Working Paper Sustainability and Innovation
No. S 11/2020

Corinne Faure
Marie-Charlotte Guetlein
Joachim Schleich

Effects of rescaling the EU energy label on household preferences for top-rated appliances

A slightly revised version of this paper has been published:

Faure, C.; Guetlein, M.-C.; Schleich, J. (2021): Effects of rescaling the EU energy label on household preferences for top-rated appliances. In: Energy Policy, 156, 112439. doi: 10.1016/j.enpol.2021.112439

Fraunhofer ISI

CHEETAH

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 723716.

Abstract

The European Union has decided to replace its current A+++ to D labelling scheme for cold appliances with a rescaled A to G labelling scheme. Employing a demographically representative discrete choice experiment on refrigerator adoption using an online survey among more than 1000 households in Germany, this paper explores the effects of the rescaled scheme compared to the old scheme on the stated uptake of top-rated refrigerators. Since in practice both schemes will be shown for a transitory period, the paper also analyses the effects of displaying both labels simultaneously. The findings from estimating a mixed logit model suggest that showing the rescaled A to G label alone significantly increases valuation of top-rated refrigerators compared to showing the current A+++ to D label alone. In comparison, when the A+++ to D and the rescaled A to G schemes are shown simultaneously, no benefits of introducing the rescaled label are found. Thus, policymakers should strive to enforce the application of the rescaled label scheme as quickly as possible and to shorten transitory periods where both labels are shown simultaneously.

Key words: energy efficiency; energy label; appliances; choice experiment.

Highlights:

- Demographically representative choice experiment in Germany on rescaled EU energy label.
- Rescaled A to G label increases willingness-to-pay for top-rated refrigerators.
- Showing rescaled label in addition to old A+++ to D label has no additional effects.
- Coexistence of rescaled A to G and old A+++ to D label should be minimized.

Table of Contents Page

1 Introduction

To enable customers to make informed product choices based on their energy use, many countries have introduced energy labelling schemes as a key policy to achieve energy and climate policy targets. For the European Union (EU), the "Labeling Directive" (European Economic Community, 1992) mandates manufacturers and retail stores the use of comparative energy labels for household appliances. The initial labelling schemes included seven energy efficiency classes visualized by horizontal bars of different colors and length ranging from the green class-A label for the appliances with the best energy performance to the red class-G label for appliances with the worst energy performance. To account for technological progress in energy efficiency, the EU subsequently introduced additional classes A+, A++ and A+++ for some appliances, including refrigerating appliances (European Commission, 2003; European Union, 2010) and in 2011 eventually removed the lowest classes E to G from the label for these appliances to maintain seven energy efficiency classes on the label. As a result, for these appliances, from 2011 onwards, the three efficiency classes A+, A++ and A+++ were separately depicted with different shades of green on the label, while Class A labels received a yellow color code. Further, a tightening of minimum energy performance standards (MEPS) in 2010 (e.g., European Commission, 2009) led to refrigerating appliances with an energy efficiency class lower than A+ no longer being able to be sold in the EU market after 2014. Therefore, since 2014, all refrigerators and freezers on the market have had a green color code; however, the label still includes horizontal bars for energy classes that no longer meet the MEPS requirements (Class A to D).

This updated A+++ to D EU energy-labelling scheme of 2011 (in combination with a tightening of the MEPS in 2010) is generally believed to have increased the market share of top-rated appliances (e.g., Bertoldi et al., 2016; Bjerregaard and Møller, 2019; Schleich et al., 2020). However, the scaling policy may have reduced the effectiveness of the label, because consumers may have erroneously interpreted an A+ label as a top energy efficiency rating (Heinzle and Wüstenhagen, 2012). In response to these concerns, the European Union (European Union, 2017) decided to reintroduce the initial classification using only the letters A to G. This new labelling scale involves rescaling of appliances so that an appliance classified A+++ (A++, A+ respectively) could for instance be classified B (E, G, respectively). For cold appliances (refrigerators and freezers), this rescaled scheme applies from 1 November 2020 onwards (European

Union, 2019). In practice, during a transition period, appliances may be simultaneously labeled with both the existing A+++ to D and the rescaled A to G label.

In this paper we analyze the effects of the rescaled A to G labelling scheme on household preferences for top-rated refrigerators compared to the A+++ to D scheme. We further explore the effects of simultaneously using both labels on preferences, since it is unclear whether such simultaneous labelling would serve to help consumers make a decision, or unintentionally confuse them. To do so, we employ an online stated preferences discrete choice experiment (DCE) on refrigerator adoption using a demographically representative sample of 1099 households in Germany.

Several studies have previously relied on stated preferences experiments through online surveys to study hypothetical adoption of household appliances, typically finding participants to respond positively (i.e., purchase more efficient appliances) to the information provided on energy labels, or to certification by an energy label. For the US EnergyGuide labels, these studies include Davis and Metcalf (2016) for air conditioners, Newell and Siikamäki (2014) for water heaters, Ward et al. (2011) and Li et al. (2016) for refrigerators. For labels in China, similar results were obtained by Shen and Saijo (2009) for air conditioners and refrigerators, by Zha et al. (2020) for refrigerators and washing machines, and by Zhou and Bukenya (2016) for air conditioners, and in Malaysia for refrigerator-freezers by Saidur et al. (2005). In addition, a few studies have analyzed moderating factors of the EU label effectiveness. For Germany, Andor et al. (2019) find participants with low cognitive ability to be particularly responsive to the energy label. Also for Germany, Andor et al. (2020) conclude that adding information about annual energy costs (in addition to kWh displayed) to the EU energy label increases the uptake of energy-efficient refrigerators. Relying on identical surveys in eight EU countries, Guetlein et al. (2019) find participants with high energy literacy to more strongly value appliances with a top energy label.

Our study is closest to Heinzle and Wüstenhagen (2012), who find, for samples of 90 and 97 persons by treatment, that - prior to its actual implementation - adding the categories A+, A++, and A+++ to the existing A to G labelling scheme in 2011 lowered the importance of energy efficiency in the hypothetical adoption of TVs by households in Germany. This finding may have possibly been partially driven by participants being familiar with the A to G labelling scheme, but not with the (at the time new) A+++ to D labelling scheme. The current rescaling of the label may not have this problem, since consumers will likely

be familiar with the rescaled label (identical to the label used before 2011 and to labels used for other product categories such as ovens, tumble dryers or air conditioners). Our results will therefore not confound effects that are due to the superiority of the A to G label with those due to familiarity with the label since we focus on the novel situation of moving back to an old version of the label. Further, in contrast to Heinzle and Wüstenhagen (2012), we focus on refrigerators and use larger samples of about 400 persons per treatment group. Finally, our study is the first to explicitly test the effects of a simultaneous usage of old and rescaled labels.

We organize the remainder of the paper as follows. Section 2 describes the methodology including the survey, the DCEs for the various labelling schemes, and the econometric model. Section 3 presents the results. Finally, the concluding Section 4 discusses the findings and possible policy implications.

2 Methodology

2.1 Survey

We tested the effects of introducing the rescaled label through an online survey conducted in Summer 2018 in Germany as part of a larger study on household preferences for energy efficient technologies and policies. Participants were members of NORSTAT's online household panel and received a participation fee for completing the survey. They were selected via quota sampling to be representative of the German adult population (between 18 and 65 years of age) on the criteria of gender, income, regional population distribution, and age. After some initial screening questions on the quota criteria and before answering questions on individual and household characteristics, respondents were randomly assigned to participate in two successive DCEs on different technologies or policies, including the DCE on labelling schemes described in detail in section 2.2.

Out of 1178 panel members who participated in the DCE on labelling schemes, 1099 (93%) completed the entire questionnaire. Incomplete responses were excluded from the analysis. The median time that respondents took to complete the survey including both the DCEs was 19 minutes. Table A1 in the Appendix presents the descriptive statistics for our sample and treatment groups together with the descriptive statistics. As we will detail in section 2.2., the treatment groups differed depending on whether they saw the existing A+++ to D label, the rescaled A to G label, or both labels. There are slightly fewer women than

men in the second group compared to the other two groups. The difference is significant (Pearson's chi-squared test; p = 0.091). We therefore test whether gender has an effect on results presented in Section 3 but do not find any significant gender effects. Otherwise, no notable differences are observed across groups.

2.2 Discrete choice experiment

The following framing was used to introduce the choice experiment (see also Appendix Figure A1):

"In this part of the survey, we invite you to make a series of hypothetical ***choices between different refrigerators****. There are no right or wrong answers to these questions.*

Imagine that ***your refrigerator has broken down and you need to buy a new one.*** *On the following pages, we will show you different refrigerator purchase options. We would like to know* ***which refrigerator you would choose, if these were your only options.***
Please assume that all refrigerator options fit properly in your kitchen and are currently available in colour and finish of your choice."[1]

This framing, adapted from Ward et al. (2011) and Li et al. (2016), was chosen to mitigate hypothetical bias since participants are told that they have to project themselves in a situation where their existing refrigerator had broken down (a realistic possibility that justifies choosing a new system).

Following the framing, participants read a detailed description of the attributes on which these refrigerators differed: size, energy class, warranty, customer ratings, and purchase price (see Table 1 for the attributes and attribute levels and Table A2 in the Appendix for an example of a choice card shown to participants). The levels for refrigerator size and price were chosen to represent the variety of refrigerator size and price options available on the market in Germany at the time of the study. We did not include extreme options such as mini-refrigerators or luxury price refrigerators to ensure that the options proposed were realistic and could be seriously considered by the majority of consumers. Based on previous literature using DCEs on refrigerators and on information

1 Instructions in 'bold ' were also seen in 'bold' by participants.

available on websites selling refrigerators, we also included warranty and customer ratings as attributes that are relevant for refrigerator purchase decisions, and that are independent from the other attributes included. Length of warranty was chosen to vary from 2 to 6 years. Customer ratings have been shown to have a great impact on purchase decisions (Chevalier and Mayzlin, 2006; Moe and Trusov, 2011); in our DCE, they were described through the typical five-star representation used in many online stores, with three possible levels, ranging from 3.5 stars to 4.5 stars.

To explore the effects of different label schemes, we randomly assigned participants to three separate groups differing by the type of energy label shown: (1) A+++ to D label (specifically, choosing between refrigerators labeled A+++, A++ or A+)[2], (2) rescaled A to G label (choosing between refrigerators labeled B, E, or G), or (3) simultaneous use of both labels (i.e., choosing between refrigerators labeled A+ under the A+++ to D label and G under the rescaled one, or A++ and E, or A+++ and B). In the first two groups, participants were told that "Refrigerators come with a label that looks like the following", which was then followed by a stylized picture of the label shown (see Figure 1 for the pictures). For the third group, participants were told that "Refrigerators come with two labels, the current label, and a new label that is replacing the current label." And were shown both labels named as "current label" and "new label". For all respondents, the picture of the label(s) was followed by a sentence explaining that "The colour **"green"** indicates a lower energy consumption while the colour **"red"** indicates a higher energy consumption compared to refrigerators with the same volume and features."[3] Finally, participants were informed that they would choose among refrigerators with energy class A+++, A++, or A+ (for groups 1 and 3) or B, E, and G (Group 2). Participants in Group 3 also were informed that the energy classes A+++, A++, or A+ correspond to classes B, E, and G according to the new label (all energy classes were bolded in text). The equivalence from one label system to the next was established based upon information provided on the European Commission website about the rescaled label and to represent the range of energy classes that will be available on the market upon introduction of the rescaled label.

---

2 Following a tightening of the MEPS (European Commission, 2009), only refrigerators labelled A+ or better were allowed to be sold on the EU market.

3 The words "**"green"** indicates a lower energy consumption" were color-coded in green and "**"red"** indicates a higher energy consumption" color-coded in red in the survey.

The equivalence from one label system to the next is not obvious and depends on various product characteristics. In particular, the Energy Efficiency Index (EEI) based on which energy classes are determined is calculated differently under the current and the rescaled labelling schemes (see Commission Delegated Regulation (EU) 1060/2010 and Commission Delegated Regulation (EU) 2019/2016.) Thus, appliances with the same energy efficiency class under the current labelling scheme might be labelled differently under the rescaled labelling scheme. At the same time, requirements for the rescaled label are such that the currently most efficient appliances (as of 2020) will be rated B at best (i.e., class A will remain empty at first). The least efficient appliances currently available on the market – including some A+-labelled appliances – will fall into the lowest category, G. As a consequence, we decided to translate all A+++ appliances into B appliances, A++ into E, and A+ into G.

Figure 1: Stylized pictures of the labels as shown to participants

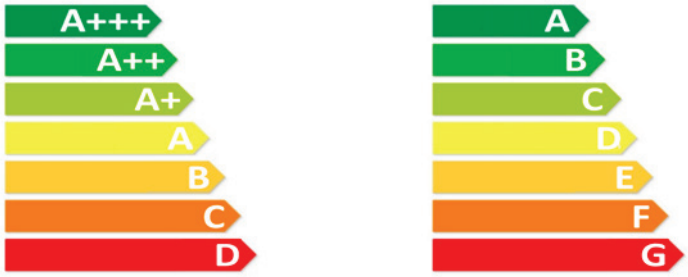

Table 1: Levels of different attributes considered in the refrigerator choice experiment

Attribute	Levels
Size	220 L, 240 L, 260 L, 280 L, 300 L, 320 L
Energy class	Group 'A+++ to D label': A^{+}, A^{++}, A^{+++} / Group 'rescaled label': G, E, B / Group 'both labels': A^{+}, A^{++}, A^{+++} and G, E, B
Warranty	2 years; 4 years; 6 years
Customer rating	3.5/5 stars; 4.0/5 stars; 4.5/5 stars
Purchase price	250 €, 350 €, 450 €, 550 €, 700 €, 850 €

Participants were then asked to respond to 6 scenarios, each consisting of a choice between two refrigerator purchase options. Because respondents were forced to choose a refrigerator among the alternatives proposed (to replace a broken one), this DCE did not include the status quo option of keeping the cur-

rent refrigerator). The scenarios were obtained through the application of a Bayesian efficient design (Sándor and Wedel, 2001) using the NGENE software (ChoiceMetrics, 2014). Bayesian efficient designs assume orthogonal attributes and are used to reduce the large number of possible attribute level combinations to an efficient design based on prior information. In this study, the priors were based on a separate pretest conducted in the UK with 50 respondents.

2.3 Econometric model

We use a mixed logit model to analyze the data from the DCE. Unlike conditional logit models, mixed logit models allow coefficients to vary across individuals and do not rely on the Independence from Irrelevant Alternatives (IIA) assumption (Revelt and Train, 1998). In a panel setting where n denotes individuals, t choice situations and j choice alternatives, the utility function can be expressed as

$$U_{njt} = \beta'_n X_{njt} + \varepsilon_{njt}, \qquad n = 1, \dots, N, \quad j = 1, \dots, \mathrm{J} \quad t = 1, \dots, T \tag{1}$$

X_{njt} is a vector of attributes of the alternative and β_n a vector of random parameters that varies across participants. β_n can be characterized by a density function $f(\beta_n|\theta)$ with a vector of parameters θ (Train, 2003). Finally, ε_{njt} denotes the error term. Our DCE consists of 12 choice situations divided into two blocks and with two choice alternatives each, hence T=12 and J=2.

Assuming that ε_{njt} is distributed iid extreme value, the probability to observe participant n chose a sequence of alternatives $s = (j_1, j_2, \dots, j_T)$ is given by

$$S_n(\theta) = \int \prod_{t=1}^{T} \frac{\exp(\beta'_n X_{nit})}{\sum_{j=1}^{J} \exp(\beta'_n X_{njt})} f(\beta_n|\theta) d\beta_n \tag{2}$$

The log likelihood function can be written as:

$$LL(\theta) = \sum_{n=1}^{N} \ln \mathrm{S}_n(\theta) \tag{3}$$

Equation (2) cannot be solved analytically. Instead, the probability is approximated through simulations. Let $P_n(\beta_n) = \prod_{t=1}^{T} \frac{\exp(\beta'_n X_{nit})}{\sum_{j=1}^{J} \exp(\beta'_n X_{njt})}$ denote the conditional probability that participant n chooses a sequence of alternatives for a

known β_n. We obtain the simulated log likelihood by running a simulation with *R* Halton draws (Train, 2003):

$$SLL(\theta) = \sum_{n=1}^{N} \ln\left\{\frac{1}{R}\sum_{1}^{R} \mathrm{P}_n(\beta^r)\right\} \tag{4}$$

where β^r is the r[th] draw from $f(\beta|\theta)$. We used *R* = 250.

The marginal WTP for an attribute x may then be estimated as.

$$\widehat{WTP_x} = -\frac{\hat{\beta}_x}{\hat{\beta}_p} \tag{5}$$

where $\hat{\beta}_x$ is the estimated random parameter associated with attribute x, and $\hat{\beta}_p$ is the estimated price parameter.

To allow testing for differences in labelling schemes, we use the observations from all three groups and include interaction terms reflecting the various labelling schemes for the label attributes.[4] The utility function may then be written as:

$$U_{njt} = \beta_1 price + \beta_{n,2} size + \beta_{n.3} A2 + (\beta_4 rescaled + \beta_5 both) \times A2 + \tag{6}$$

$$\beta_{n,6} A3 + (\beta_7 rescaled + \beta_8 both) \times A3 + \beta_{n,9} warranty + \beta_{n,10} star4 +$$

$$\beta_{n,11} star45 + \varepsilon_{njt}$$

A2 and A3 are dummy variables indicating an A++- or an A+++-labelled alternative, respectively. Thus, we use energy class A+ as a baseline. Similarly, a rating of 3.5 is used as the baseline for the customer rating scheme. *rescaled* and *both* are dummy variables indicating whether participants saw the rescaled energy labelling scheme or both energy labelling schemes simultaneously. For participants who only saw the current label, *rescaled* and *both* are equal to zero. The parameter of the price attribute and all interaction terms in equation (6) are specified as fixed parameters. All other parameters are modelled as random parameters and are assumed to be normally distributed as is standard in the literature.

4 We note that estimating three models separately would not produce correct standard errors to test for differences in the effectiveness of the label schemes.

3 Results

Table 2 presents the results of the mixed logit model. The upper part reports the means of the parameter estimates while the lower part shows their standard deviations. Half the standard deviations of the parameter estimates are statistically significant, implying heterogeneity of these parameters across respondents and supporting the use of a mixed logit model rather than a conditional logit model.

Turning first to the preferences for the attributes, as expected, the parameter estimate associated with *price* is negative and statistically significant. A higher purchasing price lowers respondents' willingness to select a particular refrigerator.

Turning next to the parameters for energy class, we find the coefficients associated with A2 and A3 to be positive and statistically significant. Thus, for the old energy class label, respondents value refrigerators with better energy classes than A+. In particular, they are willing to pay about 116 Euros more for a refrigerator labelled as A++ and about 331 Euros more for a refrigerator labelled as A+++ than for a refrigerator labeled as A+ (see Table 3). The findings for the interaction terms suggest that under the rescaled label, respondents value refrigerators with the highest energy class more than under the old label (Wald test, $\chi = 32.63$, $p<0.01$), with a difference in WTP amounting to about 233 Euros. There appears to be no difference between valuations of the highest energy class for respondents seeing the old label or both old and rescaled labels simultaneously.

Table 2: Results for mixed logit model†

	Pooled sample
Means of parameter estimates	
price	-0.005***
	(0.000)
size	0.003***
	(0.000)
warranty	0.142***
	(0.000)
star4	0.202***
	(0.000)
star45	0.543***
	(0.000)
A2	0.557***
	(0.000)
A3	1.580***
	(0.000)
A2_rescaled	0.170
	(0.311)
A3_rescaled	1.114***
	(0.000)
A2_both	-0.061
	(0.713)
A3_both	-0.002
	(0.992)
Standard deviations of parameter estimates	
size	-0.000
	(0.919)
warranty	-0.030
	(0.760)

	Pooled sample
star4	0.023
	(0.805)
star45	-1.175***
	(0.000)
A2	1.503***
	(0.000)
A3	2.000***
	(0.000)
Log likelihood	-3770.468
Number of participants	1099
Number of observations	13188

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; †p-values are shown in parentheses

Moreover, we find no evidence that the rescaled label or showing both labels has any effect on participants' valuation of refrigerators with the second highest energy class.[5] Therefore, Table 3 reports the same marginal willingness to pay for A2 for all groups.

We now turn to the findings for the remaining attributes. The mean of the parameter estimate of *size* is statistically significant and positive, implying that on average respondents prefer larger refrigerators to smaller ones. Table 3 suggest that respondents are willing to pay 0.62 Euros for an additional liter in volume. Similarly, respondents prefer refrigerators with longer warranties; on average, they are willing to pay about 30 Euros for an additional year of warranty. Next, the findings for the coefficients associated with *star4* and *star45* imply that respondents value refrigerators with higher customer ratings compared to a 3-star baseline rating. More specifically, on average respondents are willing to

5 Re-estimating the model in equation (6) with the rescaled label as baseline confirms that showing both labels instead of only the new label has no effect on the valuation of the second highest energy class (current A++). Results from the model with the rescaled label as baseline are shown in Table A2 in the Appendix.

pay an additional 42 and 114 Euros for a 4-star-rated and 4.5 star-rated refrigerator compared to a 3-star rated refrigerator, respectively.[6]

Table 3: Marginal willingness to pay estimates†

		Old label	Rescaled label	Both labels
size	0.62			
warranty	29.72			
star4	42.24			
star45	113.55			
A2†		116.46	116.46	116.46
A3†		330.63	563,64	330.63
Number of participants		367	365	367

†Only significant parameter estimates in Table 2 were used in calculating WTP estimates.

4 Conclusions and Policy Implications

The European Union has decided to replace its current A+++ to D labelling scheme for cold appliances by a rescaled A to G labelling scheme. The new rescaled label will be in place from 1 November 2020 on, but for a transitory period, both the old A+++ to D and the rescaled A to G label may be shown, presumably to allow customers to get used to the new label. Employing a demographically representative DCE on refrigerator adoption using an online survey among households in Germany, we explored the effects of the rescaled scheme compared to the existing scheme on the stated uptake of top-rated refrigerators. We also analyzed the effects of showing both label schemes simultaneously rather than separately.

Main findings

Our findings from estimating a mixed logit model suggest that overall, households prefer better classified refrigerators to worse classified ones. In addition, when the rescaled labels are shown alone, the rescaled A to G labelling scheme significantly increases the estimated WTP for top-rated refrigerators

6 Because we found women to be slightly underrepresented among respondents who saw the rescaled label, we estimated a mixed logit model with additional interaction terms between gender, energy classes and labelling schemes. We found no evidence that gender had an effect on the valuation of higher energy classes in any of the treatment groups.

compared to the previous scheme. In contrast, when the A+++ to D scheme and the rescaled A to G scheme are shown simultaneously, no benefits of introducing the rescaled label are found. Moreover, we find no evidence that showing the rescaled label instead of the current label has an effect on the valuation of the second-best energy class (current A++).

There are a few explanations for these results. First, and this was the main criticism of the introduction of the A+ to A+++ classes (see Heinzle and Wüstenhagen 2012), there is a greater perceived difference between B and G (or E) than between different shades of A+s, all of which appearing very good and only marginally differently so. The rescaling makes it apparent that appliances graded A+ under the current scheme are the worst on the market (a reality that was hidden under the old label scheme). Second, the rescaled label also implies that the differences between appliances are more pronounced on the scale. Even the top rated appliances are only rated B, whereas the lowest-rated appliances are rated G. Therefore, whereas under the old label there were only two classes difference between an A+++ and an A+ appliance, under the rescaled label, there are five classes difference between the best and the worst appliances. Third and relatedly, as a consequence of this broadening of of the scale, only top-rated appliances are in the green zone (instead of all of them being in the green zone under the old label), with former A++ appliances moving to the yellow zone and former A+ appliances to the red zone. Previous research has shown the impact of labels' color coding and especially of so-called traffic light color-coding on consumer decisions (e.g., Thøgersen and Nielsen, 2016; Tourangeau et al., 2007) and this new color classification might also explain why only top-rated appliances benefit from the rescaled label.

Our results are consistent with those from Heinzle and Wüstenhagen (2012) who had found that the introduction of the A+++ to D label might lead to confusion and might have detrimental effects on the adoption of top rated appliances. Similarly, our results show a superiority of the A to G label over the current A+++ to D label. Because the rescaled label is already familiar to EU consumers, it is unlikely that effects are due to lack of familiarity with either of the labels. Besides, if familiarity played a role, it would in this case play against the A to G label. As a consequence, our approach enables a more conservative test of the superiority of the A to G label compared to the A+++ to D label. Finally, we find that the simultaneous use of the A+++ to D label and the rescaled label eliminates the beneficial effects of introducing the rescaled label, possibly because the inconsistent information is dismissed.

Limitations

Hypothetical bias is a common concern for DCEs (e.g., Hensher (2010) because study participants are not in an actual purchase situation. This bias is particularly affecting the WTP estimates which are often inflated. Unfortunately, it was not practicable within this study to incentivize decisions (i.e., inform respondents that they might have to actually purchase the refrigerator chosen). Instead, we used a so called 'cheap talk' framing (e.g., Ward et al., 2011 or Li et al., 2016) that first reassures respondents about the fact that there are no right or wrong answers and most importantly frames the decision as a necessary replacement of a broken appliance. This is a realistic scenario in which respondents can easily project themselves, and that helps increase the realism of the task. Further, our focus in this study was not on the absolute WTP levels for various energy classes but rather on comparisons between these valuations across label scheme treatment groups. As a consequence, hypothetical bias should be less of a concern than it is when the focus in on WTP levels.

Implications for policy makers

Last but not least, our findings have clear policy implications. First, they suggest that the introduction of the rescaled A to G scheme will indeed increase the take-up of top-rated appliances and make the less efficient appliances on the market considerably less attractive. This is of course important because it speaks for the superiority of the A to G label and the appropriateness of returning to this label in the EU. Second, results on the simultaneous use of the existing and the rescaled label schemes indicate that regulation should either not allow a transition period where retailers may simultaneously use both label schemes on an appliance or at least reduce this transition period to a minimum. Our results therefore suggest that policy makers should make it a priority to enforce the application of the rescaled label scheme as quickly as possible.

Our findings also have some implications for the next steps in the implementation of the EU energy label regulations. The current regulation anticipates a regular scaling update (probably every 10 years) of the label classes (European Union, 2017) within the A to G label scheme to account for technological progress and for the fact that over time, more and more appliances will be developed for the top energy classes. A regular scaling update is therefore anticipated, implying that top appliances would at best be classified Class B and all other appliances would be rescaled in lower classes. For each of these scaling updates, the issue of simultaneous labelling of the appliances in the retail stores

will occur. Our findings suggest that such simultaneous labelling should be avoided or at least limited in time, which is consistent with the recommendations made in the regulation (European Union, 2017). Further, the fact that consumers discount the less favorable label information points to the necessity to support the updates through communication campaigns as planned in the regulation.

Acknowledgements

This research benefitted from funding by the European Union's Horizon 2020 Framework Programme under the project CHEETAH – CHanging Energy Efficient Technology Adoption in Households (Grant agreement ID: 723716). We would like to thank Gengyang Tu (University of Exeter, United Kingdom) for his contribution in designing the choice experiment, Heike Brugger (Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany) for her helpful comments on an earlier draft of the paper, and Antoine Durand and Barbara Schlomann (both of Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany) for their feedback and advice on technology-related matters.

References

Andor, M., Frondel, M., Gerster, A., Sommer, S., 2019. Cognitive reflection and the valuation of energy efficiency. Energy Economics 84, 10427.

Andor, M., Gerster, A., Sommer, S., 2020. Consumer Inattention, Heuristic Thinking and the Role of Energy Labels. The Energy Journal 41(1). https://doi.org/10.5547/01956574.41.1.mand

Bertoldi, P., Lorente, J., Labanca N., 2016. Energy Consumption and Energy Efficiency Trends in the EU-28 2000–2014. Joint Research Centre. EUR 27972 EN; doi 10.2788/581574

Bjerregaard, N., Møller, F., 2019. The impact of EU's energy labeling policy: An econometric analysis of increased transparency in the market for cold appliances in Denmark. Energy Policy 128, 891–899.

Chevalier, J. A., Mayzlin, D., 2006. The effect of word of mouth on sales: Online book reviews. Journal of Marketing Research 43(3), 345–354.

ChoiceMetrics, 2014. Ngene 1.1.2.: User manual and reference guide. The cutting edge in experimental design. Choice Metrics Pty Ltd.

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(3), 589–625.

European Economic Community, 1992. Council Directive 92/75/EEC of 22 September 1992 on the indication by labelling and standard product information of the consumption of energy and other resources by household appliances. OJ L 297, 13.10.1992, p. 16–19. Site: https://eur-lex.europa.eu/eli/dir/1992/75/oj

European Commission, 2003. Commission Directive 2003/66/EC of 3 July 2003 amending Directive 94/2/EC implementing Council Directive 92/75/EEC with regard to energy labelling of household electric refrigerators, freezers and their combinations (Text with EEA relevance), OJ L 170, 9.7.2003, p. 10–14. http://data.europa.eu/eli/dir/2003/66/oj

European Commission, 2009. Commission Regulation (EC) No 643/2009 of 22 July 2009 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for household refrigerating appliances (Text with EEA relevance). OJ L 191, 23.7.2009, p. 53–68. http://data.europa.eu/eli/reg/2009/643/oj

European Union, 2010. Commission Delegated Regulation (EU) No 1060/2010 of 28 September 2010 supplementing Directive 2010/30/EU of the European Parliament and of the Council with regard to energy labelling of household refrigerating appliances Text with EEA relevance OJ L 314, 30.11.2010, p. 17–46. http://data.europa.eu/eli/reg_del/2010/1060/oj

European Union, 2017. Regulation (EU) 2017/1369 of the European Parliament and of the Council of 4 July 2017 setting a framework for energy labelling and repealing Directive 2010/30/EU (Text with EEA relevance). OJ L 198, 28.7.2017, p. 1–23. http://data.europa.eu/eli/reg/2017/1369/oj

European Union, 2019. Commission Delegated Regulation (EU) 2019/2016 of 11 March 2019 supplementing Regulation (EU) 2017/1369 of the European Parliament and of the Council with regard to energy labelling of refrigerating appliances and repealing Commission Delegated Regulation (EU) No 1060/2010 (Text with EEA relevance). OJ L 315, 5.12.2019, p. 102–133. http://data.europa.eu/eli/reg_del/2019/2016/oj

Guetlein, M.-C., Faure, C., Tu, G., Schleich, J., 2019. The effects of energy literacy and household income on consumer choice of energy-efficient appliances – Insights from a multi-country discrete choice experiment and welfare analysis. Annual Meetings of the European Association of Environmental and Resource Economics, Manchester.

Heinzle, S. L., Wüstenhagen, R., 2012. Dynamic Adjustment of Eco-labeling Schemes and Consumer Choice–the Revision of the EU Energy Label as a Missed Opportunity? Business Strategy and the Environment 21(1), 60–70.

Hensher, D. A., 2010. Hypothetical bias, choice experiments and willingness to pay. Transportation Research Part B: Methodological 44(6), 735-752.Li, X., Clark, C. D., Jensen, K. L., Yen, S. T., 2016. The effect of mail-in utility rebates on willingness-to-pay for ENERGY STAR certified refrigerators. Environmental and Resource Economics 63(1), 1–23.

Moe, W. W., Trusov, M., 2011. The value of social dynamics in online product ratings forums. Journal of Marketing Research 48(3), 444–456.

Newell, R. G., Siikamäki, J., 2014. Nudging Energy Efficiency Behavior: The Role of Information Labels. Journal of the Association of Environmental and Resource Economists 1(4), 555–98.

Revelt, D., Train, K., 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. Review of Economics and Statistics 80 (4), 647–657.

Saidur, R., Masjuki, H.H., Mahlia, T.M.I., 2005. Labeling design effort for household refrigerator-freezers in Malaysia. Energy Policy 33(5), 611–618.

Sándor, Z., Wedel, M., 2001. Designing conjoint choice experiments using managers' prior beliefs. Journal of Marketing Research 38, 430–444.

Schleich, J., Durand, A., Brugger, H., 2020. How effective are EU minimum energy performance standards and energy labels for cold appliances? Fraunhofer ISI Working Paper Sustainability and Innovation. Forthcoming.

Shen J., Saijo, T., 2009. Does an energy efficiency label alter consumers' purchasing decisions? A latent class approach on a stated choice experiment in Shanghai. Journal of Environmental Management 90(11), 3561–3573.

Thøgersen, J., Nielsen, K. S., 2016. A better carbon footprint label. Journal of Cleaner Production 125, 86–94.

Tourangeau, R., Couper, M. P., Conrad, F., 2007. Color, Labels, and Interpretive Heuristics for Response Scales. Public Opinion Quarterly 71(1), 91–112, https://doi.org/10.1093/poq/nfl046

Train, K. E., 2003. Discrete choice methods with simulation. Cambridge: Cambridge University Press.

Ward, D. O., Clark, C. D., Jensen, K. L., Yen, S. T., Russell, C. S., 2011. Factors influencing willingness-to-pay for the ENERGY STAR® label. Energy Policy 39(3), 1450–1458.

Zha, D., Yang, G., Wang, W., Qag, Q, Zhou, D., 2020. Appliance energy labels and consumer heterogeneity: A latent class approach based on a discrete choice experiment in China. Energy Economics 90, 104839. https://doi.org/10.1016/j.eneco.2020.104839.

Zhou, H., Bukenya, J. O., 2016. Information inefficiency and willingness-to-pay for energy-efficient technology: A stated preference approach for China Energy Label. Energy Policy 91, 12–21.

Appendix

Figure A1: Framing of discrete choice experiment on refrigerators†

Introduction

Imagine that **your refrigerator has broken down and you need to buy a new one.** On the following pages, we will show you different refrigerator purchase options. We would like to know **which refrigerator you would choose, if these were your only options.**
Please assume that all refrigerator options fit properly in your kitchen and are currently available in colour and finish of your choice.

The refrigerators only differ on the following attributes:

1. Size: The total **internal space** of each refrigerator is 220, 240, 260, 280, 300, or 320 litres. 20 litres corresponds to one small compartment. The picture below shows a 320-litre and a 220-litre refrigerator.

2. Energy class: Refrigerators come with a label that looks like the following :

The colour **"green"** indicates a lower energy consumption while the colour **"red"** indicates a higher energy consumption compared to refrigerators with the same volume and features.
You will choose among refrigerators with energy class **B, E, or G**.

3. Warranty: The **warranty** for each refrigerator is 2, 4, or 6 years.

4. Customer rating: Ratings are provided by customers who have bought the same refrigerator. You may assume that the refrigerators you can choose from have **average ratings** of 2.5, 3.5, or 4.5 stars out of 5 stars.

5. Purchase price: Each **refrigerator costs 250€, 350€, 450€, 550€, 700€, or 850€**.

† Displayed in German to participants

Figure A2: Example of a scenario shown to participants in the refrigerator choice experiment†

† Displayed in German to participants

Table A1: Descriptive statistics

Variable	Sample	Group 1 (A+++ to D label)	Group 2 (rescaled label)	Group 3 (both labels)	National census
Median age†	45	44	45	45	46
Female (18–65)	51%	53%	46%	53%	49%
States of former East Germany	15%	14%	15%	17%	15%
Median net houshold income (in €)††	29,300	29,300	29,300	29,300	22,647
Number of parti-cipants	1099	367	365	367	

† The national median age is the median age of the entire population, based on census data.

†† The median net household income is calculated using the midpoints of a scale with 12 income categories. The national median net household income is the median equivalized net household income of the entire population in 2018.

Source: Eurostat

Table A2: Results for the mixed logit model with rescaled label as baseline†

	Pooled sample
Means of parameter estimates	
price	-0.005***
	(0.000)
size	0.003***
	(0.000)
warranty	0.145***
	(0.000)
star4	0.200***
	(0.000)
star45	0.563***
	(0.000)
A2	0.724***
	(0.000)
A3	2.736***
	(0.000)
A2_old_label	-0.166
	(0.329)
A3_old_label	-1.135***
	(0.000)
A2_both	-0.221
	(0.190)
A3_both	-1.141***
	(0.000)
Standard deviations of parameter estimates	
size	-0.000
	(0.797)
warranty	-0.063
	(0.275)

	Pooled sample
star4	0.024
	(0.810)
star45	-1.192***
	(0.000)
A2	1.523***
	(0.000)
A3	-2.016***
	(0.000)
Log likelihood	-3768.491
Number of participants	1099
Number of observations	13188

*$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; †p-values are shown in parentheses

Authors' affiliations

Joachim Schleich

Fraunhofer Institute for Systems and Innovation Research (Fraunhofer ISI), 76139 Karlsruhe, Germany

Grenoble Ecole de Management, 38000 Grenoble, France

Virginia Polytechnic Institute & State University, Blacksburg, VA 24061, USA

Corinne Faure

Grenoble Ecole de Management, 38000 Grenoble, France

Marie-Charlotte Guetlein

Grenoble Ecole de Management, 38000 Grenoble, France

Contact: Joachim Schleich

Fraunhofer Institute for Systems
and Innovation Research (Fraunhofer ISI)
Breslauer Strasse 48
76139 Karlsruhe
Germany
Phone: +49 721 6809-203
E-Mail: joachim.schleich@isi.fraunhofer.de
www.isi.fraunhofer.de

Karlsruhe 2021