years or older. The sample is drawn from the 300,000-person Online Research Unit survey panel. ORU attempts to recruit a nationally representative sample of the Australian population through online and offline (telephone and post) recruitment methods. The survey mandates age, sex, and postcode location quotas to generate a nationally representative sample. Survey respondents receive a financial incentive at a physical Australian post to ensure they are Australian residents. For this survey, respondents receive approximately AUD 1.50. To address the possibility of missing data, respondents are replaced if they do not complete the survey.

2.1.1 Power

We calculate that a sample size of approximately 750 per experimental group provides 80% power at a 5% significance level with a margin of error of 3.65% in survey responses. This produces a total sample of 4,500 respondents assessing the six EPFS designs.

2.1.2 Randomisation

Respondents will first be screened on the basis of age, sex, and postcode location to help ensure a nationally representative sample for each experimental group. Subsequently, the survey panel randomly allocates respondents to each of the six experimental groups. Allocation will take place by 1) selecting the least-filled experimental group, and 2) using a random sequence instrument to allocate among experimental groups of equivalent least-filled sample sizes.

2.2 Qualitative focus groups

Our study is also interested in eliciting qualitative information regarding the comparison of all six EPFS designs as well as the individual and shared features of those designs. We will analyse the results of three focus groups. One focus group will be populated with members of COTA Australia (a peak non-profit organisation that represents the interests of older Australians) that will be based in Canberra, one focus group of individuals recruited from a survey panel in Adelaide, and one focus group of individuals recruited from a survey panel in Sydney. Each focus group discussion will include five to seven individuals and last approximately two hours. Moreover, after reaching a consensus on the most preferred EPFS design, each focus group will be asked about the inclusion of additional pricing information. This information will not be tested in the framed field experiment.

2.3 Data collection and processing

Given that this is a mixed methods study, data will be collected from the ORU survey panel as well as the three focus groups. Survey questions will be binary or categorical in nature, and only one question (E.2) will offer a free response field. All survey data processing and analysis steps will be performed using STATA script and will involve manual checks at each stage to ensure there are no errors introduced. We will not analyse any data until after the survey reaches the 4,500-person target.

For each focus group, the facilitator will generate high-level discussion notes. Notably, each group will receive the same set of semi-structured questions intended to guide discussion. Focus group facilitators will collect each group's set of preferences for the EPFS design comparisons and articulate areas of agreement and disagreement among participants.

We expect there to be limited or no missing framed field experimental data because respondents who return incomplete surveys are replaced. On the basis of screening and subsequent random sequence generation for allocating experimental groups we expect balanced samples for sex, age, and postcode location. It is possible, however, some samples will skew across some other demographic (such as household income or housing tenure) and thus we will conduct other balance checks.

3. Analysis and synthesis of results

3.1 Analysis of results

In the case of the framed field experiment, we will use hypothesis testing (Section 3.1.5) to elicit evidence of statistically significant differences between any of the six experimental groups for each primary outcome question. If we do detect a statistically significant result, we will use *post-hoc* tests to determine the relative preferential ranking of each EPFS design. If there are no statistically significant results arising from hypothesis testing, we will conclude a null result and consider descriptive statistics and cross-tabulations alone in forming a recommendation. Changes to the model form will be made on the basis of purely statistical considerations, seeking consistency of the data with model assumptions. In choosing an appropriate model form we will not allow consideration of effects on research outcomes.

In the case of the qualitative focus groups, thematic analysis will be applied to focus group data. Such data will be used to inform the interpretation of the primary and secondary outcomes (Section 3.1.1) of the framed field experiments. This information will therefore be treated as exploratory rather than confirmatory.

3.1.1 Primary outcome measures

Our primary outcome is engagement with, and demonstrated favourability toward, each EPFS design. We are interested in the preferential ranking of each EPFS within the set of behaviourally informed EPFSs.

One primary outcome measure is 'proportion of respondents favouring each EPFS design'. This is calculated by computing the mean of the five responses to each of the 5-item Likert scale questions for Survey Question E2 across the six experimental groups. The five responses to Question E2, "How strongly do you agree or disagree with the following statements about the fact sheet?" include, "It is easy to understand", "It is interesting", "It provides useful information", "If I had two of these side by side for different plans, it would be easy to see which plan was best for me", and "It would help me make decisions that affect my household budget".

A second primary outcome is 'likelihood of respondents to change energy plans on the basis of information presented by EPFSs for two or more energy plans'. This is calculated by computing the mean response across the six experimental groups on an 11-point sliding scale question (E4), "Imagine you have a fact sheet like this in front of you for your current electricity plan AND a different fact sheet for a different plan. If it looked like you would be better off switching to the different plan, how likely is it you would switch?"

A third primary outcome is 'change in respondent confidence after exposure to an EPFS design'. This is calculated by computing the sample mean difference in response between Question S9 and Question E5. For each question, we will calculate the composite mean for the three 11-point sliding scale statements responding to the question, "After reading the fact sheet, how confident do you NOW feel in the following?" Statements include, "Your ability to make choices about electricity plans, such as which plan or company to choose", "That there is enough easily understood information available to you online or through other channels to make decisions about electricity plans", and "That electricity companies will offer you the best plan for your needs". The third primary outcome measure will be difference between the composite sample means of Questions S9 and E5.

3.1.2 Secondary outcome measures

Our secondary outcomes relate to the extent to which subgroups reveal their preferences toward status quo and behaviourally informed EPFSs.

Secondary outcome measures relating to a respondent's agency in making household energy decisions will be collected (Question S1). In addition, outcome measures relaying a respondent's history of switching energy plans or companies will be collected (S7 and S8). Moreover, outcome measures relaying demographic information will be collected for each of the following: household income (D1), level of education attained (D2), marital status (D3), household size (D5 and D6), number of children in household (D6), property ownership status of accommodation (D7), and housing tenure in most recent residence (D8). In addition, we will collect outcome measures relating to disability status (D4); receipt of government pension, allowance or benefit (D4); and the status of non-English language spoken in a household (D4). Respondents will also be asked to indicate how they feel about their current financial situation using a 3-item list (D9).

We will also collect secondary outcome measures incorporating qualitative information from the online survey. Respondents will be asked to provide a free response to the statement, "Please tell us which part of the fact sheet you think provides the most useful information" (E1).

By virtue of the differences in three EPFS types, the online survey will elicit slightly different information regarding an energy plan's billing information, key features, and key facts (E3). For the behaviourally informed EPFS designs (experimental groups 1-4; Question E3), we will compute the averages for each of three 5-item Likert scale questions asking respondents about the usefulness of 1) estimated yearly bill, 2) key features, and 3) key facts. For the behaviourally informed EPFS design incorporating a bar graph detailing daily usage amounts (experimental group 5; Question E3A), we will compute the averages for each of three 5-item Likert scale questions asking respondents about the usefulness of 1) expected daily usage, 2) key features, and 3) key facts section. For the status quo EPFS design (experimental group 6; Question E3B), we will compute the average for a single 5-item Likert scale responses to the question, "How useful would it be if the fact sheet included an estimate of the average annual bill or average daily usage for a household on this plan?"

3.1.3 Balance checks

Covariates collected and balanced pre-randomisation include age, sex, and postcode location to generate a nationally representative sample. To ensure further balance, we will perform Pearson chi-squared tests across secondary outcome measures to judge whether observed covariate imbalances are larger than would be expected from chance alone. This involves regression of an intervention condition on independent variables collected post-randomisation, including income, education level, property ownership, etc. A p-value of 0.01 or less will prompt a review of the random assignment procedure and possible data-handling mistakes. If the review finds no errors, we will report the imbalance test and proceed on the assumption that the imbalance is due to chance, and report estimates with and without covariate adjustment.

3.1.4 Treatment Effect

Our primary analysis will entail a regression analysis of experimental group on the response variable for each primary outcome question (E2, E4 and E5). We will fit a linear regression model to the data with a primary outcome measure as the response variable and with the experimental condition as well as three covariates (sex, age, and household income) as explanatory variables for each primary outcome question. This initial specification is an initial model form that will be adapted as required consistent with model assumptions. For example, because Questions E2 and E5 uses a Likert scale to measure responses to a set of similar statements, for each of these questions we will calculate the simple pooled sample mean and variance for each question's respective set of statements. In the case of Question E5, we will take the additional step of calculating the sample mean differences between Questions E5 and S9 (which measure any self-reported change in a respondent's confidence) across the three statements that both nearly identical questions test.

Please note that while Question E4 represents an ordinal variable, we will treat it as a continuous variable in part because of its ordinal range (11 points between 0 and 10) and given the robustness of OLS model even for ordinal variables (Judkins & Porter 2015).

3.1.5 Hypothesis tests

- We will compare all experimental groups to each other. Therefore, we will conduct an initial one-way Analysis of Variance (ANOVA) on the set of all six experimental groups to see if there is evidence of any statistically significant differences arising from this set on the response variable (Montgomery 2013) for each of Questions E2, E4 and E5. As mentioned above, ANOVA will test pooled sample means and variances for E2 and E5 and the individual mean and variance for E4.

- If the test yields evidence of an effect for Questions E2, E4 and/or E5 then we will conduct *post hoc* tests to compare the treatment effect of each group to determine the relative preferential ranking of each EPFS design (Montgomery 2013). In view of our research questions, these tests will include one linear regression in which experimental groups will be represented as dummy variables and at least two direct comparisons: EPFS designs testing 1) base rate versus base rate and discount rate, and 2) graphical representation of houses versus that of household items.

- Pre-specified subgroup analyses (Section 3.1.7) will be conducted at a minimum for the status quo group versus the two pooled superior treatment groups and each superior treatment group against each other.

3.1.6 Covariates

Covariates included in the above model are age, sex, and income of each respondent. Age and sex were balanced before randomisation to an experimental group, whereas income was collected after randomisation. We do not expect income to be altered by experimental group.

3.1.7 Subgroups

A number of secondary subgroup analyses using linear regressions (comparing treatment effects and interaction effects) will also be performed and will be considered exploratory. Previous AER research (Briarbird 2017) has indicated that demographic groups that may be most responsive to EPFSs include large, busy working families; individuals with low incomes; and recent immigrants to Australia. Therefore subgroup analysis will be particularly concerned with income; household size; number of children; and if a household speaks at least one language other than English. Moreover, subgroup analysis will be conducted at a minimum on the two most preferred EPFS designs as well as the status quo EPFS design.

3.2 Synthesis of results

After data collection and processing of findings for both the framed field experiment and the focus groups, each set of results will be synthesised. Focus-group results will be used to inform the interpretation and discussion of framed field results.

4. Reporting

4.1 Deviation from this pre-analysis plan.

If our final report contains analyses that deviate from this plan we will make it clear that these analyses not pre-specified and provide justification for them. Conversely, if we omit pre-specified analyses we will make these available as supplementary material. In either case, deviation from the pre-analysis plan will be driven solely by statistical considerations and will not be influenced by any consideration of differences in findings on the research outcomes that would occur under different model forms or methods of analysis. If findings on the research outcomes are sensitive to different plausible model forms or methods of analysis then we will report this fact in our analysis.

4.2 Outcome tables

Baseline characteristics and balance

		Status quo	Bar graph	Base rate + houses	Base rate + household items	Base rate / discount rate + houses	Base rate / discount rate + household items
Decision-maker (joint)*							
Sex (female)*							
Age*	20/under						
	21-24						
	25-34						
	35-44						
	45-54						
	55-64						
	65-74						
	75+						
	Prefer no say						
State*	Sydney						
	Other NSW						
	Melbourne						

	Other VIC	
	Brisbane	
	Other QLD	
	Perth	
	Other WA	
	Adelaide	
	Other SA	
	Hobart	
	Other TAS	
	Darwin	
	Other NT	
	ACT	
Citizen*	Australian	
	Permanent resident	
	NZ living as permanent resident	
History	Switch comp.	
	Switch plans	
	Did not switch comp.	
	Did not switch plans	
Income	>$20K	
	$20K - $40K	
	$40K - $60K	
	$60K - $80K	
	$80K-$100K	
	$100K-$120K	
	$120K-$150K	
	$150K +	
	Don't know	
	Prefer no say	

Education < Year 12		
	Year 12	
	Trade/TAFE	
	Diploma	
	Univ. Degree	
Marital Single, NM		
	Married	
	De-facto	
	Widowed	
	Divorced	
	Separated	
	Prefer no say	
Language (non-English)		
Disability (household)		
Government benefit		
House # > age 18		
	< age 18	
Property Own		
	Rent	
	Share	
	Other	
Tenure <=1 year		
	2-3 years	
	4-5 years	
	6-9 years	
	>=10 years	
Financial Comfortable		
	Struggle	
	Pressure	

* Screening outcome measure

Main outcome table 1 – detecting an effect among comparison groups

	Mean scores for each experimental group (SD)							
	Status quo	Bar graph	Base rate + houses	Base rate + household items	Base rate / discount rate + houses	Base rate / discount rate + household items	F (df)	Significance (p-value)
E2* (continuous on 1-5)								
E4 (ordinal on 0-11)								
E5* (continuous on 1-5)								

*Note that the mean scores for Questions E2 and E5 are pooled across Likert-scales for five and three statements, respectively.

Main outcome table 2 – comparisons of treatment effect for each experimental group

	Coefficient (S.E.)	95% Confidence Interval	Significance (p-value)
Bar graph			
Base rate + houses			
Base rate + household items			
Base rate / discount rate + houses			
Base rate / discount rate + household items			

5. References

Briarbird. 2017. "Enhanced website feasibility report: Energy Made Easy website" on behalf of Australian Energy Regulator.

Judkins, D.R. and Porter, K.E. "Robustness of ordinary least squares in randomized clinical trials". Statistics in Medicine. Wiley.

Montgomery, D.C. 2013. *Design and Analysis of Experiments: International Student Version*. 8th ed. Wiley.