Long-term Energy Cost Labelling for Appliances: Evidence from a Randomised Controlled Trial in Ireland

E. Denny

Received: 25 May 2021 / Accepted: 25 March 2022 / Published online: 10 May 2022
© The Author(s) 2022

Abstract
Given the longevity of investments in energy-consuming products (such as household appliances, vehicles, and properties), underinvestment in energy efficiency can have long-lasting negative economic and environmental consequences. Previous research has indicated that underinvestment may be due to imperfect information in relation to the long-term benefits of investing in energy efficiency. This paper presents the results of a cluster randomised controlled trial examining an intervention which aims to overcome this information deficit by providing long-term energy cost information on appliances in an electrical retail chain in Ireland. Two treatments are considered: a label showing 10-year energy cost information based on typical usage for four appliance categories (fridge freezers, dishwashers, washing machines, and tumble dryers); and a second treatment which supplements this label with a QR code where consumers can gain personalised cost estimates based on their expected appliance usage. Results indicate that neither of the treatments resulted in an increase in the average energy efficiency of appliances sold. Also, engagement of customers with the QR code was extremely low. Given that the newly designed EU energy labels incorporate QR codes for personalisation, this low usage suggests that this element of the new labels may be ineffective in increasing the uptake of energy efficiency. Finally, a customer survey suggests that while the treatment increased the stated importance of energy efficiency in decision-making, this did not translate into an increase in efficiency of products purchased, i.e., stated preferences for energy efficiency did not translate into revealed purchasing preferences.

Keywords Energy efficiency · Energy labels · Household appliances · Randomised control trial · Personalised information

It is well recognised that consumers often make sub-optimal choices when making investment decisions. Prior research has shown private investment “mistakes” in purchasing health insurance (Abaluck & Gruber, 2011), financial investments (Barber et al., 2005), and...
food choices (Bollinger et al., 2011) amongst others. Allcott and Knittel (2019) argue that these systematic mistakes could be the result of imperfect information regarding costs and benefits and/or customers failing to pay attention to certain attributes. In particular, it has been shown that individuals tend to underestimate values for prices, taxes, and quantities when either information is transmitted opaquely or due to customer inattention (Wichman, 2017).

Given the longevity of investments in energy-consuming products, such as household appliances, vehicles, and properties, “mistakes” in energy-related purchases can have long-lasting effects with both negative economic and environmental consequences. Investments in energy efficiency (EE), which appear to be highly beneficial from both a private and social perspective, are often not made. This is the so-called energy efficiency gap or energy efficiency paradox (Jaffe & Stavins, 1994). Some authors suggest that almost 40% of the potential energy savings worldwide are not achieved due to underinvestment in greater energy efficiency (IEA, 2007). One of the primary reasons given for this underinvestment is informational failures which include asymmetry and imperfect information.

The literature on the energy efficiency gap raises the following question: Would providing energy cost information in monetary terms at the point of sale address this information deficit and encourage consumers to purchase more energy-efficient appliances? As the energy efficiency gap posits, if customers do indeed have imperfect information in relation to energy consumption and energy costs, then an informational intervention addressing this deficit should encourage a greater uptake of energy-efficient appliances.

This paper presents the results of a cluster randomised controlled trial (RCT) providing long-term energy cost information on appliances in an electrical retail chain in Ireland. Two treatments were devised in the experiment: The first displayed a label showing 10-year energy cost information in euros on four appliance categories (fridge freezers, dishwashers, washing machines, and tumble dryers). The second supplemented this label with a QR code where consumers could use their smartphone to gain personalised energy consumption information based on their expected usage of the appliance. Monthly sales data across the four appliance categories were collected over an eight-month period to determine if the interventions resulted in an increase in the sales of energy-efficient appliances in the treatment stores. The RCT was also supplemented by an in-store customer survey to gather individual-level data.

There are three key findings from this paper; the first is that neither of the treatments resulted in an increase in the average energy efficiency of appliances sold, i.e., displaying the 10-year running cost information (and supplementing this with a QR code for personalised information) did not encourage a greater uptake of energy-efficient appliances. This is consistent with previous revealed preference research indicating that cost labels alone on large white good appliances are not sufficient to nudge energy-efficient purchases. The second key finding is that the engagement of customers with the QR code to personalise the energy cost information was extremely low, with just 20 customers in total who used the QR code over the period of the trial. This represents less than one customer per month in each of the treatment stores. Given that the newly designed EU energy labels, due to be rolled out in 2021, incorporate QR codes for personalisation, this low uptake suggests that this element of the new labels may be ineffective in increasing the uptake of energy efficiency. Thirdly, the customer survey suggests that while the treatments increased the stated importance of energy efficiency in decision-making, this did not translate into an increase in efficiency of products purchased for the sampled customers, a result which holds for all demographic groups sampled.
The main contribution of this paper is to provide experimental evidence on the impact of long-term average and personalised energy cost information in monetary terms on energy efficiency investment for appliances. The null results found in this paper are consistent with previous experimental research on revealed behaviour and strengthen the field by overcoming some limitations of previous work, by testing usage of a novel QR code, including consumer-level data, testing a label with a longer time horizon and including multiple appliances.

Prior research using discrete choice experiments explores monetary labelling for appliances (Andor et al., 2020; Davis & Metcalf, 2016; Heinzle, 2012; Min et al., 2014; Sammer & Wüstenhagen, 2006), vehicles (Brazil et al., 2019; Codagnone et al., 2016), and properties (Carroll et al., 2016a). The potential for monetary labels has also been examined in the field, for example, using annual fuel economy labels for vehicles (Allcott & Knittel, 2019) and annual energy bill estimates in property sales and rentals (CONSEED, 2019).

Of particular relevance to this paper are field trials examining the impact of monetary labels for appliances on actual purchasing decisions. Allcott and Taubinsky (2015) and Andor et al. (2019b) analyse the impact of energy cost information on lightbulb sales in the USA and Germany respectively with the former finding no impact and the latter finding a positive impact. For larger appliances, Kallbekken et al. (2013) conducted a field experiment testing the effectiveness of a monetary label and sales-person training on fridge freezers and tumble dryers sales in Norway; del Mar Solà et al. (2021) conducted a similar field trial amongst small retailers in Spain with fridge freezers, dishwashers, and washing machines; Carroll et al. (2016b) test the effectiveness of a new five-year energy cost label on tumble dryers sales in Ireland and DECC (2014) examined a new energy cost label for washing machines and tumble dryers in the UK. Interestingly, the vast majority of these field trials show null or non-significant results, i.e., they demonstrate that displaying energy cost information for appliances alone did not increase uptake of more efficient products. While Kallbekken et al. (2013) did find positive effects, these were only when labels were combined with sales-person training. The energy cost information alone was not sufficient to increase energy efficiency, a result supported by Allcott and Sweeney (2016). del Mar Solà et al. (2021) also find that sales-person training is important for effectiveness of monetary labels on fridge freezers (although was not significant for washing machines) and no effect was found for dishwashers.

A recent paper by Stadelmann and Schubert (2018) compares the effectiveness of the EU energy label (in kWh) and a new monetary label in an online randomised controlled trial (the labels were shown in isolation) on tumble dryers, freezers, and vacuum cleaners. They find no statistical difference in the effectiveness of the two labels for tumble dryers and freezers (and for vacuum cleaners the existing EU label was more effective in promoting energy efficiency than the monetary label).

Thus, results of studies so far have been mixed which could be due to different methodological approaches, different appliances, and different contexts. Also, given the complexities associated with running experiments in a live retail environment, many of these revealed preference papers have limitations, including non-randomised design, results for just one appliance type and/or no customer-level data, and a lack of baseline data. These are limitations which this paper aims to address; thus, the null findings in this paper strengthen the null findings of the previous studies and provide stronger evidence that energy cost labelling alone may not be sufficient to drive increases in consumer investment in high-efficiency household appliances. In addition, this paper indicates that a QR code approach to providing personalised information will not be widely used and is unlikely to encourage greater energy efficiency uptake.
The remainder of the paper is structured as follows: The “Framework and Research Questions” section provides the framework for the research and outlines the research questions; the “Methodology” section describes the methodology and the “Data Summary” section presents a summary of the data collected; the empirical analysis strategy is presented in the “Empirical Strategy” section and the results and limitations in the “Results and Discussion” section, with the conclusions in the “Conclusions” section.

Framework and Research Questions

This paper is grounded in the energy efficiency paradox literature which hypothesises that the diffusion of energy efficiency technologies is slower than would be socially optimal (Jaffe & Stavins, 1994). Informational failures, market failures, and behavioural failures are amongst the explanations presented to explain this gap (del Mar Solà et al., 2021; Frederiks et al., 2015; Gerarden et al., 2015; Jaffe et al., 2004; Linares & Labandeira, 2010). This paper is primarily concerned with the first of these aspects, informational failures (Allcott & Sweeney, 2016; Labandeira et al., 2012; Phillips, 2012), and more specifically, imperfect information (or biased expectations) regarding energy savings (the difference between the energy consumption of an efficient and inefficient good). This imperfect information may be driven by lack of knowledge on energy prices and/or unfamiliarity with the unit of electricity (the kilowatt-hour) (Brounen et al., 2013; Sovacool & Blyth, 2015).

The theoretical framework adopted in this paper is adapted from Allcott and Greenstone (2012) and del Mar Solà et al. (2017) and describes an individual’s choice between an efficient good with energy intensity $e_1$ and an inefficient good $e_0$ (where $e_1 < e_0$). The decision is spread across two time periods, with investment in period 1 and consumption in period 2. It considers that the individual will choose the more efficient good if

$$\omega \cdot p \cdot m_1 \left( e_0 - e_1 \right) \left( 1 + r \right) - c_i + b_i > I$$

where $p$ is the price per unit of energy, in this context, price per kWh of energy consumption, and $r$ is the risk-adjusted discount rate between period 1 and 2. The incremental costs and benefits of adopting the more efficient good are represented by $c$ and $b$ respectively, and the additional investment cost for the more efficient product is denoted by $I$. To capture the energy efficiency gap or “investment inefficiencies” (Allcott & Greenstone, 2012), energy savings are scaled by a parameter ($0 < \omega < 1$), which is equivalent to the implied discount rate. This parameter $\omega$ and the costs and benefits are denoted with a subscript $i$ to allow for heterogeneity amongst individuals (Houde & Myers, 2021b).

As is clear from Eq. 1, the choice to invest in the more efficient product is a function or more than just a comparison of incremental energy savings and investment costs. Other costs (captured by $c_i$) could include unobserved transaction or adoption costs, and benefits ($b_i$) could include warm glow, comfort improvements, and status effects amongst others and are likely to be highly agent specific and influenced by demographic factors, environmental attitude, and culture (see Houde & Wekhof, 2021, for a discussion on market barriers to energy efficiency investment which indicate the importance of co-benefits such as comfort and ecological concerns).

Behavioural failures also play an important role and may widen the “gap” through, for example, hyperbolic discounting, loss aversion, and status-quo bias. Market failures could contribute to aspects such as energy prices being lower than marginal cost, for example,
due to environmental externalities not being considered; capital market failures in respect of accessing capital at the risk-adjusted discount rate; and imperfect information leading to the under-valuation of energy savings due to lacking information on energy intensity of different products and/or lacking information on energy prices (for a full literature review on the factors contributing to the energy efficiency gap in the residential sector, please see del Mar Solà et al. 2020).

The focus of this study is around this last market failure and overcoming this information gap by displaying estimates of long-term energy cost estimates at the point of sale. While the standard EU labels provide energy consumption information in kWh per annum, it is hypothesised that households are missing vital information to translate this kWh information into energy cost estimates for each appliance. The theoretical underpinnings for providing cost estimates, which are eloquently summarised in Wichman (2017), stem from extensive literature that indicates that obtaining the required and relevant information to make perfectly informed decisions is costly, that consumers may be inattentive to or unaware of (changes in) prices or taxes, that inattention could depend on attributes that are hidden from consumers, that consumers may use heuristics for decision-making when price and quantity information is unclear or uncertain, and/or that consumers may have biased assessments of prices, expenditure, and consumption.

In this study, this missing information is provided directly through a new energy cost label which displays estimates of energy costs for a 10-year period. This long-term horizon is chosen based on findings in Heinzle (2012) and Carroll et al. (2021) who demonstrate in a stated preference study that long-term energy cost forecasts increase the willingness to pay for energy efficiency, whereas short-term forecasts have a null or negative effect. As shown in Houde (2014), there is seen to be heterogeneity in how households engage with labels and energy information with some attentive to energy costs, others to rating (in the case of that paper, the EnergyStar certification), and others to both. In the context of this paper and the proposed 10-year monetary label, for households who are inattentive to energy consumption and energy prices, the new information in the label may introduce energy costs into their consideration set at the point of purchase. For attentive households with biased energy cost expectations, it is expected that demand for higher energy efficiency will increase only if the costs displayed through the labels are higher than their prior expectations. For both sets of households, it is anticipated that this would translate into higher demand for energy efficiency. This is tested in research question 1 (RQ1):

RQ1: Does providing long-term energy cost information (based on typical usage) increase the purchase of higher efficiency appliances?

The second research question relates to a household’s-specific quantity of energy services. As shown in a stated choice experiment by Davis and Metcalf (2016), when energy cost information is individually tailed, it has significant effects on stated choices between energy-using products. Research question 2 (RQ2) aims to determine if this effect is evident when examining revealed behaviour, i.e., are individuals more attentive to energy efficiency if they receive a personalised prediction of energy costs based on their usage patterns. The second research question therefore examines whether personalised energy cost information is more impactful than average energy cost information.

RQ2: Does providing long-term energy cost information (based on personalised household usage) increase the purchase of higher efficiency appliances?
It is likely that there is considerable household to household variation in the valuation of energy efficiency, and similarly in other aspects of the decision-making process, for example, discount rate, opportunity cost/utility (which may be a function of age/education, etc.), and warm glow effect. Thus, the third research question aims to explore the role of customer characteristics on the effectiveness of the long-term energy cost labels.

RQ3: Does effectiveness of the monetary label differ by customer characteristics?

Research questions 1 and 2 are addressed using a cluster randomised controlled trial (RCT) which tests the impact of a new long-term energy cost label on appliance sales while research question 3 is examined using an in-store consumer survey.

Methodology

This paper presents the results of a cluster randomised controlled experiment combined with an in-store consumer survey. The methodology for both approaches is discussed in this section. The study received ethical approval from the ethics committee of the Faculty of Arts, Humanities, and Social Sciences at the author’s institution. A requirement of this ethics application was a detailed pre-analysis plan (PAP). This included an outline hypothesis and a detailed description of the methodology for the experimental design, the implementation, and the econometric specification. The latter was based on the approach followed by Carroll et al. (2016b). This ethics approval with pre-analysis plan was submitted on 9th November 2015. The methodology was later updated to include a comparison to del Mar Solà et al. (2021) as that paper had not been published at the time the PAP was written.

Cluster Randomised Controlled Trial

The field partner for the RCT is an electrical retailer, DID Electrical, with 22 physical showroom stores located around the island of Ireland plus a web store (see map of locations in Fig. 5 in Appendix Figs. 5 and 6). These stores stock a wide range of household appliances from electric kettles and televisions to larger fridge freezers and washing machines. The customer experience is typical for an electrical retailer where sample models are displayed on the showroom floor and customers collect the product or arrange delivery following payment. The stores are a mix of large and small-to-medium outlets and are located in both city centre and suburban geographical areas.

Power tests were completed ahead of the experimental design using the software Optimal Design v3.01 (Spybrook et al., 2011) setting “Cluster randomised trials with cluster level outcomes.” Given the relatively small number of stores in the retail partner’s chain, the power of this study is low with a minimal detectable effect size of 1.4 with 22 clusters ($\alpha = 0.05$ and $P = 0.8$). The low power in this RCT is a limitation and the results should therefore be considered in light of this low power. Further discussion on this limitation is provided in the “Limitations and Future Work” section.

In the experimental design, stratified randomisation is conducted at the store level to assign stores to treatment and control groups. Stores were first stratified by size (using sales data from the baseline period) and then by average income level in the store district (using data from the 2016 national census (CSO, 2016)). From each group, approximately
one-third of the stores were randomly assigned to the control group, one-third to treatment group 1 (10-year energy cost label), and one-third to treatment group 2 (10-year energy cost label with QR code for personalised information). The web store was omitted from the study due to concerns over contamination of the control group. Table 1 illustrates the store allocations to treatment and control.

Table 1: Randomised allocation of stores to within clusters

| Regional income level | Store size        | Store names                  |
|-----------------------|-------------------|------------------------------|
|                       | Low/medium        | Control store 1             |
|                       |                   | Control store 2             |
|                       |                   | Treatment 1 store 1         |
|                       |                   | Treatment 1 store 2         |
|                       |                   | Treatment 2 store 1         |
|                       |                   | Treatment 2 store 2         |
|                       | Large             | Control store 3             |
|                       |                   | Control store 4             |
|                       |                   | Treatment 1 store 3         |
|                       |                   | Treatment 1 store 4         |
|                       |                   | Treatment 2 store 3         |
|                       |                   | Treatment 2 store 4         |
|                       |                   | Control store 5             |
|                       |                   | Control store 6             |
|                       |                   | Treatment 1 store 5         |
|                       |                   | Treatment 2 store 5         |
|                       |                   | Treatment 1 store 7         |
|                       |                   | Treatment 2 store 6         |
|                       |                   | Treatment 2 store 7         |

Store names are suppressed to protect commercially sensitive information. Approximate geographic dispersion of stores is illustrated in Appendix Figs. 5 and 6.

Figure 1 illustrates the new labels that were introduced in the treatment stores. Labels were designed in collaboration with the retailer, a graphic designer, and through feedback from two customer focus groups. The only difference in the design of the label between the two treatment groups is the text describing the function of the QR code. For treatment 1, the instructions for the QR code stated “Scan here to compare models” and for treatment 2 stores, the instructions for the QR code were “This is an average running cost. To see how much this appliance would cost YOUR household, scan here.” Customers who scan the QR code on their smart phone in treatment 1 are directed to a digital version of the label where they can track and compare the models they have scanned based on average consumption information. In treatment 2 stores, customers who scan the QR code are directed to a webpage where they can personalise the running cost information based on the behaviour of their household (and also track and compare models). This webpage was designed specifically for this study and a screenshot of the landing page for treatment 2 is shown in Fig. 2.

The labels presented in Fig. 1 were displayed alongside the existing EU energy label (example in Figs. 5 and 6 in Appendix 1). The “estimated 10-year running cost” displayed on each label in the treatment stores was calculated using the following formula:

\[
\text{Estimated 10-year running cost} = (\frac{\text{kWh}}{a}) \times \frac{\text{€}}{\text{kWh}} \times 10 \text{ years}
\]  

(2)

where kWh/a is the estimated annual consumption for the appliance provided by the EU energy label, and €/kWh is the average Irish electricity price (which was €0.1763/kWh at the time of the study).

For treatment 2, customers could scan the QR code and, using the webpage shown in Fig. 2, adapt the estimated running cost based on their own typical usage. For example, for...
a washing machine, the personalised 10-year energy cost estimate would be calculated as follows:

\[
\text{Personalised 10- year running cost} = \left( \frac{\text{kWh}/a}{220 \text{ cycles}} \right) \times Cycles_i \times 52 \text{ weeks} \times 10 \text{ years}
\]  

(3)

where “220 cycles” is the average number of annual washing machine cycles as assumed in the official EU energy label for washing machines.\(^1\) Dividing the kWh/a consumption by “220 cycles” provides the estimated consumption per cycle for the appliance. This is multiplied by the inputted number of weekly cycles by the household on the webpage, \(Cycles_i\), and multiplied by 52 weeks in a year for 10 years.

For the other appliances, from the official EU energy label guide, the average number of annual cycles for dishwashers is 280 cycles per annum\(^2\) and for tumble dryers is 160 cycles per annum.\(^3\) The methodology for calculating personalised fridge freezer consumption is more involved and is conducted by converting the assumed average household size used in

\(^1\) https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32010R1061.

\(^2\) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32010R1059.

\(^3\) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32012R0392.
Long-term Energy Cost Labelling for Appliances: Evidence from…

the EU energy label (2.4 persons) to the tailored household size entered by the household in the landing page for fridge freezers in Fig. 2. Full details on how personalisation is conducted for fridge freezers are provided in Appendix 2.

The cost calculations do not include a discount rate and assume fixed electricity prices. The concept of discounting is unlikely to be understood by the general public and as such it is not included as it is likely to lower the clarity of the label and therefore buyer engagement. In addition, while it is likely that buyers will understand energy price inflation, it is excluded in the interest of transparency. However, as discussed in Carroll et al. (2016b), the exclusion of these two factors together is likely to be less biasing since one of has an upward effect on future energy costs (energy price rises) and the other has a downward effect (discounting). The labels displayed in the treatment stores (and the QR landing page) provide the details of the energy cost assumptions used.
Fig. 3 Timeline of data collection. Author's design
Figure 3 illustrates the timeline of the treatment. The author met with staff in the head office of DID Electrical who made initial introductions to store-level managers. DID Electrical provided the make and model of all appliances available for purchase and unique labels were created for each model based on its kWh/a energy consumption. The author hired 16 undergraduate research assistants (from the author’s institution) to travel to the treatment stores and install the unique labels on all washing machines, tumble dryers, dishwashers, and fridge freezers. This installation took place during May and June 2017. Research assistants then returned to the stores throughout the trial period (July–Nov 2017) to install labels on newly added appliances to each store’s line-up and to conduct consumer surveys (as described in the “Consumer Survey” section).

**Consumer Survey**

Given the low level of power in the RCT (which was anticipated in advance due to the pre-analysis plan), an additional avenue for analysis, with individual-level outcomes, was implemented. This took the form of surveys of customers in all stores to provide additional insights and individual-level outcomes. It should be noted that this second element, survey collection, does not constitute a randomised sample; rather, it presents a sample of customers who were willing to engage with surveyors in store and is therefore likely to be biased towards customers with time and interest. This limitation is discussed further in the “Limitations and Future Work” section.

The research assistants received brief training in advance of store visits on appropriate behaviour in-store. Research assistants approached customers shopping for any of the four appliances (washing machines, tumble dryers, fridge freezers, and dishwashers), in each of the control and treatment stores to request participation in the survey. Participants were approached on the open shop floor in the vicinity of the large white-goods section prior to the point of sale. Participation was on a voluntary basis with an incentive that participants would be entered into a raffle for a €500 voucher. Research assistants used a paper-based two-page survey (Appendix 3). The survey instrument was approved by the ethics committee at the author’s institution and included an information section and consent details in the opening statement. The survey was completely anonymous and participants registered for the raffle in a separate box by providing their details on a postcard (given to them by the research assistant). No email addresses or identifying data were collected within the survey.

Under the terms of the ethical approval, participation in the survey was voluntary and participants could skip questions if they wished, leading to some variation in response rates for each question.

**Data Summary**

This section provides a high-level summary of the sales data in the baseline period and summary statistics from the customer surveys. Table 2 provides a summary of the sales of appliances in the baseline period

In the analysis, and similar to the approach taken in del Mar Solà et al. (2021), appliances are grouped into two dichotomous categories: “high-efficiency” models and “all other models” (Eq. 5). Given that washing machines are only available in A+ and above ratings, only A+++ rated models are considered “high efficiency.” For the other appliances, a wider
range of models are available and thus “high-efficiency” models are considered to be A++ and A+++.

\[
\text{High Efficiency}_m = \begin{cases} 
1 & \text{if EU energy rating} = \text{A+++} \forall m \\
1 & \text{if EU energy rating} = \text{A+} \forall m \neq \text{WM} \\
0 & \text{if EU energy rating} = \text{A+} \text{ for } m = \text{WM} \\
0 & \text{if EU energy rating} < \text{A+} \forall m
\end{cases}
\]

where \( m \) is the product type (washing machine, tumble dryer, fridge freezer, or dishwasher), and WM is washing machine.

Table 3 illustrates a summary of the characteristics of each of the products in control and treatment stores in the baseline period. As can be seen, there are no significant differences in the share of high-efficiency products between control and treatment stores except in the case of fridge freezers where average energy consumption of appliances sold (kWh/\( a \)) is marginally lower in the treatment stores in the baseline period.

Table 4 summarises the data collected in the customer surveys. This is a self-selected pool of participants for the consumer survey and as such, extrapolating the findings to the wider population should be done with caution. It is however of interest to compare those customers in the control and treatment stores as there is no reason to believe that survey volunteers in control stores differ from those in the treatment stores.
This paper is concerned with the aspects which impact on appliance sales within a retail environment and specifically: (RQ1) Does providing long-term energy cost information (based on average usage) increase the purchase of higher efficiency appliances? and (RQ2) does providing long-term energy cost information (based on personalised usage) increase
Table 4  Summary of customer survey data

| Variable                  | Control stores | Treatment stores |
|---------------------------|----------------|------------------|
|                           | # responses    | % of sample      | # responses    | % of sample      |
| Total sample              | 387            | 635              |
| Age                       | 374            | 97%              | 610            | 96%              |
| 18-29                     | 10%            | 7%               |
| 30-39                     | 24%            | 23%              |
| 40-49                     | 27%            | 27%              |
| 50-59                     | 24%            | 21%              |
| 60-69                     | 13%            | 16%              |
| 70 and over               | 3%             | 7%               |
| Education (N)             | 372            | 610              |
| Primary school            | 2%             | 3%               |
| Secondary school          | 24%            | 25%              |
| Diploma/certificate       | 23%            | 19%              |
| Primary degree            | 35%            | 33%              |
| Master degree             | 16%            | 19%              |
| PhD degree                | 1%             | 1%               |
| Monthly income (N)        | 328            | 478              |
| €0–1000                   | 6%             | 7%               |
| €1001–2000                | 16%            | 10%              |
| €2001–3000                | 25%            | 23%              |
| €3001–4000                | 23%            | 22%              |
| €4001–5000                | 16%            | 19%              |
| €5001–6000                | 5%             | 10%              |
| €6001-7000                | 5%             | 4%               |
| More than €7000           | 6%             | 6%               |
| Male                      | 374            | 609              |
| Homeowner                 | 373            | 614              |
| Landlord                  | 373            | 614              |
| Made a purchase 1         | 387            | 635              |
| Adults in household (#)   | 371            | 604              |
| Mean                      | 2.3            | 604              |
| Std dev                   | 1.19           | 2.2             |
| Children in household (#) | 329            | 452              |
| Mean                      | 0.97           | 1.26             |
| Std dev                   | 1.15           | 1.32             |
| Patience2                 | 374            | 607              |
| Mean                      | 6.63           | 6.53             |
| Std dev                   | 2.21           | 2.25             |
| Risk3                     | 372            | 607              |
| Mean                      | 5.60           | 5.75             |
| Std dev                   | 2.09           | 2.37             |

1Customers shopping for any of the four appliances of interest were approached in store and asked to participate in the survey; however, not all shoppers proceeded to make a purchase. This variable represents the proportion who made a purchase.

2The survey included a question eliciting a subjective measure of each individual’s patience level; the question was phrased “Are you generally an impatient person, or someone who always shows great patience? Please rate yourself on a scale of 0 to 10 where 0 represents ‘very impatient’ and 10 is ‘very patient’.”

3The survey included a question eliciting a subjective measure of each individual’s risk preferences; the question was phrased “How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please indicate where you are on a scale of 0 to 10 where 0 represents ‘unwilling to take risks’ and the 10 represents ‘fully prepared to take risk’.”
the purchase of higher efficiency appliances? No data is available on the factors motivating a consumer’s decision to purchase these appliances; thus, the focus is on the total number of appliances purchased by customers in a given store in a given month.

The effects of treatment are estimated in two ways to facilitate comparison with the different methods utilised in two comparable papers, del Mar Solà et al. (2021) and Carroll et al. (2016b). They are first estimated using a probit model, which follows the approach in del Mar Solà et al. (2021), and considers the likelihood of a purchased appliance being a high-efficiency model (as defined in Eq. 4 previously) in the control versus the treatment stores.

\[ P(y = 1|X) = \beta_1 + \beta_2 T1 + \beta_3 T2 + \beta_4 Price + \beta_5 Price^2 + \sum_i \beta_i N_i \]  

(5)

where \( y \) is the energy efficiency level, with \( y = 1 \) if the product is considered high efficiency (as per Eq. 4) and zero otherwise. \( X \) represents the explanatory variables (treatment/control group and technical specifications of the appliances). \( T1 \) and \( T2 \) are dummy variables taking the value of 1 for stores in treatment group 1 and 2 respectively. \( Price \) is the price of the appliance and \( N \) represents technical characteristics of each model such as noise level, capacity, and delay functionality. Controls are also included for store and month of year.

The sales data contains a high number of zero observations, for example see Fig. 4 for washing machine sales for a month. An observation of zero indicates that no units of a given appliance model were purchased in a particular store in a particular month. Given the high level of zeros in the data, a negative binomial count regression model is employed, using the same methodology as Carroll et al. (2016b). The effects of the treatments are then estimated using a difference-in-difference approach.

Following the approach in Carroll et al. (2016b), an exposure variable (\( z \)) is included to control for differences in observational settings between stores, to capture, for example, differences in footfall across different stores. Monthly total sales in each store are used to capture this. The expected kWh/annum of model \( m \) in store \( j \) during month \( t \) is then:

\[ \mu_{mjt} = \exp \left[ \beta_3 + \beta_1 w_t + \beta_2 e_m + \beta_3 z_j + \beta_4 (w_t \ast e_m) + \beta_5 (w_t \ast z_j) + \beta_6 (e_m \ast z_j) + \beta_7 (w_t \ast e_m \ast z_j) + \sum_i \beta_i N_i r + s_{jt} \right] \]

(6)
where \( w \) is a dummy variable representing the trial period, \( z \) is a dummy for treatment store and there are \( N \) appliance characteristics, \( s \) is the store and \( t \) is the month. The main interest is the relationship between kWh of appliances sold in the treatment versus the control stores during the trial period. As such, the effect for the control group pre-trial is described by \( \beta_2 \) and \( \beta_4 \) indicates if there is a significant change in this effect during the trial period. The coefficient for the three-way interaction coefficient (\( \beta_7 \)) is the key coefficient of interest and indicates any additional change for the treatment group during the trial period (Carroll et al., 2016b).

It is anticipated that consumers prefer greater energy efficiency; thus, a negative relationship between energy consumption (kWh per annum) and sales (a negative and significant \( \beta_2 \)) is anticipated holding all other product characteristics constant. If the new labels in the treatment increase the value that purchasers place on energy efficiency, the magnitude of this relationship will increase (\( \beta_7 \) also negative and significant). The model is estimated for treatments 1 and 2 combined and also separately for each of treatment 1 and treatment 2.

Research question 3 (Does effectiveness of the monetary label differ by customer characteristics?) is assessed using the data collected in the in-store customer surveys. Given the ordered and discrete nature of question responses, an ordered logit is conducted to consider the impact of customer characteristics on stated energy efficiency importance.

\[
EE_i = \beta_1 + \beta_2 T_i + \sum_{r} \beta_r Z_{ir} \tag{7}
\]

where \( EE_i \) is the stated importance of energy efficiency for individual \( i \) on a scale of 0 to 10 (where 0 is “I didn’t consider it at all” and 10 is “it is the most important aspect of my decision”). \( T_i \) is a dummy variable which takes a value of 1 if the customer was “treated,” i.e., was in a treatment store and stated they had seen the new energy cost label. \( Z_i \) are demographic and socio-economic characteristics of individual \( i \).

The customer survey also gathered information on the actual appliance purchased by individual customers; thus, the impact of the label on individual purchases was estimated using a probit model where the dependent variable (\( y \)) takes a value of 1 if the purchased appliance was “high efficiency” (as defined in Eq. 4) and 0 otherwise:

\[
P(y = 1|X) = \beta_1 + \beta_2 T_i + \sum_{r} \beta_r Z_{ir} \tag{8}
\]

**Results and Discussion**

If the treatment(s) worked, then an increase in the purchase of more energy-efficient products in treatment stores is expected resulting in a reduction in average kWh/\( a \) in these stores. Table 5 indicates the average change in kWh/\( a \) in control and treatment stores in the baseline and trial periods. While most stores experience improvements in efficiency between baseline and trial period, this is apparent in both the control and treatment stores. There is no marked improvement in either treatment group when compared to control stores. This is the first indication that both treatments were unsuccessful in incentivising improved energy efficiency at a store level.

Table 5 illustrates averages for the treatment 1 and treatment 2 stores separately and also pooled (“pooled T1 and T2”). This is due to data from the webpage for treatment 2 which indicated that there were just 20 customers in total who used the QR code in treatment 2.
stores over the period of the trial. This represented less than one customer per month in each of the treatment 2 stores. Given this extremely low usage rate, and the fact that the physical labels in treatments 1 and 2 are identical except for the text instructions for the QR code (see Fig. 1), the subsequent analysis is estimated with both treatment groups pooled. Appendix Tables 15, 16 and 17 illustrates the models estimated with treatment split into treatments 1 and 2 and the results are consistent.

Table 5 Average efficiency of appliances sold in baseline and trial periods

| Average kWh/a of products sold | Baseline period average kWh/a | Trial period average kWh/a | Change in average kWh/a | % change in average kWh/a |
|-------------------------------|-------------------------------|----------------------------|------------------------|--------------------------|
| **Washing machine**           |                               |                            |                        |                          |
| Control stores                | 187.04 (30.84)                | 182.76 (29.98)             | −4.28                 | −2.23%                   |
| Treatment 1 stores (T1)       | 186.63 (30.73)                | 182.52 (29.11)             | −4.11                 | −2.20%                   |
| Treatment 2 stores (T2)       | 187.28 (30.01)                | 183.03 (28.62)             | −4.25                 | −2.27%                   |
| Pooled T1 and T2              | 186.99 (30.52)                | 182.79 (28.85)             | −4.20                 | −2.25%                   |
| **Tumble dryer**              |                               |                            |                        |                          |
| Control stores                | 469.77 (135.18)               | 469.45 (147.36)            | −0.32                 | −0.07%                   |
| Treatment 1 stores (T1)       | 462.91 (136.84)               | 465.23 (145.38)            | +2.32                 | +0.5%                    |
| Treatment 2 stores (T2)       | 468.87 (135.01)               | 468.58 (145.79)            | −0.29                 | −0.06%                   |
| Pooled T1 and T2              | 466.08 (135.87)               | 466.98 (145.58)            | +0.9                  | +0.19%                   |
| **Fridge freezer**            |                               |                            |                        |                          |
| Control stores                | 295.85 (82.70)                | 291.80 (76.42)             | −4.05                 | −1.37%                   |
| Treatment 1 stores (T1)       | 290.34 (81.40)                | 288.25 (74.65)             | −2.09                 | −0.72%                   |
| Treatment 2 stores (T2)       | 291.27 (81.17)                | 288.74 (74.11)             | −2.53                 | −0.87%                   |
| Pooled T1 and T2              | 290.82 (81.27)                | 288.50 (74.36)             | −2.32                 | −0.80%                   |
| **Dishwasher**                |                               |                            |                        |                          |
| Control stores                | 270.59 (28.26)                | 268.18 (28.80)             | −2.41                 | −0.89%                   |
| Treatment 1 stores (T1)       | 269.15 (28.95)                | 268.49 (28.94)             | −0.66                 | −0.25%                   |
| Treatment 2 stores (T2)       | 271.59 (28.23)                | 269.41 (28.74)             | −2.18                 | −0.80%                   |
| Pooled T1 and T2              | 270.41 (28.60)                | 268.95 (28.84)             | −1.46                 | −0.54%                   |

Standard deviation in parentheses
Table 6 illustrates the marginal effects from separate probit models for each of the four appliances to determine the probability of sales of high-efficiency products. The dependent variable is the energy efficiency of the product with a dichotomous specification with high-efficiency models determined as per Eq. 4. It can be seen that there is no significant effect for the treatment group, i.e., sales in treatment group stores during the trial period were not more likely to be high efficiency compared with the control group stores. The price coefficient is significant and of the expected sign, as is the case for the other product attributes.

Table 6 presents the marginal effects from a probit model, following a similar empirical approach to del Mar Solà et al. (2021), and finds no significant treatment effects. This result supports the findings in del Mar Solà et al. (2021) who found that label only treatment, i.e., a label with no sales-person intervention, did not have an impact for fridge freezers and dishwashers, although they did find a modest impact for washing machines.4

4 Tumble dryers were not considered in their analysis.
One of the limitations of del Mar Solà et al. (2021) is a lack of sales data in stores prior to the trial period. It can be seen from Table 5 that despite the random allocation of stores to control and treatment groups, the average kWh/a in treatment stores in the baseline period is marginally lower than in control stores for each appliance. Therefore, a difference in difference regression estimation is also tested to control for differences in the baseline (following the methodology in Carroll et al. (2016b)), with the coefficient of interest being $e \times z \times w$ (the kWh/a in the treatment stores in the trial period) from Eq. 6. As discussed previously, given the low usage of the QR code, treatments 1 and 2 stores are pooled and are considered together as “treatment”; however, the analysis is also conducted for the treatment groups separately and is presented in Tables 15 and 16.

From Table 7, it is apparent that again a null treatment effect is found for all products (shaded cells are the main coefficients of interest). Carroll et al. (2016b) also found a null treatment effect for monetary labels for tumble dryers; however, one of the limitations of that paper was a lack of store randomisation. Thus, this paper overcomes this non-randomisation finding and extends the analysis to three further appliance categories but still finds a null treatment effect.

One limitation with the data is that the sales system of the partner retailer, DID Electrical, does not capture the difference between no sales versus no availability at the store level (both show as zeros in the data). Therefore, a single sale on a particular model in one store automatically generates zero observations across the remaining stores, many of which may not have stocked the relevant model. Despite store visits by the research assistants, it was not possible to accurately monitor availability of all models in all stores at all times. Thus, a robustness check is run which focuses on the top 95% of appliance sales during the trial period and omits models which are not a regular feature in the typical retail stock, for example, models which are available for order only. This approach considerably reduces the number of zeros in the data; however, this lack of information on individual appliance availability at the store level should be considered a limitation of the analysis. Table 8 presents the results of the diff-in-diff model for the core group of appliances (i.e., the 95% of models sold).

Again, a null result from treatment is indicated, even when only core models are considered. A further robustness check is performed to account for label monitoring. As can be seen from the distribution of stores around the island of Ireland (Figs. 5 and 6 in Appendix 1), the majority of stores are in the midlands and east coast. While every effort was made to ensure the research assistants frequented the retail stores on a regular basis to ensure comprehensive label displays, this was logistically more difficult for the Galway and Bandon stores, the two most westerly treatment stores in the map in Fig. 5. Thus, a robustness check is run which considers only those treatment stores within the Leinster region, i.e., within 1.5-hour travel distance from the main university campus, omitting Galway and Bandon. Results are presented in Table 9.

The results of the restricted sample of core models and geographically “close” stores also indicate a null result for treatment. In summary, across all model specifications (both the probit model and the negative binomial model with robustness checks), there are no significant effects of the treatment.

The analysis so far has concentrated on store-level data in the cluster randomised controlled trial. In addition to the RCT, an in-store customer survey was also conducted across all stores. This targeted customers who were shopping for washing machines, tumble dryers, fridge freezers, and dishwashers (the survey instrument is available in Appendix 3). It should be noted that this is a special subsample of the population of customers in the cluster RCT and specifically represents customers who were willing to engage in the
Table 7  Negative binomial model results

| VARIABLES       | (1)                | (2)                | (3)                | (4)                |
|-----------------|--------------------|--------------------|--------------------|--------------------|
|                 | Washing Machine   | Tumble Dryer       | Fridge-Freezer     | Dishwasher         |
| $e$ (kwh/a)     | 0.00239**          | 0.00154***         | 0.00433***         | 0.0179***          |
|                 | (0.00120)          | (0.000429)         | (0.000592)         | (0.00170)          |
| $z$ (treatment) | -0.0438            | 0.269              | 0.566***           | 0.882              |
|                 | (0.274)            | (0.270)            | (0.165)            | (0.592)            |
| $w$ (trial period) | 0.583*            | -0.758***          | -0.904***          | 1.978***           |
|                 | (0.300)            | (0.287)            | (0.210)            | (0.707)            |
| $e \times z$   | 0.000323           | -0.000499          | -0.00199***        | -0.00318           |
|                 | (0.00149)          | (0.000524)         | (0.000572)         | (0.00214)          |
| $e \times w$   | -0.00444***        | 0.00108*           | 0.00238***         | -0.00827***        |
|                 | (0.00164)          | (0.000561)         | (0.000715)         | (0.00254)          |
| $z \times w$   | 0.150              | -0.0568            | -0.0882            | -0.466             |
|                 | (0.380)            | (0.372)            | (0.256)            | (0.855)            |
| $e \times z \times w$ | -0.000843       | 7.23e-05           | 0.000284           | 0.00168            |
|                 | (0.000207)         | (0.000726)         | (0.000883)         | (0.000308)         |
| Price           | -0.00718***        | -0.00556***        | -0.00418***        | -0.00526***        |
|                 | (0.000300)         | (0.000342)         | (0.000174)         | (0.000276)         |
| Price$^2$       | 2.31e-06***        | 1.59e-06***        | 9.26e-07***        | 1.98e-06***        |
|                 | (1.96e-07)         | (1.45e-07)         | (6.00e-08)         | (1.64e-07)         |
| Capacity (kg)   | 0.193***           | 0.231***           |                   |                   |
|                 | (0.0138)           | (0.0271)           |                   |                   |
| Fridge capacity (l) | 0.00202***     |                   |                   |                   |
|                 | (0.000395)         |                   |                   |                   |
| Freezer capacity (l) | 0.00128***    |                   |                   |                   |
|                 | (0.000479)         |                   |                   |                   |
| Noise (db)      | -0.0552****        | -0.183***          | 0.0158***          | -0.0165*           |
|                 | (0.00300)          | (0.0128)           | (0.00584)          | (0.00917)          |
| Delay function  | -0.341***          | -0.379***          | -0.115***          |                   |
|                 | (0.0382)           | (0.0451)           |                   | (0.0337)           |
| Constant        | -0.163             | 8.365***           | -4.712***          | -6.190***          |
|                 | (0.292)            | (0.906)            | (0.273)            | (0.706)            |

$^*$ $p < 0.1; ^{***} p < 0.05, ^{****} p < 0.001$, robust standard errors clustered at store level in parentheses, other controls include dummy variables for store and month (not reported for ease of presentation). The dependent variable is number of sales of model $i$ in store $j$ in month $t$. All models are estimated in STATA version 13 ("nbreg" command). Store monthly sales are used for the exposure variable. The Wald statistic is for a likelihood ratio test of the joint significance of the regressors.
survey. This subsample likely overrepresents customers who are not time constrained and who potentially are more motivated in the topic of energy efficiency. The limitations of this sample are discussed further in the “Limitations and Future Work” section.
The survey asked participants about the importance of energy efficiency in their choice of appliance through the question “How important is/was the energy efficiency of the appliance when making your decision?” Responses were given on a scale of 0 to 10 where 0 represents “I didn’t consider it at all” and 10 is “it was the most important aspect of my
Long-term Energy Cost Labelling for Appliances: Evidence from…

Table 10  Stated importance of energy efficiency in the customer survey

|               | EEImportance | All customers |             | Customers who made a purchase |             |
|---------------|--------------|---------------|-------------|-------------------------------|-------------|
|               | N            | Mean          | Stdev       | N                            | Mean        | Stdev       |
| All products  | 868          | 7.77          | 2.18        | 215                          | 7.58        | 2.31        |
| Washing machine | 220        | 7.83          | 2.22        | 75                           | 7.64        | 2.37        |
| Tumble dryer  | 102          | 8.08          | 1.74        | 30                           | 8.10        | 1.75        |
| Fridge freezer | 330         | 7.74          | 2.19        | 57                           | 7.26        | 2.35        |
| Dishwasher    | 87           | 7.38          | 2.39        | 25                           | 6.88        | 2.45        |
| Multiple products | 129      | 7.78          | 2.26        | 28                           | 8.11        | 2.35        |

Responses were given on a scale of 0 to 10 where 0 represents “I didn’t consider it at all” and 10 is “it was the most important aspect of my decision”; the mean is the average response to this question.

Table 11  Stated importance of energy efficiency in control and treatment stores

|               | EEImportance | Observations | Mean | Std. err. | Std. dev | 95% conf. interval | t   | df   | P             |
|---------------|--------------|--------------|------|-----------|----------|--------------------|-----|------|----------------|
|               |              |              |      |           |          |                    |     |      | (Ha: control = treated) |
| Control       | 373          | 7.88         | 0.11 | 2.09      | 7.66     | 8.09               | 1.42| 984  | 0.15           |
| Treatment     | 613          | 7.67         | 0.09 | 2.32      | 7.49     | 7.86               |     |       |                |

Statistics are estimated in STATA version 13 (“ttest” command)

decision” and the resulting variable is called EEImportance with discrete values ranging from 1 to 10. Table 10 presents the responses to these questions.

It can be seen from Table 10 that the average stated importance of energy efficiency is high for all product types (above 7.2 in almost all cases), although marginally lower for those customers who ultimately purchased an appliance versus those who did not.

It is interesting to examine if the new labels increased the stated importance of energy efficiency in the treatment stores versus the control stores. Table 11 presents the results of a t-test of responses to this survey question between customers in the control and treatment stores. It can be seen from Table 11 that there is no significant difference in the stated energy efficiency importance of respondents between control and treatment stores.

The intention of the study was to “treat” all customers of the four appliances in the treatment stores, i.e., all customers in treatment stores are considered “treated.” However, question 8 in the customer survey specifically asked customers if they had “seen the labels showing the 10-year running costs of the appliances.” The responses to this question in the survey indicate those who were actually “treated” rather than “intention to treat.” Therefore, Table 12 presents the results of a sub-group of customers in the treatment stores, who reported to have “seen the label.” This cohort is compared with the control group customers (from control stores where there were no labels). It is seen that customers who “saw the label” in the treatment stores are more likely to report a higher stated importance of energy efficiency, a result which is statistically significant.
Table 12 highlights a limitation of this paper as it can be seen that just 183 of respondents in the treatment stores reported having “seen” the label. This represents just 30% of those sampled in the treatment stores. While recognising that the survey sample is not a representative sample of shoppers, it does highlight a concern regarding the salience of energy efficiency labelling on appliances. This limitation is discussed further in the “Limitations and Future Work” section.

The customer survey also collected demographic information on customers (see Table 4 for summary) and Table 13 illustrates the relationship between stated energy efficiency importance and key customer characteristics. Given the ordered and discrete nature of the variable $EEImportance$, Table 13 presents the results of an ordered logit model and shows odds ratios (rather than the coefficients of the logit model) for ease of interpretation. An odds ratio greater than 1 implies that the explanatory variable increases the odds of a higher value of $EEImportance$. A value of less than 1 implies the variable reduces the odds of having a higher value of $EEImportance$.

It is apparent from Table 13 the stated importance of energy efficiency is higher in treatment stores (significant at the 5% level). It is also evident that increased age and education are related to an increased probability of reporting a higher importance of energy efficiency. The odds ratio for landlords suggests that they are less likely to report high importance of energy efficiency, a result which is intuitive, although the $p$-value is just outside the 10% significance level. It can also be seen that people who report themselves as having more patience (a proxy for a lower personal discount rate) are more likely to report a higher importance of energy efficiency.

The findings illustrated in Table 12 and 13 indicate that there is a positive correlation between monetary labelling and stated importance of energy efficiency. The results presented here are from a convenient sample of in-store customers which is a very specific and limited sample. Therefore, it is useful to also consider results from studies with similar research questions which use broader samples such as Andor et al. (2017), Davis and Metcalf (2016), Min et al. (2014), Sammer and Wüstchenhagen (2006), and Heinzle (2012) which show that stated valuation of energy efficiency is increased by monetary labelling.

Recognising that the survey sample in this paper is not a random or representative sample, it is nevertheless interesting to consider if the stated preferences for energy efficiency in survey responses translated into revealed behaviour. This can be examined by looking at the actual purchases of the individual customers who participated in the survey, i.e., did the labels increase the energy efficiency of individual customer purchases in treatment stores (Eq. 8 previously). Table 14 illustrates the results of a probit model where the dependent

---

Table 12  
Stated importance of energy efficiency for those who saw the label

|               | Observations | Mean | Std. err. | Std. dev | 95% conf. interval | t    | df | P (Ha: control = treated) |
|---------------|--------------|------|-----------|----------|--------------------|------|----|--------------------------|
| Control       | 373          | 7.88 | 0.11      | 2.09     | 7.66 - 8.09        | -2.47| 554| 0.01                     |
| Saw label †   | 183          | 8.32 | 0.12      | 1.68     | 8.07 - 8.56        |      |    |                          |

† “Saw label” in this table refers to those surveyed in the treatment stores who stated that they were treated, i.e., stated they had seen the label. Statistics are estimated in STATA version 13 (“ttest” command).
variable takes the value 1 if the appliance purchased was deemed “high efficiency” (as per Eq. 4), and 0 otherwise. It can be seen that customers who were surveyed in the treatment stores were not more likely to purchase high-efficiency appliances compared to customers who were surveyed in the control group.

The analysis in Table 14 was also conducted for the smaller sub-group of customers in treatment stores who confirmed they saw the labels and the findings remain unchanged, and those who saw the labels were not more likely to purchase high-efficiency appliances compared to the control group. This is further supported by question 10 in the survey (which asked directly, for those who saw the label, if it had influenced their purchasing decision). An analysis of the results of this question shows that just 29% of respondents reported that the label led them to a more efficient purchase.

In summary, for treated customers in the treatment stores (those who stated they saw the labels, as per Table 11), the label increased their stated importance of energy efficiency. However, for these treated customers, we do not see greater energy efficiency in their actual purchase decisions compared to customers in control stores (Table 14). This was also tested for all product types and interacted with demographics and the results are consistent, and the label did not influence purchases for any product type or demographic group (age, income, education, gender).

### Table 13: Ordered logistic regression of energy efficiency importance

| Variable       | Odds ratio | Robust std. err. | z     | P>|z| | 95% conf. interval |
|----------------|------------|------------------|-------|------|-------------------|
| Treated        | 1.64       | 0.42             | 1.97  | 0.05 | 1.00 to 2.70      |
| Age            | 1.23       | 0.06             | 4.27  | 0.00 | 1.12 to 1.36      |
| Education      | 1.25       | 0.12             | 2.23  | 0.03 | 1.03 to 1.52      |
| Income         | 0.95       | 0.06             | -0.85 | 0.40 | 0.83 to 1.08      |
| Male (D)       | 1.14       | 0.18             | 0.84  | 0.40 | 0.84 to 1.54      |
| Landlord (D)   | 0.13       | 0.17             | -1.60 | 0.11 | 0.01 to 1.58      |
| # of adults    | 1.11       | 0.10             | 1.17  | 0.24 | 0.93 to 1.32      |
| # of children  | 1.07       | 0.06             | 1.24  | 0.22 | 0.96 to 1.19      |
| Patience       | 1.05       | 0.03             | 1.63  | 0.10 | 0.99 to 1.12      |
| Risk           | 0.99       | 0.04             | -0.17 | 0.86 | 0.92 to 1.07      |
| /cut1          | -1.12      | 0.64             |       |      |                   |
| /cut2          | -0.73      | 0.60             |       |      |                   |
| /cut3          | -0.51      | 0.58             |       |      |                   |
| /cut4          | -0.28      | 0.56             |       |      |                   |
| /cut5          | 0.22       | 0.52             |       |      |                   |
| /cut6          | 0.53       | 0.51             |       |      |                   |
| /cut7          | 1.24       | 0.52             |       |      |                   |
| /cut8          | 2.29       | 0.55             |       |      |                   |
| /cut9          | 3.14       | 0.55             |       |      |                   |
| Observations   | 516        |                  |       |      |                   |

Robust standard errors clustered at the store level. Dependent variable is EEImportance with discrete values from 0 to 10. Model is estimated in STATA version 13 ("ologit" with “or” command). D represents dummy variable.
Limitations and Future Work

As discussed in Ioannidis (2005), the probability that a research claim is true and that a statistically positive result indicates a true relationship is often inaccurate. They state that a finding is less likely to be true when, amongst other factors, there is smaller sample size and when effect sizes are smaller. Statistical power is an essential parameter to assess the value of the results, with power of 0.8 generally being considered adequate. The cluster RCT in this paper has low power (0.38 with 22 clusters, $\alpha = 0.05$ and $\delta = 0.8$). Low power implies higher rates of false negatives—meaning that there could be an effect but given the low power of the study, that this cannot be detected.

A balanced study has approximately equal observations in control and treatment groups which can be seen, is not the case in this study due to the combination of treatments 1 and 2 (due to the low uptake of the QR code). When considered separately, see Appendix Tables 15, 16 and 17, the sales across the three groups are balanced, at approximately 14,100 units sold in each of control, treatment 1 and treatment 2 groups, and the results are consistent. Nevertheless, an unbalanced study has less statistical power.

Given that the RCT in this study finds a null effect plus has low power raises the question that the result is a false negative. Ioannidis et al. (2017) find that empirical economics research is generally underpowered with typical statistical power of no more than 18% in half of the areas they studied and over 90% of reported results stemming from underpowered studies. As discussed in Christensen and Miguel (2018), one possible approach to address underpower is to employ larger data sets. In this context, power tests suggest that in order to achieve adequate power for the case study considered in this paper, the number of stores considered would need to be in excess of 60 separate stores. There is no such electrical retailer in Ireland with this many stores; thus, in order to increase power, future research should consider engaging with multiple retail chains (which may cause different problems in terms of implementation), experimentation in a larger jurisdiction with a retail chain with sufficient number of stores, or a treatment with individual-level outcomes.

The sample for the surveys in the treatment and control stores is a convenient sample of customers who were in-store at the time of sampling and who were willing to participate in the survey, and it is not a random or representative sample. This self-selection limits the ability to generalise the results of the consumer survey to the broader population of customers and increases the likelihood that the sample overrepresents customers with time.

### Table 14 Impact of label on individual customer purchases

| Probability of high-efficiency purchase | 1 | 2 |
|-----------------------------------------|---|---|
| Coef. (std err) | $z$ | $P>z$ | Coef. (std err) | $z$ | $P>z$ |
| Treated | 0.101 (0.14) | 0.71 | 0.477 | 0.087 (0.16) | 0.55 | 0.585 |
| Demographic controls included | No | - | - | Yes | - | - |
| Constant | $-0.52$ (0.08) | $-6.63$ | $0.000$ | $-0.983$ (0.431) | $-2.28$ | 0.023 |
| Observations | 201 | | | 199 | | |

Robust standard errors clustered at the store level in parentheses. Model is estimated in STATA version 13 (“probit” command). The dependent variable takes a value of 1 if the customer purchased a product of high efficiency (as defined in Eq. 4) and 0 otherwise.
and/or interest in energy efficiency. For these reasons, generalisations of the results from the customer survey presented in this paper should not be done and results should be considered within the context in which the data was collected. The focus of the discussion of the results of the surveys is on a comparison of responses collected in the control and treatment stores.

For a more robust individual-level investigation, a random and/or representative sample of individuals should be collected. This could be done in future work for example by engaging with a consumer organization to randomly select customers for surveying, by randomly sampling customers through retail websites, or to use a professional panel company to recruit a representative sample. The latter approach is adopted in Ceolotto and Denny (2021) who collect representative samples across four countries to determine the impact of personalised energy efficiency labels using a multi-country discrete choice experiment. Findings from that paper indicate that personalised energy information increases the willingness-to-pay for energy efficiency in the UK, whereas monetary information has a negative impact in Canada. No significant effect is detected in Ireland and the USA. That paper also finds that providing monetary information crowds out individuals who would buy a more efficient product for environmental reasons.

The results of Ceolotto and Denny (2021) suggest a considerable element of context-specific results for monetary labelling which supports the mixed results found across the literature. For example, cost labelling is found to be effective in revealed preference studies for lightbulbs in Germany (Andor et al., 2019b), washing machines but not dishwashers in Spain (del Mar Solà et al., 2021), tumble dryers with sales-person training in Norway (Kallbekken et al., 2013) but not for lightbulbs in the USA (Allcott & Taubinsky, 2015), or tumble dryers in the UK or Ireland (Carroll et al., 2016b; DECC, 2014). These context-specific differences are also seen in stated preference studies for example with positive impacts found for refrigerators in Germany (Andor et al., 2019a; Andor et al., 2020), televisions in Switzerland (Heinzle, 2012), lightbulbs in the USA (Min et al., 2014), air conditioners in the USA (Davis & Metcalf, 2016) but not for refrigerators in Greece (Skourtos et al., 2021) and as mentioned previously not for tumble dryers in Ireland, the USA, and Canada (Ceolotto & Denny, 2021).

The reasons for context-dependent results could stem from not only different methods and products but also other factors such as different electricity prices across countries. For example, for the countries listed above, residential electricity prices range from €0.32 per kWh in the first half of 2021 in Germany (the highest in the EU) (Eurostat, 2021) to €0.093 equivalent ($0.1054) per kWh in the USA (Statistica, 2021). However, the results of Davis and Metcalf (2016) suggest that cost labelling is effective for air conditioners in the USA despite low electricity prices; thus, other more fundamental differences between contexts may play a role such as different knowledge, beliefs, culture, trust, political views, and discount rates of individuals in different countries. Recent work by Houde and Myers (2021a) explores the relationship between local electricity prices and purchase prices and indicates that electricity prices which deviate from social marginal costs, for example due to the failure to capture externalities, can have a significant distortion on demand.

While the label was installed on each of the four appliance types in each of the treatment stores, and the label display was monitored by research assistants visiting stores throughout the trial, it is not possible to determine the extent to which customers actually engaged with it. As seen in Fig. 1, the new label was designed with a red background and an eye-catching image to ensure it stood out against the
white background of the electrical appliances. However, the results in Table 12, albeit not representative of the customer population, suggest that just 30% of shoppers surveyed noticed the new label. The facade of electrical appliances in-store is full of information in addition to energy efficiency labelling with signage and labels for pricing offers, payment instalment offers, delivery and installation information, warranties, and product information such as technical specifications, innovative features, and information on capacity and cycle times. The purpose of the new label was to attempt to draw attention to energy efficiency amongst this noise of information; however, it remains a limitation that this may not have been sufficiently achieved; thus, from a theoretical perspective, the results here could be considered a lower bound of the true effects.

While field trials provide advantages in terms of realistic settings for decision-making, they can be marred by elements external to the experimenter’s control, such as all of the other information labels listed above. For this reason, results from laboratory and discrete choice experiments provide valuable insights in that the information presented to individuals can be controlled. For example, Ceolotto and Denny (2021), using a representative sample, consider tumble dryer choices in the context of information provided on the attributes customer rating, capacity, brand, and energy efficiency (presented as letter grade, monetary cost, or personalised monetary cost). They find that energy efficiency is the attribute with highest willingness to pay across each of the four countries considered. However, it should be noted that the “invisibility” of energy efficiency information amidst all the other commercial information on in-store displays is a concern for all energy efficiency labelling, not just this new label and is the reality of the environment faced by customers when making investment decision.

Conclusions

This paper examines the impact of long-term energy cost labelling on four common household appliances: washing machines, tumble dryers, fridge freezers, and dishwashers. It considers the impact of a new energy label which provides information on estimated 10-year energy costs based on typical household usage and also a QR code which facilitates personalisation of this information. The effectiveness of the new labels are tested in a cluster randomised controlled trial in collaboration with a retail partner, where labels are displayed on appliances in-store for a randomly selected group of stores. The RCT is supplemented by an in-store customer survey.
The paper makes a number of key findings in the context of the case study examined. First, results suggest that the long-term energy cost label did not increase the efficiency of appliances purchased. This was apparent across a range of model specifications. Second, engagement with the QR code was extremely low. Given that the newly designed EU energy labels, due to be rolled out in 2021, incorporates QR codes for personalisation, this low uptake suggests that this element of the new labels could be ineffective in increasing the uptake of energy efficiency. Third, the customer survey indicates that the label increased customers’ stated preference for energy efficiency but this did not translate into revealed behaviour of more efficient purchases.

There is a large body of literature indicating that monetary labelling increases stated preferences for energy efficiency. The stated preference results presented in this paper, while not a representative sample, indicate that monetary labelling increases the stated preferences for energy efficiency for those who saw the labels; however, for the individuals sampled, this stated preference did not translate into a revealed preference for greater energy efficiency.

This paper is one of the first to examine both stated and revealed behaviour in the same setting. The results of the RCT suggest that the monetary labels did not impact on the efficiency of appliances purchased. Future work recommends more comprehensive sampling for the survey to allow for a direct comparison of stated preference and revealed preference data.

This paper overcomes some limitations of other papers in the field by examining multiple appliances, randomisation in experimental design, a rich baseline dataset, and consumer-level data. Furthermore, it is the first to combine energy cost labelling with a QR code for personalisation of estimates. Previous stated preference studies have indicated that willingness to pay for energy efficiency increases with personalised information; however, the delivery of this information to the customer is a challenge and the low usage of the QR code in this paper indicates that further work is required to determine the optimal way of delivering this information.

The logistics of physical energy cost labelling are significant and labels will vary by country/region due to differences in electricity prices. It is also unclear the extent to which customers have already selected their appliance before coming to the store, for example, through online research and information on the retailer’s website. Likewise, the customer survey in this paper highlights concerns around the lack of attention to energy efficiency labelling amidst the noise of other product information on display in-store. The combination of these factors suggests that future work should consider testing energy cost information through online interventions with retailer websites. This would also help overcome the limitations of low power in this study. Furthermore, if energy cost labels are considered in an online setting, then perhaps other costs should also be included, such as water charges, recycling costs, life-cycle emissions, and CO₂ costs.
Appendix 1

Fig. 5 Geographical dispersion of retail stores. Note: All stores were randomly allocated to control and treatment. Stores in red are control stores, and green represents treatment stores. The circle represents the Dublin area in which there were 7 randomly allocated treatment stores and two control stores.
Fig. 6 EU energy label. Source: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:314:0047:0063:EN:PDF
Appendix 2. Calculating the personalised energy consumption for fridge freezers

The key assumptions used in the calculations for the personalisation of fridge/freezer energy consumption are based on the assumptions in the EU Commission Regulation (EC) No. 643/2009 with regard to ecodesign requirements for household refrigeration appliances and the Commission Delegated Regulation (EU) No. 1060/2010 with regard to energy labelling of household refrigeration appliances (European Commission, 2010).

Assumptions underlying EU fridge freezer energy consumption labels:

- Heat capacity of water = 4.2 kJ/kg/K (assumed also for food stuffs)
- Heat capacity of air (at 20°C) = 0.83 kJ/m³/K
- Ambient temperature in kitchen is 20°C
- Food load of 1000 kg per person per annum
- Coefficient of performance (COP) of 2.5
- When fridge door is opened, there is a temperature differential of 15°C
- When freezer door is opened, there is a temperature differential of 38°C
- 1 MJ = 3.6 kWh

Assumptions used in converting energy consumption based on average household to personalised consumption for the purposes of treatment 2:

- Based on the EU average of 2.4 people per household and average number of fridge/freezer openings of 20 per day from European Commission (2009), it is assumed that each person opens the fridge on average 8 times per day and freezer 2 times per day.
- Constant coefficient of performance, i.e., COP does not vary by model due to lack of data.

Warm load calculations:

The formula used in the EU energy label for fridge freezers is given by Eq. 9

\[
\text{Energy demand (MJ)} = \frac{(\text{Food Load}) \times (\#\text{people}) \times (\text{Temperature Differential}) \times (\text{Heat Capacity Of Water & Food})}{1000}
\]

To convert energy demand in MJ to kWh electricity consumption, divide by \((3.6 \times \text{COP})\)

\[
\text{Electricity consumption kWh} = \frac{(1000\text{kg}) \times (\#\text{people}) \times (15^\circ\text{C}) \times (4.2\text{kJ/kg/K})}{1000} \frac{1}{3.6 \times 2.5}
\]

Door openings:

It is assumed there are 8 fridge openings and 2 freezer opening per person per day. Thus, the larger the household, the more openings and the greater the air capacity of the fridge/freezer to be replaced. This is calculated as:

Air capacity of fridge and freezer to be replaced:
Fridge: 

\[ m^3 \text{ of air capacity to be replaced} = (8 \times \# \text{people} \times 365) \times (\text{fridge capacity (l)}/1000) \]

Freezer: 

\[ m^3 \text{ of air capacity to be replaced} = (2 \times \# \text{people} \times 365) \times (\text{freezer capacity (l)}/1000) \]

Energy demand (kJ) of air capacity to be replaced = \( (m^3 \text{ air capacity}) \times (\text{heat capacity of air}) \times (\text{temperature differential}) \)

For fridge 

\[ \text{kJ} = (m^3 \text{ air capacity}) \times (0.83) \times (15^\circ C) \]

For freezer 

\[ \text{kJ} = (m^3 \text{ air capacity}) \times (0.83) \times (38^\circ C) \]

To convert to kWh electricity consumption divide by \((1000 \times 3.6 \times \text{COP})\)

Thus, total variable consumption is given by Eq. 10 plus replaced air capacity for fridge freezer is

\[
\text{Total variable consumption} = \frac{(1000 \times \# \text{people} \times (15 \times (4.2) \times \left( \frac{1}{5.6 \times 2.5} \right))}{1000} \\
+ \frac{(8 \times \# \text{people} \times 365) \times (\text{fridge capacity (l)}/1000)(0.83 \times 15) + (2 \times \# \text{people} \times 365) \times (\text{freezer capacity (l)}/1000)(0.83 \times 38)}{1000 \times 3.6 \times 2.5}
\]

Tidying up this expression gives Eq. 11 below

\[
\text{Total variable kWh}_\text{annum} = \# \text{people} \times (7 + 4.04 \times (\text{fridge capacity (l)}/1000) + 2.56 \times (\text{freezer capacity (l)}/1000)) \times 10 \times 0.1763
\]

(11)

For the personalised cost information provided in treatment 2, this calculated amount in Eq. 11 needs to then be multiplied by \(10\) years and the electricity cost per kWh \((€0.1763)\) and added to the annual kWh figure from the EU label for each appliance.

The EU energy label already captures implied variable consumption for an assumed family size of 2.4; thus, this element of variable consumption for the average household size is subtracted from the kWh figure given by the EU label giving.

\[
\text{Variable kWh}_\text{10 years} = (\# \text{people} - 2.4) \times \left( 7 + 4.04 \times (\text{fridge capacity (l)}/1000) + 2.56 \times (\text{freezer capacity (l)}/1000) \right) \times 10 \times 0.1763
\]

(12)

The calculated number from Eq. 12 then needs to be added to the annual kWh figure from the EU label for that particular machine.

Example for 2-person household with a 300-fridge freezer (200 l fridge, 100 l freezer) and an A++ rating (153 kWh per annum):

Using Eq. 12 above, the additional variable cost of running this machine is:

\[
(2 - 2.4) \times (7 + 4.04 \times (0.2) + 2.56 \times (0.1)) \times 10 \times €0.1763 = -€5.69
\]

This needs to be added to the average 10-year running cost of the machine.

\[
\text{Total 10 - year cost} = (153 \text{kWh} \times 10 \times 0.1763) + (-5.69) = €264.05
\]

The same machine for a six-person household is estimated at

\[
(153 \times 10 \times 0.1763) + (6 - 2.4) \times (7 + 4.04 \times (0.2) + 2.56 \times (0.1)) \times 10 \times 0.1763 = €320.92
\]
Appendix 3. In-store survey

Consumer Survey

This research examines the decision making process for buying household appliances. We are trying to find out what factors are important to you in purchasing a household appliance. This research is being carried out in Trinity College Dublin and has NO commercial purpose. Your answers are 100% CONFIDENTIAL and anonymous and your details and opinions will never be given to a third party.

This research is being led by Prof. Eleanor Denny in the Economics Department at Trinity College Dublin (dennyne@tcd.ie). Full details of the research project can be found at www.neepd.com.

You will be asked an absolute maximum of 20 questions in this survey and all questions are optional. Your participation in this survey is voluntary and you may withdraw at any time and for any reason without penalty. In the extremely unlikely event that illicit activity is reported during this study, the lead researcher will be obliged to report it to appropriate authorities.

At the end of the survey you can enter an optional draw to win a €500 One4All voucher.

I am over 18 and I consent to participating in this survey: Yes ☐ No ☐

Q1. What appliance(s) were you shopping for today? You can select more than one
   • Washing machine ☐
   • Dishwasher ☐
   • Tumble dryer ☐
   • Fridge-freezer ☐
   • None of the above ☐

Q2. Did you purchase any appliances today?
   • Yes ☐
   • No ☐
   (if no, skip to question 5)

Q3. If Yes, please indicate the make and model of the appliances purchased
   Make ___________________________
   Model ___________________________
   Energy rating (A+, A, B, C etc.) ____________

Q4. If Yes, will you be paying the electricity bills in the property where the ‘appliance’ will be located?
   • I pay all of the electricity bill ☐
   • I pay part of the electricity bill ☐
   • I do not pay any of the electricity bill ☐

Q5. Are you concerned about environmental issues such as climate change? On a scale of 0 to 10 where 0 represents ‘completely unconcerned’ and 10 is ‘extremely concerned’.

Q6. How important is/was the energy efficiency of the appliance when making your decision?
   On a scale of 0 to 10 where 0 represents ‘I didn’t consider it at all’ and 10 is ‘it was the most important aspect of my decision’.

   __________

Q7. Which of the following best describes you?
   a. I am a homeowner and am shopping for my main residence ☐
   b. I am a homeowner and am shopping for a secondary residence/holiday home ☐
   c. I am a landlord and am shopping for my rental property ☐
   d. I am a tenant and am shopping for my rental property but the landlord will cover the purchase cost of any items I buy ☐
   e. I am a tenant and am shopping for my rental property but I will cover the purchase cost myself (or share this cost with the other tenants) ☐
   f. I am shopping on behalf of a third party (e.g. other family member, workplace) ☐
   g. Other ________________________________ ☐

Q8. Did you see the labels showing the 10 year running costs of the appliances?
   • Yes ☐
   • No ☐ (if no, skip to question 12)

Q9. To what extent did seeing the running costs influence your ultimate purchasing decision?
   Scale of 0 to 10 with 0 being ‘not at all’ and 10 being ‘it was critical in my ultimate decision’
Q10. Do you think you would have chosen a different appliance if you hadn’t seen this running cost information?
   a. Yes, I would have chosen an appliance which would have ended up costing more to run
   b. Yes, I would have chosen an appliance which would have ended up costing less to run
   c. No, I would have chosen the same appliance

Q11. Did you scan the QR codes and use the app to see how much the appliances cost to operate?
   • Yes ⬜
   • No ⬜

In the next few questions we would like to ask you some questions about yourself. This is purely to identify if the people who are helping us with this survey are representative of the Irish population as a whole. Please remember that all this information is 100% anonymous and confidential. At no point in this survey will we ask you to identify yourself.

Q12. How old are you?
   • 18-29 ⬜
   • 30-39 ⬜
   • 40-49 ⬜
   • 50-59 ⬜
   • 60 – 69 ⬜
   • 70 and over ⬜

Q13. Please indicate the highest level of education you have completed?
   • Primary school ⬜
   • Secondary school ⬜
   • Diploma/certificate ⬜
   • Primary degree ⬜
   • Master degree ⬜
   • PhD degree ⬜

Q14. Gender
   • Male ⬜
   • Female ⬜
   • Other ⬜

Q15. How many people live in your household?
   ______ Adults _______ Children

Q16. How would you describe your present status?
   • Working for payment or profit ⬜
   • Unemployed ⬜
   • Student ⬜
   • Looking after home/family ⬜
   • Retired ⬜
   • Unable to work due to permanent sickness or disability ⬜
   • Other ⬜

Q17. If working for payment or profit, please classify your employment type:
   • Employer or manager ⬜
   • Professional ⬜
   • Non-manual ⬜
   • Skilled manual ⬜
   • Semi-skilled manual ⬜
   • Unskilled manual ⬜
   • Farmer/Agriculture ⬜
   • Other ⬜

Q18. Are you generally an impatient person, or someone who always shows great patience? Please rate yourself on a scale of 0 to 10 where 0 represents ‘very impatient’ and 10 is ‘very patient’.

Q19. How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please indicate where you are on a scale of 0 to 10 where 0 represents ‘unwilling to take risks’ and the 10 represents ‘fully prepared to take risk’. 

Q20. What is your household monthly income (including any social welfare payments) after tax? This survey is completely anonymous, and all questions are completely optional however the answer to this question helps us determine how representative our sample is of the Irish population.
   1) €0 - €1000 ⬜
   2) €1001 - €2000 ⬜
   3) €2001 - €3000 ⬜
   4) €3001 - €4000 ⬜
   5) €4001 - €5000 ⬜
   6) €5001 - €6000 ⬜
   7) €6001 - €7000 ⬜
   8) More than €7000 ⬜

Thank you so much for taking the time to help us with our research. If you would like to be in with the chance of winning a €500 one4all voucher please write your contact details on the card and post them in the raffle box.
Table 15  Marginal effects for probit model with treatments 1 and 2 separated

| Marginal effects          | (1) Washing machine | (2) Tumble dryer | (3) Fridge freezer | (4) Dishwasher |
|---------------------------|---------------------|-----------------|--------------------|---------------|
| Control                   | Reference           |                 |                    |               |
| Treatment 1               | 0.01238 (0.01675)   | −0.00008        | 0.05659 (0.01989)**| −0.00466 (0.01853) |
| Treatment 2               | 0.01348 (0.01037)   | −0.00038        | 0.01880            | −0.00802 (0.01359) |
| **Product attributes**    |                     |                 |                    |               |
| Price                     | 0.00049*** (0.00009) | 0.00035***      | 0.00087***         | −0.00002 (0.00003) |
| Price^2                   | −0.00000 (0.00000)  | −0.00000***     | −0.00000***        | −0.00000*** (0.00000) |
| Capacity (kg)             | 0.13148*** (0.00303) | 0.00604***      |                    |               |
| Fridge capacity (l)       |                     |                 | −0.00042***        |               |
| Freezer capacity (l)      |                     |                 | −0.00239***        |               |
| Noise (db) (db)           | −0.00444*** (0.00037) | 0.00445***      | −0.00619***        | −0.13058*** (0.00457) |
| Delay function            | −0.29718*** (0.01034) | 0.01330***      | 0.19516***         |               |
| Observations              | 12,460              | 6466            | 13,314             | 10,223        |
| Log Likelihood            | −3263.17            | −693.24         | −5204.74           | −4878.69      |
| Wald chi sq               | 14,472.19           | 74,638.02       | 41,722.31          | 38,764.92     |

* p < 0.1; *** p < 0.05, **** p < 0.001, standard errors in parentheses, other controls include dummy variables for store and month (not reported for ease of presentation). The dependent variable is high efficiency which is specified in Eq. 4.
Table 16  Negative binomial regression for treatment 1

| Variables          | Washing machine | Tumble dryer | Fridge freezer | Dishwasher |
|--------------------|-----------------|--------------|---------------|------------|
| e (kwh/yr)        | 0.00303***      | 0.00126***   | 0.00362***    | 0.0179***  |
|                   | (0.000899)      | (0.000333)   | (0.000514)    | (0.00122)  |
| z1 (treatment 1)  | 0.279           | 0.0574       | 0.503***      | 1.770***   |
|                   | (0.298)         | (0.307)      | (0.171)       | (0.605)    |
| w (trial period)  | 0.739***        | −0.792***    | −0.946***     | 1.966***   |
|                   | (0.216)         | (0.219)      | (0.148)       | (0.492)    |
| e × z1            | −0.00139        | −0.000120    | −0.00182***   | −0.00633***|
|                   | (0.00161)       | (0.000593)   | (0.000597)    | (0.00219)  |
| e × w             | −0.00534***     | 0.00112***   | 0.00249***    | −0.00822***|
|                   | (0.00119)       | (0.000424)   | (0.000516)    | (0.00177)  |
| z1 × w            | −0.196          | −0.0141      | −0.0487       | −0.819     |
|                   | (0.406)         | (0.402)      | (0.255)       | (0.836)    |
| e × z1 × w        | 0.00118         | 3.78e-05     | 0.000212      | 0.00293    |
|                   | (0.00221)       | (0.000783)   | (0.000896)    | (0.00304)  |
| Price             | −0.00718***     | −0.00556***  | −0.00418***   | −0.00525***|
|                   | (0.000299)      | (0.000343)   | (0.000174)    | (0.000274) |
| Price^2           | 2.32e-06***     | 1.59e-06***  | 9.22e-07***   | 1.98e-06***|
|                   | (1.96e-07)      | (1.45e-07)   | (5.96e-08)    | (1.62e-07) |
| Capacity (kg)     | 0.193***        | 0.231***     | 0.231***      | 0.00202*** |
|                   | (0.0138)        | (0.0271)     | (0.004095)    | (0.000479) |
| Fridge capacity (l)|                 |              | 0.00129***    | (0.000479) |
|                   |                 |              | (0.000479)    |            |
| Freezer capacity (l)|               |              |                |            |
| Noise (db)        | −0.0553***      | −0.183***    | 0.0159***     | −0.0165*   |
|                   | (0.00300)       | (0.0129)     | (0.00583)     | (0.00918)  |
| Delay             | −0.341***       | −0.379***    | −0.115***     | −0.115***  |
|                   | (0.0381)        | (0.0451)     | (0.0337)      | (0.0337)   |
| Constant          | −0.276          | 8.512***     | −4.510***     | −6.218***  |
|                   | (0.250)         | (0.890)      | (0.258)       | (0.615)    |
| α                 | 1.95            | 1.55         | 2.36          | 1.83       |
| Observations      | 14.652          | 7436         | 18.743        | 15.554     |
| Log likelihood    | −18.976.8       | −8431.40     | −18.457.63    | −13.810.11 |
| Wald Chi sq       | 3054.35         | 1033.01      | 1519.75       | 1310.62    |
| Prob > chi2       | 0.000           | 0.000        | 0.000         | 0.000      |

* p < 0.1; *** p < 0.05, **** p < 0.001, robust standard errors clustered at store level in parentheses, other controls include dummy variables for store and month (not reported for ease of presentation). The dependent variable is number of sales of model $i$ in store $j$ in month $t$. All models are estimated in STATA version 13 (“nbreg” command). Store monthly sales are used for the exposure variable. The Wald statistic is for a likelihood ratio test of the joint significance of the regressors.
Table 17  Negative binomial regression for treatment 2

| Variables | (1) Washing machine | (2) Tumble dryer | (3) Fridge freezer | (4) Dishwasher |
|-----------|---------------------|-----------------|--------------------|---------------|
| $e$ (kwh/a) | 0.00206** (0.000943) | 0.00136*** (0.000340) | 0.00317*** (0.000508) | 0.0147*** (0.00126) |
| $z2$ (treatment 2) | −0.328 (0.279) | 0.220 (0.293) | 0.112 (0.175) | −0.922 (0.600) |
| $w$ (trial period) | 0.565** (0.225) | −0.780*** (0.218) | −0.951*** (0.150) | 1.554*** (0.495) |
| $e \times z1$ | 0.00174 (0.00152) | −0.000397 (0.000564) | −0.000369 (0.000618) | 0.00326 (0.00217) |
| $e \times w$ | −0.00431*** (0.00123) | 0.00112*** (0.000427) | 0.00254*** (0.000520) | −0.00676*** (0.00179) |
| $z2 \times w$ | 0.360 (0.385) | −0.0415 (0.398) | −0.0380 (0.256) | 0.398 (0.841) |
| $e \times z2 \times w$ | −0.00209 (0.00212) | 3.00e-05 (0.000770) | 7.44e-05 (0.000901) | −0.00140 (0.00035) |
| Price | −0.00717*** (0.000300) | −0.00557*** (0.000343) | −0.00419*** (0.000174) | −0.00526*** (0.000275) |
| Price^2 | 2.31e-06*** (1.96e-07) | 1.59e-06*** (1.46e-07) | 9.27e-07*** (5.97e-08) | 1.99e-06*** (1.63e-07) |
| Capacity (kg) | 0.192*** (0.0138) | 0.231*** (0.0271) | 0.00205*** (0.000396) | 0.00132*** (0.000480) |
| Fridge capacity (l) | 0.00205*** (0.000396) | 0.00132*** (0.000480) | 0.00205*** (0.000396) | 0.00132*** (0.000480) |
| Freezer capacity (l) | 0.192*** (0.0138) | 0.231*** (0.0271) | 0.00205*** (0.000396) | 0.00132*** (0.000480) |
| Noise (db) | −0.0553*** (0.00299) | −0.183*** (0.0129) | 0.0159*** (0.00584) | −0.0161* (0.00918) |
| Delay | −0.341*** (0.0381) | −0.379*** (0.0450) | −0.116*** (0.0336) | −0.116*** (0.0336) |
| Constant | −0.0839 (0.259) | 8.449*** (0.889) | −4.391*** (0.258) | −5.325*** (0.619) |
| $\alpha$ | 1.95 (0.259) | 1.55 (0.889) | 2.36 (0.258) | 1.84 (0.619) |
| Observations | 14,652 (3048.7) | 7436 (1047.5) | 18,743 (1505.92) | 15,554 (1317.07) |
| Log likelihood | −18,976.86 (−18,976.86) | −8430.71 (−8430.71) | −18,464.50 (−18,464.50) | −13,814.93 (−13,814.93) |
| Wald Chi sq | 3048.7 (3048.7) | 1047.58 (1047.58) | 1505.92 (1505.92) | 1317.07 (1317.07) |
| Prob > chi2 | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) |

* $p < 0.1$; *** $p < 0.05$, **** $p < 0.001$, robust standard errors clustered at store level in parentheses, other controls include dummy variables for store and month (not reported for ease of presentation). The dependent variable is number of sales of model $i$ in store $j$ in month $t$. All models are estimated in STATA version 13 ("nbreg" command). Store monthly sales are used for the exposure variable. The Wald statistic is for a likelihood ratio test of the joint significance of the regressors.
Acknowledgements Thanks are extended to staff and management at DID Electrical, the retail partner, for facilitation of the randomised controlled trial. Thanks also to Declan O’Sullivan, Darragh Blake, and Joris Vreek for their inputs on the development of the QR codes and the interactive website. Thanks also to Richard Roberts, Natasha Brown, the team of research assistants at TCD, Dr. James Carroll, and faculty at the Department of Economics at TCD for their valuable insights.

Funding Open Access funding provided by the IReL Consortium This research was funded by the Irish Research Council under the New Horizons programme, Research Project REPRO/2015/28 entitled NEEP.

Data Availability There are two data sets used in the analysis, one which contains sales data from the retail partner, and the second which contains the data from the customer survey. In the event that the paper is accepted for publication, anonymised and scaled (to protect commercially sensitive data) datasets can be provided.

Declarations

Ethical Approval The study received ethical approval from the Ethics Committee of the Faculty of Arts, Humanities, and Social Sciences at Trinity College Dublin, the University of Dublin.

Conflict of Interest The author declares no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abaluck, J., & Gruber, J. (2011). Choice inconsistencies among the elderly: Evidence from plan choice in the Medicare Part D program. American Economic Review, 101, 1180–1210.

Allcott, H., & Greenstone, M. (2012). Is there an energy efficiency gap? The Journal of Economic Perspectives, 26, 3–28.

Allcott, H., & Knittel, C. (2019). Are consumers poorly informed about fuel economy? Evidence from two experiments. American Economic Journal: Economic Policy, 11, 1–37.

Allcott, H., & Sweeney, R. L. (2016). The role of sales agents in information disclosure: Evidence from a field experiment. Management Science, 63, 21–39.

Allcott, H., & Taubinsky, D. (2015). Evaluating behaviourally motivated policy: Experimental evidence from the lightbulb market. American Economic Review, 105, 2501–2538.

Andor, M. A., Frondel, M., Gerster, A., & Sommer, S. (2019a). Cognitive reflection and the valuation of energy efficiency. Energy Economics, 84, 104527.

Andor, M. A., Gerster, A., & Götte, L. (2019b). How effective is the EU energy label? Evidence from a real-stakes experiment. Environmental Research Letters, 14, 044001.

Andor, M., Gerster, A., & Sommer, S. (2017). Consumer inattention, heuristic thinking and the role of energy labels (Ruh Economic Papers No. 671). RWI – Leibniz-Institut für Wirtschaftsforschung.

Andor, M. A., Gerster, A., & Sommer, S. (2020). Consumer inattention, heuristic thinking and the role of energy labels. Energy Journal, 41, 83–122.

Barber, B. M., Odean, T., & Zheng, L. (2005). Out of sight, out of mind: The effects of expenses on mutual fund flows. Journal of Business, 78, 2095–2119.

Bollinger, B., Leslie, P., & Sorensen, A. (2011). Calorie posting in chain restaurants. American Economic Journal: Economic Policy, 3, 91–128.

Brazil, W., Kallbekken, S., Saelen, H., & Carroll, J. 2019. The role of fuel cost information in new car sales. Transportation Research Part D, 74, 93–103.
Brounen, D., Kok, N., & Quigley, J. M. (2013). Energy literacy, awareness, and conservation behavior of residential households. Energy Economics, 38, 42–50.

Carroll, J., Aravena, C., Boeri, M. & Denny, E. 2021. “Show me the energy costs”: Short and long-term energy cost disclosure effects on willingness-to-pay for residential energy efficiency. The Energy Journal, 43(2)

Carroll, J., Aravena, C., & Denny, E. (2016a). Low energy efficiency in rental properties: Asymmetric information or low willingness-to-pay? Energy Policy, 96, 617–629.

Carroll, J., Denny, E., & Lyons, S. (2016b). The effects of energy cost labelling on appliance purchasing decisions: Trial results from Ireland. Journal of Consumer Policy, 39, 23–40.

Ceolotto, S., & Denny, E. (2021). Putting a new ‘spin’ on energy labels: Measuring the impact of reframing energy efficiency on tumble dryer choices in a multi-country experiment (TCD working paper series). https://www.tcd.ie/Economics/research/papers/. Accessed 10 Nov 2021

Christensen, G., & Miguel, E. (2018). Transparency, reproducibility, and the credibility of economics research. Journal of Economic Literature, 56, 920–980.

Codagnone, C., Veltri, G. A., Bogliacino, F., Lupiáñez-Villanueva, F., Gaskell, G., Ivchenko, A., Orteleva, P., & Mureddu, F. (2016). Labels as nudges? An experimental study of car eco-labels. Economia Politica, 33, 403–432.

CONSEED. (2019). Report on field trial evidence on the effectiveness of providing information on energy costs on energy related decisions in households. European Commission.

CSO. (2016). National Census. Central Statistics Office. https://www.cso.ie/en/census. Accessed 7 Sep 2016

Davis, L. W., & Metcalf, G. E. (2016). Does better information lead to better choices? Evidence from energy-efficiency labels. Journal of the Association of Environmental and Resource Economists, 3, 589–629.

DECC. (2014). Evaluation of the DECC/John Lewis energy labelling trial. Department of Energy & Climate Change UK.

Del Mar Solá, M., De Ayala, A., Foudi, S., & Galarraga, I. (2017). Understanding consumer decision making in the context of energy efficiency. CONSEED Deliverable 1.1. https://www.conseedproject.eu/wp1. Accessed 10 Nov 2021

Del Mar Solá, M., De Ayala, A., & Galarraga, I. (2021). The effect of providing monetary information on energy savings for household appliances: A field trial in Spain. Journal of Consumer Policy, 44, 279–310

Del Mar Solá, M., De Ayala, A., Galarraga, I., & Escapa, M. (2020). Promoting energy efficiency at household level: A literature review. Energy Efficiency, 14, 1–22.

European Commission. (2009). Accompanying document to the proposal for a commission regulation implementing Directive 2005/32/EC with regard to household refrigerating appliances. Full impact assessment, Document 52019SC0341, SWD/2019/0341 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52019SC0341. Accessed 7 May 2022

European Commission. (2010). Commission Regulation (EC) No. 643/2009 with regard to ecodesign requirements for household refrigeration appliances and Commission Delegated Regulation (EU) No. 1060/2010 with regard to energy labelling of household refrigeration appliances, Official Journal of the European Union L315/187.

EUROSTAT. (2021). Electricity price statistics. Available from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Electricity_price_statistics [Online]. Accessed 4 Nov 2021

Frederiks, E. R., Stenner, K., & Hobman, E. V. (2015). Household energy use: Applying behavioural economics to understand consumer decision-making and behaviour. Renewable and Sustainable Energy Reviews, 41, 1385–1394.

Gerarden, T., Newell, R., & Stavins, R. N. (2015). Deconstructing the energy-efficiency gap: Conceptual framework and evidence. American Economic Review, 105, 183–186.

Heinze, S. (2012). Disclosure of energy operating cost information: A silver bullet for overcoming the energy-efficiency gap? Journal of Consumer Policy, 35, 43–64.

Houde, S. (2014). How consumers respond to environmental certification and the value of energy information (NBER Working paper). https://doi.org/10.3386/w20019.20019. Accessed 7 May 2022

Houde, S., & Myers, E. (2021a). Are consumers attentive to local energy costs? Evidence from the appliance market. Journal of Public Economics, 201, 104480.

Houde, S., & Myers, E. (2021b). Heterogeneous (mis-) perceptions of energy costs: Implications for measurement and policy design (NBER Working paper, No. 25722).

Houde, S., & Wekhof, T. (2021). The narrative of the energy efficiency gap (ETHZ working paper, 21/359). https://doi.org/10.3929/ethz-b-000495755. Accessed 7 May 2022

IEA. (2007). Mind the Gap. OECD Publishing Paris.

Ioannidis, J. P. (2005). Why most published research findings are false. PLoS Medicine, 2, 0696–0701.
Ioannidis, J. P., Stanley, T. D., & Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal, 127*, F236–F265.

Jaffe, A., Newell, R., & Stavins, R. (2004). Economics of energy efficiency. *Encyclopedia of Energy, 2*, 79–90.

Jaffe, A. B., & Stavins, R. N. (1994). The energy-efficiency gap What does it mean? *Energy Policy, 22*, 804–810.

Kallbekken, S., Saelen, H., & Hermansen, E. T. (2013). Bridging the energy efficiency gap: A field experiment on lifetime energy costs and household appliances. *Journal of Consumer Policy, 36*, 1–16.

Labandeira, X., Labeaga, J. M., & López-Otero, X. (2012). Estimation of elasticity price of electricity with incomplete information. *Energy Economics, 34*, 627–633.

Linares, P., & Labandeira, X. (2010). Energy efficiency: Economics and policy. *Journal of Economic Surveys, 24*, 573–592.

Min, J., Azevedo, I. L., Michalek, J., & De Bruin, W. B. (2014). Labeling energy cost on light bulbs lowers implicit discount rates. *Ecological Economics, 97*, 42–50.

Phillips, Y. (2012). Landlords versus tenants: Information asymmetry and mismatched preferences for home energy efficiency. *Energy Policy, 45*, 112–121.

Sammer, K., & Wüstenhagen, R. (2006). The influence of eco-labelling on consumer behaviour – Results of a discrete choice analysis for washing machines. *Business Strategy and the Environment, 15*, 185–199.

Skourtos, M., Damigos, D., Tourkolias, C., & Kontogianni, A. (2021). Efficient energy labelling: The impact of information content and style on product choice. *Energy Efficiency, 14*, 1–19.

Sovacool, B. K., & Blyth, P. L. (2015). Energy and environmental attitudes in the green state of Denmark: Implications for energy democracy, low carbon transitions, and energy literacy. *Environmental Science & Policy, 54*, 304–315.

Spybrook, J., Bloom, H., Congdon, R., Hild, C., Martineze, A., & Raudenbush, S. (2011). *Optimal design plus empirical evidence*. Documentation for the “Optimal Design” software. http://wagrantfoundation.org/library/uploads/2015/11/OD-Documentation-V3.pdf. Accessed 5 June 2016

Stadelmann, M., & Schubert, R. (2018). How do different designs of energy labels influence purchases of household appliances? A field study in Switzerland. *Ecological Economics, 144*, 112–123.

STATISTICA. (2021). Average retail electricity prices in the United States from 1990 to 2020. https://www.statista.com/statistics/183700/us-average-retail-electricity-price-since-1990/20%5bOnline%5d. Accessed 4 Nov 2021

Wichman, C. J. (2017). Information provision and consumer behavior: A natural experiment in billing frequency. *Journal of Public Economics, 152*, 13–33.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.