Australian Government
Department of the Prime Minister and Cabinet

BETA

Safety meets savings
Technical Report

This Technical Report includes mathematical notation that may not be fully accessible for those using assistive technology. Please email BETA@pmc.gov.au if you require assistance engaging with this material.

Pre-registration, pre-analysis plan and ethics

The Randomised Controlled Trial (RCT) component of this project was publicly pre-registered on the American Economic Association's Social Science Registry (AEARCTR-0014411) and on the BETA website. Both registrations were completed after a pilot study (described below). We launched the full study before finalising the preregistration, but pre-registered before the full sample was collected.

The ethical aspects of the research were reviewed and approved by Macquarie University Low Risk Committee (520241809158945).

The analyses of the RCT data were consistent with the pre-analysis plan. All exploratory analyses are clearly designated. The pre-analysis plan is available on the BETA website.

Pilot study

To test our design, we ran a pilot study with about 200 participants, who each responded to 6 unique choice sets, totalling about 1200 observations. The purpose of the pilot was to test that the Discrete Choice Experiment (DCE) was giving us expected results, and that participants were correctly interpreting our questions. We also used the pilot to estimate variability in the outcome measure for the RCT to adjust the power calculations.

Results of the pilot study included the following.

- We made no changes to the DCE.
- No changes were made to the survey questions.
- The pilot study revealed a much larger standard deviation (SD) in the outcome variable than initially assumed (SD = $28,599.28), which significantly altered our power calculations and approach (see below for detailed information).
- The pilot data was not included in the RCT analysis.

Population and sample selection

The population of interest was adults in Australia who own at least one residential property built before 1990. We monitored a number of characteristics to ensure the sample included good coverage of the characteristics of interest, including: owner-occupiers, landlords, those with mortgages and those who own outright, those who live in regional/remote areas and those who live in urban centres, those from low and those from high income brackets. We also monitored age, gender, state or territory of residence, and culturally and linguistically diverse (CALD) status. For this project a participant was classified as CALD if they either mainly spoke a language other than English at home or were born in a non-English speaking country.

The target sample size for the overall study was 4,500. The sample was recruited by Octopus Surveys, from their panel of participants. The study was conducted online using Qualtrics (the survey software), and our final sample size was 4,403. Full sample details are available in Appendix A of the main report.

Screening and replacement participants

We started the survey by asking respondents a few questions to check their eligibility (see Consort diagram, Figure 1). To participate in the study, respondents had to own one or more residential properties that were built prior to 1990 (and be over 18 years old and live in Australia). Buildings constructed pre-1990 are more likely to contain asbestos, and this cohort was therefore most relevant to our project partner, the Asbestos and Silica Safety and Eradication Agency (ASSEA).

After data collection, we checked the quality of the data. Any *duplicate* nonsense open-ended responses were be excluded from the data set (as they are likely to be bots) and replaced by the recruitment provider. There were no duplicates under these criteria. We had no other automatic exclusions.

We conducted two robustness checks. Our main analysis used the data set minus *duplicate* nonsense responses, as described above. Our first robustness check was with the *full* sample (that is, including duplicate nonsense responses, even though they are likely to be bots).

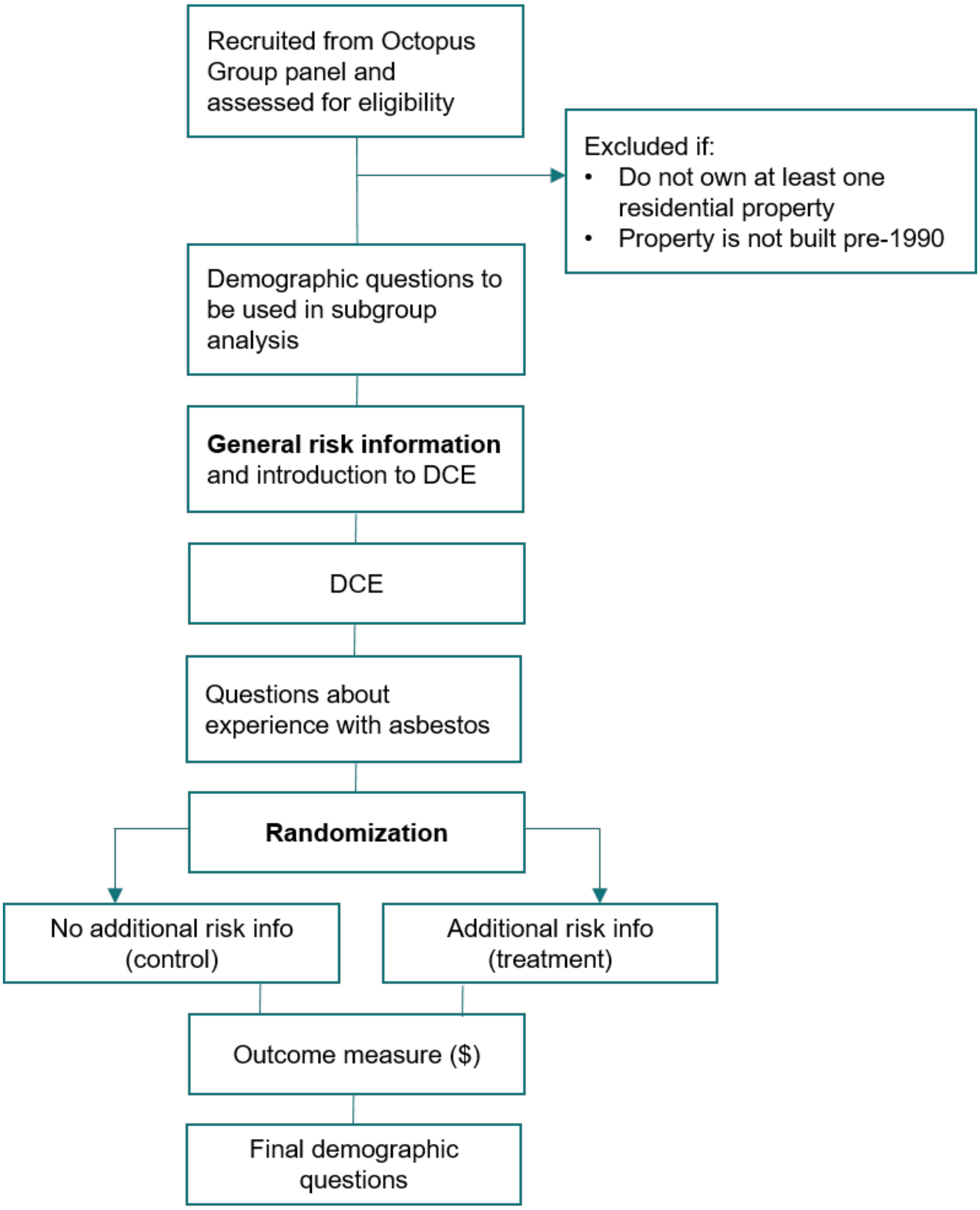
For the second robustness check, we examined responses that receive a score less than one on Qualtrics's bot score, or who provide nonsense/suspicious open-ended responses, or who are very fast (less than 2 minutes). These responses were examined in detail and excluded if there were multiple indicators that they were non-serious/nonsense responses. This sample – with *possible* bots removed – was our second robustness check. Of the 99 participants removed from this second robustness check, 97 did not provide responses to any questions and two were duplicates.

Randomised controlled trial

Trial design

Figure 1 below illustrates the trial design, including sequencing of initial risk information, the DCE, and the additional risk information (as part of the RCT).

Figure 1. CONSORT diagram illustrating trial design



Interventions and randomisation

This trial was an individually randomised online experiment. Participants were randomised to 1 of 2 arms (no additional risk information vs additional risk information). Randomisation was done by Qualtrics, by giving each participant a 1/2 probability of being assigned to each trial arm.

Outcome measure

The primary outcome measure was the price (in dollars) participants said they could afford to pay if they had to remove asbestos. At the individual level this was asked as a single question, and participants entered a dollar value in response. At a group level this was the mean dollar value in each group.

- If a participant skipped this question, their response were coded as 'missing' (n=54).
- If they entered $0, this counted as a legitimate response.

Hypothesis

People who do not receive additional risk information will indicate a different maximum price (in dollars) than those who receive additional risk information (B ≠ A two-tailed test).

Power calculations

Due to resource constraints, our sample was fixed at around 4,500 individuals. For this study, alpha was set to 0.05, and hypothesis tests were two-sided.

ASSEA indicated the smallest effect size of interest would be around $500.

Initially, we powered our study assuming a SD of $2000, based on expected variability in the outcome measure. This allowed us to calculate sample sizes with 90% power, targeting a $350 difference between groups (Cohen's d = 0.175). Given these assumptions, our original sample size of 4000 participants was sufficient, and we planned to conduct sensitivity analyses with smaller subsamples (such as 2000 participants) to examine whether the effect remained practically significant and robust under more typical, not overpowered study conditions.

However, our pilot study revealed a much larger SD in the outcome variable than initially assumed (SD = $28,599), which significantly altered our power calculations. With this variability incorporated, we calculated the design had 80% power to detect an effect size of $1740 (Cohen's d = 0.061).

Method of analysis

The principal analysis of the effect of the intervention consisted of a covariate-adjusted comparison of our primary outcome. This estimate, confidence intervals and p-values were derived from a linear regression model using robust (HC2) standard errors and with the following specification:

$$Y_i = \beta_0 + \beta_1 Z_i + \beta_2 X_i + \beta_3 Z_i X_i + \epsilon_i$$

Where *i* was an index for each individual in the trial, *Y* was the individual's dollar response, β_0 was the intercept, *Z* was treatment assignment indicator, β_1 was the coefficient representing the average treatment effect for the intervention relative to control, *X* was a mean-centred covariate (see Covariates section below), *ZX* was the interaction of the treatment indicator with the mean-centred covariate indicator, and ε was the individual error term.

Covariates

We included participants' household income in the model as a baseline dummy-coded, mean-centred covariate. It was recoded into low, medium and high, with low as the reference group. The question asked for household income in 8 brackets: <$30,000; 30,000 - 50,000; 50,001 - 70,000; 70,001 - 100,000; 100,001 - 130,000; 130,001 - 150,000; 150,001 - 200,000; over 200,001. We recoded incomes below 70,001 as low, incomes between 70,001 and 130,000 as medium, and those 130,001 and above as high.

Robustness checks

We checked for robustness by winsorising the outcome variable to reduce the influence of extreme outliers. This approach involved capping the most extreme values (99th percentile), allowing us to assess whether the main analysis results remained consistent when the variability caused by outliers was minimised.

Missing data approach

Since participants could skip the outcome question, there was the possibility of differential attrition. To address this, we took a multi-stage approach. At the end of data collection, we examined the rate of missingness in our primary outcome variables. Only 54 respondents (<1%) had missing data on the outcome variable. As it was less than the threshold of 10% indicated in our pre-analysis plan, we conducted complete case analysis.

Tables

Table 1. RCT primary analysis

Condition	Marginal Means ($)	Estimate ($)	Standard error ($)	95% CI ($)	p-value
Control	16,566.40	-	-	-	-
Treatment	22,595.30	6,028.95	5,234.00	-4,232.54; 16,290.44	0.25

n = 4105. OLS model adjusted for income category with HC2 robust standard errors.

Table 2. RCT winsorised analysis

Condition	Marginal Means ($)	Estimate ($)	Standard error ($)	95% CI ($)	p-value
Control	14,703.10	-	-	-	-
Treatment	14,696.60	-6.52	576.51	-1,136.80; 1,123.75	0.99

n = 4105. OLS model adjusted for income category with HC2 robust standard errors using winsorised outcome data.

Discrete choice experiment

Interventions and randomisation

The DCE component of this trial was randomised at the level of attributes and options. We specified the following:

- each participant responded to seven choice sets (the second and last were the same to calculate intra-respondent reliability, see further details below)
- each choice set contained two options for removal and an ‘I would not remove the asbestos’ option
- there were five attributes for each removal option:
    - loan
    - grant
    - tax offset
    - lottery
    - quoted cost
- the levels for each attribute are specified in Table 1.

We calculated a D-efficient experimental design based on these specifications, using Qualtrics’ “Conjoint Analysis” feature. We implemented a conditional constraint on the grant and quoted cost attributes to ensure the grant amount never exceeded the quoted cost. While this may have affected perfect design balance, it maintained scenario realism and enhanced response validity. While the remaining combinations varied in how ‘realistic’ they were, we didn’t anticipate that this would create problems for the analyses.

Table 3. Attributes and levels used in the present study

Attribute	Number of levels	Value of levels
Quoted cost	4	$5,000; $10,000; $20,000; $30,000
Grant	4	None; $5,000; $10,000; $15,000
Tax offset	2	Not available; available
Interest-free loan	4	None; 5 year; 10 year; 15 year
Lottery	2	Not available; available

Outcome measures

The primary outcome measure was the option each individual chose, in each of six (plus one repeated) choice sets. Each choice set contained two options with varying incentives and costs for removing asbestos. Each decision was coded as ‘1’ if option was chosen, and as ‘0’ if it was not chosen. If a participant chose ‘I would not remove’, both options in that choice set were coded as 0. Using this outcome, we calculated the influence each attribute had on participants’ choices.

Hypotheses

We did not have specific confirmatory hypotheses for the DCE.

Method of analysis

We had a range of analysis options specified in the pre-analysis plan.

The primary model was a 'full' mixed logit model first including all levels of all incentives and random slopes except on cost. The model converged successfully, therefore no alternative specifications were required.

The probability of choosing an option (0/1) was modelled using a mixed logit model with the following variables:

- a continuous variable for quoted cost (rescaled by dividing by $30,000) – fixed effect
- a continuous variable for grant (rescaled by dividing by $15,000) – random effect
- a binary indicator for tax offset (0 = no tax offset, 1 = tax offset) – random effect
- an indicator for loan, dummy-coded (0 = no loan, three levels of loan) – random effect
- a binary indicator for lottery (0 = no lottery, 1 = lottery) – random effect.

The random effects were coded as uncorrelated. To aid the mixed effect model in convergence we normalised the cost and grant variables by their respective maximum values – cost was divided by $30,000 and grant by $15,000, scaling each range between 0 and their maximum potential value. We did not centre the variables as zero represented a meaningful value in our context.

We also ran another model where we recoded all incentives to binary and used random slopes except on cost. This was to calculate an 'average' assessment of the impact of each incentive vs no incentive as well. This was to allow us to assess how much individual participants varied in their sensitivity to the presence of the incentives.

We note that while our pre-analysis plan initially specified normalising both cost and grant variables by the maximum cost value ($30,000), we modified this approach to instead normalise each variable by its own maximum value. This change allowed for clearer interpretation as each variable was scaled relative to its full range in the DCE, while still maintaining the benefits of normalisation for model convergence.

Thus we set:

$$X_i{}^{norm} = \frac{X_i}{P_i^{\max}}$$

Where *X* is the value for either the cost or grant attribute and *Pmax* is the maximum value of the respective attribute:

- $30,000 for the cost attribute
- $15,000 for the grant attribute.

We then back-converted the normalised and centred value for reporting.

Intra-respondent reliability correction

We also calculated and corrected for the intra-respondent reliability (IRR) to control for potential measurement error. The IRR was calculated by comparing participants' responses to choice set 2 and a repeated choice set 7, which were identical except for the reversal of the left and right positions of the alternatives. This approach enabled us to measure the consistency of participants' choices, which was 80% (following the approach outlined by Clayton et al. 2023).

The IRR was calculated as the proportion of consistent responses between the original and repeated choice sets. The IRR was then used to adjust the observed Marginal Means (MM) and Average Marginal Component Effects (AMCEs). We used AMCEs to estimate the causal effect of changes in each attribute level on the probability of choosing an option. AMCEs provide directly interpretable estimates of how changing one feature affects choice probability, holding all other features constant. We calculated both unadjusted AMCEs and AMCEs adjusted for response reliability to provide more conservative estimates that account for measurement error in respondents' choices. Since our IRR was 80%, indicating good response consistency, the adjusted AMCEs are larger than the unadjusted ones. This adjustment provides estimates of what the true effects might be after accounting for response noise in the DCE. We did this adjustment by applying the following correction formula:

$$MM_{corrected} = \frac{MM_{observed\,-\tau}}{1 - 2\tau}$$

where τ is the swapping error rate derived from the IRR. The swapping error rate refers to the probability that a participant's response to a choice set does not reflect their true preference but is instead randomly "swapped" or chosen incorrectly.

$$AMCE_{corrected} = \frac{AMCE_{observed\,-\tau}}{1 - 2\tau}$$

Subgroup analyses

For the subgroup analysis, we used a simpler model specification with fixed effects for all attributes and only a random intercept for participant, rather than the full random slopes specification used in the main analysis. This approach was chosen due to the smaller sample sizes in subgroups, while still accounting for the panel structure of the data through the random intercept. We fit the DCE model to subsets of the data where the subset was larger than *n* = 500 (after excluding 'would not remove' responses). These subsets were:

- People who are renting out their property
- People who are not renting out their property but may still own multiple properties
- People whose property is mortgaged
- People whose property is owned outright
- People in different age brackets
- People in different income brackets (lower, medium, higher)
- People who are planning to remove asbestos and those who are not planning to remove asbestos

For each subgroup, we calculated IRR separately to account for potential measurement error across these groups and also used AMCEs as per the main analysis.

Tables

Table 4. DCE main analysis

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-4.31	0.06	-70.99	-0.19
Grant (per $1000)	4.33	0.08	52.98	0.19
5-year loan	0.31	0.03	9.91	0.41
10-year loan	0.37	0.03	10.99	0.49
15-year loan	0.34	0.03	9.86	0.44
Tax offset	0.45	0.02	19.51	0.59
Lottery	0.05	0.02	2.31	0.07

n = 4,403. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 52,836 (6 choices per participant with each choice yielding 2 observations). We have removed the duplicated choice set. AMCE = Average Marginal Component Effect.

Table 5. DCE subgroup: Income category = low (<70,001)

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.80	0.10	-39.44	-4.94
Grant (per $1000)	3.68	0.15	23.80	4.78
5-year loan	0.25	0.07	3.83	0.33
10-year loan	0.26	0.07	3.92	0.33
15-year loan	0.23	0.07	3.59	0.31
Tax offset	0.29	0.05	6.20	0.37
Lottery	0.10	0.05	2.18	0.13

n = 888. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 10656 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 6. DCE subgroup: Income category = medium ($70,001-$130,000)

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.73	0.08	-49.05	-4.85
Grant (per $1000)	3.93	0.12	31.85	5.11
5-year loan	0.25	0.05	4.79	0.32
10-year loan	0.33	0.05	6.38	0.43
15-year loan	0.34	0.05	6.53	0.44
Tax offset	0.38	0.04	10.42	0.50
Lottery	0.04	0.04	1.12	0.05

n = 1300. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 15600 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 7. DCE subgroup: Income category = high (>$130,001)

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.73	0.06	-62.46	-4.85
Grant (per $1000)	3.87	0.10	39.41	5.03
5-year loan	0.30	0.04	7.14	0.39
10-year loan	0.39	0.04	9.24	0.51
15-year loan	0.34	0.04	7.95	0.44
Tax offset	0.48	0.03	16.11	0.62
Lottery	0.03	0.03	1.09	0.04

n = 1960. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 23520 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 8. DCE subgroup: Age = 18-34 years

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.60	0.08	-46.60	-4.67
Grant (per $1000)	3.71	0.13	29.02	4.81
5-year loan	0.24	0.06	4.39	0.31
10-year loan	0.38	0.06	6.80	0.49
15-year loan	0.28	0.06	4.99	0.36
Tax offset	0.44	0.04	11.26	0.57
Lottery	0.09	0.04	2.26	0.11

n = 1124. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 13488 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 9. DCE subgroup: Age = 35-49 years

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.78	0.07	-57.83	-4.91
Grant (per $1000)	3.94	0.11	36.77	5.11
5-year loan	0.38	0.05	8.39	0.50
10-year loan	0.41	0.05	9.01	0.54
15-year loan	0.43	0.05	9.28	0.55
Tax offset	0.41	0.03	12.77	0.54
Lottery	0.03	0.03	0.96	0.04

n = 1671. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 20052 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 10. DCE subgroup: Age = 50-64 years

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.84	0.09	-42.98	-4.98
Grant (per $1000)	3.90	0.14	27.23	5.06
5-year loan	0.17	0.06	2.75	0.22
10-year loan	0.27	0.06	4.54	0.36
15-year loan	0.19	0.06	3.21	0.25
Tax offset	0.47	0.04	10.96	0.61
Lottery	0.00	0.04	0.02	0.00

n = 1018. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 12216 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 11. DCE subgroup: Age = 65+ years

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-4.16	0.13	-32.08	-5.40
Grant (per $1000)	4.17	0.21	20.21	5.41
5-year loan	0.12	0.08	1.46	0.16
10-year loan	0.08	0.08	0.98	0.11
15-year loan	0.13	0.08	1.59	0.17
Tax offset	0.25	0.06	4.21	0.33
Lottery	0.06	0.06	1.03	0.08

n = 556. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 6672 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 12. DCE subgroup: Home owned outright

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.51	0.08	-46.68	-4.56
Grant (per $1000)	3.40	0.12	27.87	4.42
5-year loan	0.13	0.05	2.41	0.16
10-year loan	0.15	0.05	2.95	0.20
15-year loan	0.14	0.05	2.62	0.18
Tax offset	0.33	0.04	9.03	0.43
Lottery	0.04	0.04	1.12	0.05

n = 1320. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 15840 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 13. DCE subgroup: Home owned, paying off mortgage

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.83	0.05	-78.89	-4.98
Grant (per $1000)	4.04	0.08	50.72	5.24
5-year loan	0.32	0.03	9.44	0.42
10-year loan	0.39	0.03	11.59	0.51
15-year loan	0.36	0.03	10.65	0.47
Tax offset	0.44	0.02	18.36	0.57
Lottery	0.04	0.02	1.88	0.06

n = 3073. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 36876 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 14. DCE subgroup: Not a landlord (may still own multiple properties)

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.71	0.05	-80.05	-4.82
Grant (per $1000)	3.82	0.08	50.88	4.96
5-year loan	0.28	0.03	8.92	0.37
10-year loan	0.37	0.03	11.48	0.48
15-year loan	0.33	0.03	10.21	0.42
Tax offset	0.37	0.02	16.35	0.48
Lottery	0.03	0.02	1.16	0.03

n = 3498. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 41976 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 15. DCE subgroup: Landlord renting out their property

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.81	0.10	-36.37	-4.95
Grant (per $1000)	3.82	0.17	22.55	4.96
5-year loan	0.20	0.07	2.82	0.26
10-year loan	0.19	0.07	2.67	0.25
15-year loan	0.21	0.07	2.87	0.27
Tax offset	0.61	0.05	11.83	0.79
Lottery	0.10	0.05	2.05	0.14

n = 687. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 8244 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 16. DCE subgroup: No asbestos removal plans

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-4.53	0.14	-33.31	-5.88
Grant (per $1000)	4.72	0.21	22.25	6.13
5-year loan	0.34	0.09	3.98	0.44
10-year loan	0.32	0.09	3.71	0.41
15-year loan	0.31	0.09	3.63	0.40
Tax offset	0.41	0.06	6.86	0.53
Lottery	0.00	0.06	0.02	0.00

n = 577. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 6924 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

Table 17. DCE subgroup: Asbestos removal plans

Condition	Estimate (percentage points)	Standard error (percentage points)	z-value	AMCE
Cost (per $1000)	-3.24	0.10	-31.17	-4.21
Grant (per $1000)	4.05	0.18	23.00	5.26
5-year loan	0.31	0.08	4.13	0.41
10-year loan	0.44	0.08	5.82	0.58
15-year loan	0.49	0.08	6.48	0.64
Tax offset	0.45	0.05	8.47	0.59
Lottery	0.04	0.05	0.78	0.05

n = 583. Logistic mixed model corrected for intra-respondent reliability. Number of observations = 6996 (6 choices per participant with each choice yielding 2 observations). AMCE = Average Marginal Component Effect.

References

Clayton, K, Horiuchi, Y, Kaufman, AR, King, G and Komisarchik, M (2023) Correcting measurement error bias in conjoint survey experiments. Harvard University Working Paper.