Financing the Economic Transition — On the Role of Labels for Retail Investors*

Markus Dertwinkel-Kalt⁺

Astrid Eller[‡] Christoph Feldhaus[§] Andreas Löschel[∥] Leon Landes[¶]

First Version: December 30, 2024 Current Version: April 3, 2025

Abstract

The economic transition from a fossil-fueled economy to a green economy requires substantial financing, including contributions from retail investors. What determines retail investors' willingness to pay (WTP) for sustainable investments is, however, an open question. In this paper, we address this question by examining the role of information and its presentation—particularly through labels—for retail investors' WTP for sustainable investments. Using an incentivized experiment with 1,219 German retail investors, we analyze how different presentations of information on environmental impact—ranging from plain text to labels with scores from one to five, under both loose and strict labeling standards impact WTP. Our findings reveal that while retail investors exhibit a significant WTP for environmental impact, this WTP is only weakly influenced by the actual extent of this impact. Labels play a critical role: more demanding label standards enhance the sensitivity of WTP to environmental impact, whereas more lenient standards diminish this sensitivity. Moreover, warm-glow utility emerges as a key driver of the WTP for environmental impact. Higher label scores increase warm-glow utility even in the absence of changes in actual impact. Finally, the effect of warm-glow utility on WTP is stronger in label treatments.

JEL-Classification: C91, D64, G11, G41

Keywords: Retail Investors, ESG, Warm Glow, Scope Insensitivity

^{*}We thank Vincent Eulenberg, Maximilian Schmetzke, Anna Schulze-Tilling and Anna Ressi for their helpful comments. Also, we thank the German Federal Ministry of Education and Research (BMBF), Funding Line Klimaschutz und Finanzwirtschaft (KlimFi) (FKZ: 01LA2202B) as well as the German Research Foundation (project 462020252) for financial support. Ethical approval for this study was given by the German Association for Experimental Economic Research. Details about the ethical approval are available here: https://gfew.de/ethik/FoU1diBD.

[†]University of Münster & Max Planck Institute for Research on Collective Goods. Email: markus.dertwinkelkalt@wiwi.uni-muenster.de

[‡]University of Münster. Email: aeller@uni-muenster.de

[§]Ruhr University Bochum. Email: christoph.feldhaus@ruhr-uni-bochum.de

^IRuhr University Bochum. Email: leon.landes@ruhr-uni-bochum.de

^{II}Ruhr University Bochum. Email: andreas.loeschel@ruhr-uni-bochum.de

1 Introduction

Interest in sustainable investments has grown rapidly over the past few years. Sustainable funds, in particular, have seen a remarkable surge in net inflows, with over \$10 billion flowing into such funds during the third quarter of 2024 alone (Bioy et al. 2024). This surge in interest is not only evident in market movements but also in the regulatory landscape. For example, EU legislation (EU 2019) lays out different criteria to classify financial products and disclosure requirements. Still, the current shift in financial flows is insufficient to reach the net zero goal by 2050 (Klaaßen and Steffen 2023).

Investors generally exhibit a preference to invest in sustainable funds (Hartzmark and Sussman 2019; Bauer et al. 2021) and are willing to accept lower returns for more sustainable investments (Barber et al. 2021). In this paper, we study whether changing the design of sustainability information (plain text or labels) influences the demand of retail investors for sustainable investments. Additionally, we analyze the underlying motivations that drive retail investors to choose sustainable investments, with a particular focus on emotional factors like the warm-glow utility linked to such investments. Investigating the effect of labels in the context of sustainability information to participants in text form. However, in real-world investment contexts, information about an investment's sustainability is provided to retail investors mainly through a label or score, such as the Morningstar Sustainability Rating (Morningstar 2021).

To analyze these questions, we conducted an incentivized field experiment among German retail investors in April and May 2024. In our experiment, each participant chose between two funds, one with a positive environmental impact and one without any impact. We varied the upfront fee for an investment of \notin 1,000, allowing us to measure participants' WTP for an investment in the sustainable fund. Each participant's WTP was elicited twice for different sustainable funds enabling within-subject analyses. Additionally, participants were randomly assigned to one of six treatment conditions in a 2×3 experimental design, also making between-subjects comparisons possible. One variation varied the size of the beneficial climate impact associated with the sustainable investment, that is, high or low impact range. The other dimension altered the presentation format, using either a strict label, a loose label, or plain text. Given the same environmental impact connected to the investment, the label with loose standards assigns a better label score than the label with the strict standard.

Our findings, first, indicate that retail investors generally exhibit a strong WTP for investments with a positive climate impact. Second, labels influence retail investors' WTP, with higher label scores leading to a higher WTP both within and across subjects. Although participants respond to changes in the magnitude of the positive impact of a fund, their reaction to these changes is rather small. Notably, a 10-fold increase in impact results, on average, only in a 20% increase in WTP. This is in line with findings from Heeb et al. (2023) who also show that retail investors' WTP for sustainable investments does not increase significantly when the sustainability level of an investment improves further. Furthermore, we show that labeling standards significantly affect investors' responses to changes in the impact: When impact information is presented via a label with strict standards, the average sensitivity of investors increases compared to treatments without a label. However, this sensitivity is reduced when the sustainability information is presented using a label with loose standards.

Also, retail investors' WTP responds to the underlying criteria of the labels, showing a preference for larger impact values, even when identical label scores are presented. This indicates that investors do not mindlessly follow the labels but pay at least some attention to the underlying criteria. One common concern about the introduction of environmental labels is their potential misuse, as they may give agents opportunities to misrepresent their environmental performance (Lyon and Montgomery 2015). Our analysis suggests that greenwashing is feasible only to a limited extent within our sample.

Focusing on the motivation for sustainable investment, we demonstrate that warm-glow utility is influenced by labels. Following Heeb et al. (2023), we elicit the positive emotions retail investors feel when investing in the sustainable answer with the question "*Compared to investing in the conventional fund, investing in the impact fund feels...*". We show that retail investors' warm-glow utility of a sustainable investment increases with a higher label score. This effect occurs irrespective of the investment's actual climate impact. Furthermore, we show that a higher warm-glow utility is strongly related to investors' WTP for sustainable investments. Finally, we find that the presentation format matters for the strength of the link between warm-glow utility and retail investors' WTP: If information is presented with a label instead of text, the impact of warm-glow utility on WTP is even stronger.

With our study, we add to several branches of literature. First, we add to the literature investigating retail investors' preferences for sustainable investments: For example, when given information about the ESG scores of different funds, retail investors invest a higher share of their portfolio in ESG-Funds (Gutsche, Wetzler, et al. 2023; Seifert et al. 2024). Furthermore, investors have a higher WTP for funds with a positive environmental impact (Heeb et al. 2023), and pro-social and environmental preferences are closely linked to sustainable investment (Riedl and Smeets 2017). We add to this literature by investigating whether retail investors' behavior changes if sustainable information is presented with labels instead of text, specifically, investors' response to changes in the level of sustainability and their warm-glow utility.¹

Second, we contribute to analyses of the effect of labels on sustainable behavior. The effect of labels on pro-environmental behavior has been documented in many contexts, such as food choices (Lohmann et al. 2022; Schulze Tilling 2024) and energy labels (Andor et al. 2020; Rodemeier and Löschel 2024).

Third, our findings inform a broad literature that analyzes the motivation for pro-social behavior. For example, in the context of investment behavior Bonnefon et al. (2025), distinguish between value-alignment and impact-seeking. Many models assume that decision makers are essentially consequentialists (e.g. Kaufmann et al. 2024), that is, the demand for sustainable investment depends on the actual positive impact connected to the investment. Others argue that sustainable behavior is related to the positive feeling personally experienced that is connected to these actions, called warm-glow utility (Andreoni 1990). Findings across various domains show that pro-social behavior is often motivated by this effect (Imas 2014; Hartmann et al. 2017). In line with the results of Heeb et al. (2023), we show that warm-glow utility strongly

¹To our knowledge, the only other paper investigating the effects of sustainability labels in an investment context is by Bassen et al. (2019). We are, however, the first to explicitly focus on labels with a focus on warm-glow utility and retail investors' sensitivity to the level of sustainability.

influences retail investors' WTP for sustainable investments. Furthermore, to our knowledge, we are the first to show that the effect of warm-glow utility on the WTP is especially strong when labels are present.

Finally, we add to the literature that focuses on scope insensitivity, a long-discussed topic in the context of contingent valuation methods (Lopes and Kipperberg 2020). More recently, several experimental papers have shown scope insensitive behavior, suggesting that the WTP for pro-social action does not necessarily scale (strongly) with the magnitude of pro-social benefits. For example, Pace et al. (2025) demonstrate that people's WTP for carbon mitigation is concave, and Rodemeier (2024) finds that, without learning experiences, people are completely inelastic to changes in the magnitude of carbon offsets. Regarding the behavior of retail investors, Heeb et al. (2023) report insensitivity. Once again, we further support these findings. Going beyond existing literature, we show that the degree of sensitivity changes with how sustainability information is displayed. Specifically, introducing a label can increase and decrease sensitivity compared to providing sustainable information in text form, depending on the label standard.

The remainder of this paper is organized as follows: Section 2 covers the experimental design and presents the related hypotheses. In Section 3, we present the results, and Section 4 concludes the paper.

2 Experimental Design

The central part of our pre-registered² study consists of measuring participants' WTPs for two funds with a positive climate impact and eliciting participants' warm-glow utility associated with each investment. Section 2.1 describes in detail how we elicit participants' WTP for sustainable investment. Afterward, we introduce in Section 2.2 the six different treatments in our experiment and explain the experimental procedure in Section 2.3. Based on the experimental set-up, we can analyze several hypotheses about the effects of labels on investor behavior, which we introduce in Section 2.4.

2.1 WTP Measurement

In the experiment, participants had to decide between an investment of $\notin 1000$ in one out of two different funds. One fund always had a positive environmental impact, while the other fund offered no environmental benefit. All other aspects of the two funds were identical and the two funds differed only in their climate impact. Participants faced two sets of seven decisions. Within each set, only the upfront fee for an investment into one of the funds changed while the other characteristics remained constant. Initially, the upfront fee for each fund was $\notin 10$, and the bisection method (Kuilen and Wakker 2011) was used to elicit participants' WTP for the sustainable fund. After the first decision, the upfront fee for the selected fund was increased by $\notin 40$. Depending on subsequent choices, the upfront fee was either increased or decreased by half of the previous change. After seven decisions, this approach allowed us to measure

²Details about the preregistration are available here: https://aspredicted.org/9LZ_BLZ. Note that, for clarity in presentation, some hypotheses in the preregistration correspond to differently numbered hypotheses in this paper.

participants' WTP within a range of €1.25. We define the WTP as the midpoint of this range. This approach of eliciting retail investors' WTP for a sustainable investment closely follows Heeb et al. (2023).

Participants who were censored by this design, that is, those who chose the same fund in each decision, were asked to state their WTP explicitly. Furthermore, all participants were asked if they agreed with the measured WTP. Those who disagreed were able to repeat the investment decisions once.

2.2 Treatments

Table 1 summarizes the study design. Using a 2x3 between-subjects design, participants are assigned to one of six treatment groups. One variation changes the size of the positive environmental impact connected to the sustainable investment while keeping the information display method constant. The other variation corresponds to the format in which information about the positive impact of the investment is displayed to participants, with the magnitude of climate impact remaining constant within each condition.

Regarding the variation in the size of positive environmental impact, there are two treatments: low impact and high impact. In the high impact conditions, participants' WTP is measured by comparing an investment into a fund with 0t of reduced CO_2 emissions against 5tof reduced CO_2 emissions. In the second investment decision, the WTP is measured for two funds with reduced CO_2 emissions of 0t and 0.5t, respectively. To avoid ordering effects, we randomize which investment decision participants make first. In the low impact conditions, these values are scaled down by a factor of 10. For each participant, we refer to the fund with the larger impact as Fund L and the fund with a smaller impact as Fund S.

We use three different treatments, changing the way in which the positive impact of the investment is presented to participants. In the no label conditions, information about the positive climate impact is provided solely in text form, closely following the study by Heeb et al. (2023). In the strict label conditions, information about the beneficial environmental impact is displayed via a label with relatively strict standards, scored as 3 out of 5 and 1 out of 5 green leaves, respectively. In the loose label conditions, the positive environmental impact is displayed with scores of 5 out of 5 and 3 out of 5, respectively. See Figure A1-A3 in the Appendix for investment decisions featuring the same underlying trade-offs but different label conditions.

Before participants made their first investment decisions, they were informed about the presentation format of environmental impact. In the strict label and loose label conditions, participants received details about the different levels of positive climate impact represented by each label score. We informed participants that this information could be accessed with a simple button click during all investment decisions. Participants were required to answer an attention check question about the label or size of climate impact to ensure they paid at least some attention to the climate impact of the investments.

	High Impact	Low Impact
No Label	$0t CO_2$ Fund S: 0.5t CO ₂	0t CO_2 Fund S: 0.05t CO_2
INU Label	$0t \operatorname{CO}_2$ Fund L: $5t \operatorname{CO}_2$	$0t CO_2$ Fund L: $0.5t CO_2$
	adada d adad	addd dddd
Strict Label	adda dad ad	99999 999 99
Looso Lobol	addad dad ad	adddd ddd dd
Loose Label	addad adda d	addda dddda

Table 1: Treatments

2.3 Experimental Procedure

We conducted the framed field experiment with 1,219 participants from Germany.³ The survey was carried out in April and May 2024 in cooperation with the survey company Psyma Research+consulting GmbH. We coded the survey using oTree (Chen et al. 2016). Psyma was primarily responsible for recruiting the participants and managing the financial and environmental incentives for the experiment. This involved making investments into a fund and donations to an offset organization based on the decisions of selected participants.

In Figure 1, the experimental process is illustrated. Participants were invited to enter the online survey with a link. As the main focus of the study is the behavior of retail investors, participants were screened immediately after the welcoming screen. They were asked who makes financial decisions in their households and what types of assets they have already invested in at least once.⁴ Participants who indicated that they do not make financial decisions in their household or had never invested in any asset besides a savings account or a call money account were immediately screened out and excluded from the study.⁵

Afterward, participants were walked through an example investment decision. We informed participants that the decisions could have real-world consequences as the decisions of five randomly drawn participants would be implemented. Details about the incentivization can be found at the end of this section. Thus, we informed participants that it is beneficial to answer truthfully and according to their preferences. Participants had to answer attention and comprehension checks. Those who failed were given a second opportunity to retake the test before being screened out.

Then, participants faced the two sets of investment decisions, which followed the abovementioned procedure but measured WTP for a different sustainable fund (see Table 1). Whether the WTP for Fund L or Fund S was elicited first was randomized to control for order effects. After completing each set of investment decisions, participants answered four questions about their perceptions of the different funds, including expected return, anticipated risk, and warm-glow

³The survey was carried out in German. A link to English instructions for the survey can be found in the Appendix.

⁴The possible options were: savings account, call money account, stocks or bonds, actively managed funds, passively managed funds, mixed funds, other fixed-income investments, other non-fixed-income investments, cryptocurrency, none of these investments, I do not know.

⁵This approach is similar to Gutsche and Ziegler (2019) and Gutsche, Wetzler, et al. (2023). However, our screening criteria were stricter, as individuals with only a savings account, a call money account, or both were excluded from participation.



Figure 1: Experimental Procedure

emotions. Each answer was given on a 5-point Likert scale. To measure warm-glow emotion, participants respond to the following question: "*Compared to investing in a conventional fund, investing in the impact fund feels...?*" Possible answers ranged from "*much worse*" to "*much better*." Finally, participants completed a post-experimental survey, which included questions about their views on climate change.

We incentivized both the financial and environmental aspects of participants' choices. One of the two investment decisions was carried out for five randomly chosen participants. Using the BDM-Mechanism (Becker et al. 1964), a random price between the highest and lowest elicited WTP was drawn. If the WTP for the sustainable fund was higher than the randomly drawn price, an investment of €1000 minus the random price was acquired on the participants' behalf. Additionally, the positive climate impact of the investment was realized by a donation to the offset organization Atmosfair. This is in line with a broad experimental literature using donations to incentivize (e.g. Kirchler et al. 2015). Donations to carbon offsets, in particular, are used to assess environmental or climate preference (e.g. Kesternich et al. 2019).

If the elicited WTP was below the randomly drawn price, an investment of \pounds 1000 was made into the same fund, but no offsets were purchased. The randomly selected participants will receive the initial investment in addition to potential losses or gains after one year. Separating the incentivization of the financial and climate component of the investment decision mirrors the incentivization from Heeb et al. (2023). This experimental approach is necessary as the realworld benefits of investments in sustainable funds are not clear and hard to quantify (Kölbel et al. 2020).

Psyma implemented the actual investment and the donation to Atomsfair. While incentivization is standard practice in analyzing the effects of sustainability information on retail investors' behavior in recent research (e.g. Heeb et al. 2023; Gutsche, Wetzler, et al. 2023), several influential earlier studies relied solely on stated preference methods and did not incentivize participants' choices (e.g Bassen et al. 2019; Gutsche and Ziegler 2019).

2.4 Hypotheses

This design allows us to investigate several key questions about the effect of labels on retail investors' investment behavior. The natural starting point is to begin by analyzing whether the WTP for sustainable investment increases with the magnitude of the positive climate impact in general and in the absence of labels. Accordingly, we first focus on the no label treatments and replicate part of the analysis from Heeb et al. (2023). This question is explored both within subjects by comparing the two elicited WTPs of each individual participant and across subjects by comparing participants in the low impact condition with those in the high impact condition. Assuming that participants have preferences for investments with a better carbon footprint leads to the following hypotheses:

Hypothesis 1.

- *a.* Within subjects, retail investors' WTP for investing in the sustainable fund increases when the impact of the fund is larger.
- **b.** Across subjects, retail investors' WTP for investing in the sustainable fund increases when the impact of the fund is larger.

After establishing how retail investors behave in the absence of labels, we turn our attention to the effects of labels. Here, too, the impact of higher label scores on investors' WTP is assessed both within and across subjects. Similar to the previous analysis, within subject comparisons involve the two elicited WTPs for each participant. However, as the magnitude of beneficial climate impact changes between the two sets of investment decisions, this approach cannot completely isolate the effect of better label scores. To address this limitation, our design allows for comparisons between subjects facing the same climate impact but different label scores comparing participants within the same impact treatment but in different label conditions.

For example, participants in the high impact and strict label condition choose between a fund with 0t of climate impact and 5t of climate impact, represented with zero and three leaves, respectively (see Table 1). Participants in the high impact and loose label condition face the same climate impact, but the impact is represented by zero and five leaves instead. Thus, the label score is the only variable changing between these participants, while the climate impact and all other factors remain identical.

In line with findings from the literature (e.g., Lohmann et al. 2022; Rodemeier and Löschel 2024), we investigate whether higher label scores lead to more pro-environmental friendly behavior, in our case, higher WTP values.

Hypothesis 2.

- *a.* Within subjects, retail investors' WTP for investing in the sustainable fund increases when the label score is higher.
- **b.** Across subjects, retail investors' WTP for investing in the sustainable fund is higher when the label score is higher.

Assuming that label scores strongly influence participants' investment decisions, a natural follow-up question is whether impact is necessary to increase WTP. If participants rely solely on label scores without paying attention to the underlying climate impact, it suggests that higher WTP values could be achieved without actual increases in climate benefits. This scenario raises concerns about the potential for greenwashing.

We compare participants' WTP within the same label standard but across different impact treatments to analyze this. Once again, turning to Table 1, participants in the high impact and strict label condition face a choice between a fund with no climate impact and one with 5t of climate impact, depicted by zero and three leaves, respectively. Participants in the low impact and the strict label condition also face a choice between a fund with a label score of zero and a fund with a label score of three. However, for the second participant, the actual impacts of the two funds are no climate impact and 0.5t of climate impact. If investors do not focus on the actual climate impact and mindlessly follow the label scores, the WTP measurements should not differ across the impact ranges when the label scores remain the same:

Hypothesis 3. *The WTP for investing in the sustainable fund does not differ between high impact and low impact treatments when the same label scores are displayed.*

Next, we compare the relative differences between the two elicited WTPs for each participant. This comparison enables us to assess how strongly retail investors' WTP responds to changes in climate impact and label scores. Following Heeb et al. (2023), we define sensitivity as the ratio of a participant's WTP for the fund with a larger impact (Fund L) to their WTP for the fund with a smaller impact (Fund S) in the respective treatment. If WTP scales linearly with the size of the climate impact, sensitivity would equal ten, as the beneficial climate impact differs by a factor of ten across the two rounds of investments for each participant (see Table 1).

We hypothesize that labels alter investors' sensitivity. While we are agnostic about the sign of the effect of labels in general, we argue that stricter labels lead to a higher sensitivity than loose labels. This hypothesis is based on behavioral economic models that feature diminishing sensitivity, such as distortions in consumer choices by salience effects (Bordalo et al. 2013). In many salience models, decision weights are influenced by proportional differences. In our context, a shift from one to three green leaves may be perceived as a larger relative change than a shift from three to five green leaves, leading to higher sensitivity in strict label conditions than in loose label conditions:

Hypothesis 4.

- *a.* The sensitivity of retail investors differs between the label and no label conditions.
- **b.** The sensitivity of retail investors is higher in the strict label condition than in the loose label condition.

Having explored the effect of labels on retail investors' WTP and sensitivity, the next step is to examine the mechanisms through which labels influence investor behavior. We do this in two steps: Research on pro-social behavior often emphasizes that the positive emotions associated with performing beneficial actions play a key role in decision-making (Imas 2014; Hartmann et al. 2017). In line with previous studies on retail investors' behavior in the absence of labels (Gutsche, Wetzler, et al. 2023; Heeb et al. 2023), we investigate whether labels affect the positive emotions, that is, warm-glow utility, experienced when investing in the sustainable fund.

Hypothesis 5.

- *a.* Within subjects, investors' warm-glow utility associated with investing in the sustainable fund increases as the label score improves.
- **b.** Across subjects, investors' warm-glow utility associated with investing in the sustainable fund increases as the label score improves.

Next, we analyze the effect of warm-glow utility on participants' WTP. We argue that participants' WTP is influenced by warm-glow utility and that this effect is especially strong when labels are present.

Hypothesis 6.

- a. A higher warm-glow utility leads to a higher WTP.
- **b.** The effect of warm-glow utility on retail investors' WTP is stronger when a label is present.

3 Results

3.1 Descriptives

Overall, 1219 German retail investors completed our study.⁶ We follow Heeb et al. (2023) in excluding participants who indicated that climate change is not a serious issue, reducing the sample for the main analysis to 1,121 participants. Additionally, we winsorize the WTP at the 5th and 95th percentiles to account for potential outliers. Both steps are pre-registered and do not drive the results as shown in the robustness checks (see section 3.8).

In line with previous work focusing on sustainable investments (e.g. Døskeland and Pedersen 2016; Gutsche, Wetzler, et al. 2023), our sample tilts male, with only 37.2 % of participants identifying as female. The average participant's age is 52 years, with the oldest being 88 and the youngest 18 years old. Compared to the general German population, our sample is comparatively well-educated, with 73.0 % having at least one university degree. The median income lies between €50.000 and €60.000, which is slightly above the median household income of Germany. Descriptive statistics summarizing the sample's characteristics are presented in the Appendix. Table B1 in the Appendix provides descriptive statistics for each treatment group and indicates that randomization was largely successful.⁷

⁶A total of 4,963 people entered the survey. Of these, 2,130 people were screened out as they did not meet the definition of retail investors. Additionally, 591 people failed the attention and comprehension check, and 1,023 did not complete the survey or were flagged as speedsters. Attrition rates did not vary significantly across treatments.

⁷A Kruskal-Wallis test detects significant differences across treatments for income and net worth (see Table B2 in the Appendix). Specifically, the mean income is slightly but significantly higher in the loose label conditions. Potential concerns that these differences are driving the results are addressed in section 3.8.



Figure 2: Average WTP by Treatment

3.2 Retail Investors' WTP for Climate Impact without Labels

Our findings reveal that retail investors exhibit a substantial and economically meaningful WTP for sustainable investments. Figure 2 illustrates the average WTP by treatment and for both investment decisions. We start our analysis by investigating the effect climate impact has on retail investors' WTP, focusing exclusively on the no label conditions. This replicates the joint evaluation from Heeb et al. (2023), and accordingly, the average WTP is at a similar level in our study, equaling approximately €33 in the no label conditions. Furthermore, our results support Heeb et al. (2023), showing that when retail investors are shown two funds with differences in the size of climate impact, their WTP responds to these relative changes. This is true both for the high impact as well as the low impact treatments: The WTP for fund L is larger than for fund S (Wilcoxon signed rank test⁸, p<0.01) indicating support for H1a.

However, our findings about the response to the absolute level of impact are more mixed. While Heeb et al. (2023) report no significant differences between the high and low impact treatment, our results reveal marginally significant differences in WTP for Fund S (Mann-Whitney U test, p < 0.1). For Fund L, no significant differences emerge (Mann-Whitney U test, p = 0.93). Thus, at best, there is limited evidence that a higher climate impact leads to higher WTP across subjects (see H1b).

Result 1. When comparing two sustainable investments, the WTP increases with the magnitude of climate impact (within subject). We find little evidence that changes in the absolute size of the impact influence retail investors' WTP (between subjects).

Overall, in the no label condition, investors show limited sensitivity to changes in the magnitude of climate impact. In the high impact treatment, the average WTP increases by about 3 %, despite a 10-fold increase in the magnitude of the beneficial climate impact. In the low impact group, the corresponding increase in WTP is about 20 %.

⁸As the data contains ties and zeros, we use the normal approximation to calculate p-values.

3.3 The Effects of Labels on Retail Investors' WTP

Now, we turn to the effect of labels on retail investors' WTP for sustainable investments. Overall, the patterns observed are similar to those in the no label treatments. Regarding H2a, in all treatments with a label, the WTP for fund L is significantly larger than for fund S (Wilcoxon signed rank test, p < 0.01). In other words, the WTP increases within subjects when participants see a higher label score.

As explained above, within subjects, both the label score and the climate impact change across investment decisions. Accordingly, it is important to make comparisons between subjects to isolate the effect of higher label scores. Therefore, the impact treatment is kept constant, and we compare the same investment decision, Fund L or Fund S, across participants. For example, we compare the WTP for retail investors in the strict standard and high impact treatment to participants in the loose standard and high impact treatment. As laid out in section 2.4, the only thing changing between these participants is the label scores, as the magnitude of climate impact is equal for all participants in the high impact condition.

In all cases, the average WTP in the loose label condition is higher than in the strict label condition. In other words, across subjects, the average WTP for the sustainable fund increases with a higher label score, all else being equal. This effect is not statistically significant for the comparisons of Fund L (Mann-Whitney U-Test, p=0.13 & p=0.28). However, for the comparisons of Fund S, the differences between the label conditions are statistically significant (Mann-Whitney U-Test, p<0.01). Furthermore, using Fisher's method for combining p-values shows joint significance (p<0.01) of all four tests, indicating support for H2b that between subjects, a higher label score leads to a higher WTP.

Result 2. A higher label score increases retail investors' WTP.

3.4 The Effect of Environmental Impact on WTP with Labels

To analyze the extent to which participants respond to the absolute level of climate impact when labels are present, we compare the participants in the high and low impact treatments for a given label treatment. This tests whether retail investors mindlessly follow the label scores or pay attention to the underlying criteria (H3).

The statistical analysis shows that a higher level of impact is associated with a higher WTP, even when the same label score is shown: Comparing the same investment decision (Fund L or Fund S) across impact treatments while holding the label scores constant returns significant results for all four tests (Mann-Whitney U-test, p<0.05).

Result 3. A larger impact increases retail investors' WTP even when the same label scores are displayed.

This result is surprising because, in the no-label treatments—where participants are directly given the magnitude of climate impact—no such effect is observed. Our finding that retail investors do not mindlessly follow label scores suggests that greenwashing through labels might only be feasible to a limited extent in the financial context, especially if information about the underlying label conditions is provided in an accessible and understandable way.



Figure 3: Average non-negative Sensitivity by Label Treatment

3.5 Retail Investors' Sensitivity to Climate Impact

The analysis above already shows that retail investors' WTP for sustainable investment reacts to changes in impact. However, on average, it scales only to a limited extent with the magnitude of the beneficial impact. As the experimental design provides two WTP measurements for each investor, we can calculate a sensitivity measure for each participant. Following Heeb et al. (2023), sensitivity is defined as the division of the WTP for Fund L by the WTP for Fund S. Figure 3 shows the average sensitivity for the three label conditions.⁹ Thus, sensitivity indicates the factor by which participants' WTP changes in response to a tenfold increase of the size to climate impact; e.g., a sensitivity of 2 would indicate that the WTP for Fund L is double that of Fund S.

We find that the inclusion of a label can both increase and decrease investors' sensitivity to the size of the climate impact compared to the no label conditions. Specifically, participants' sensitivity is higher in the strict label conditions but lower in the loose label conditions. Furthermore, the sensitivity in the strict label condition is significantly higher than in the loose label condition and no label condition (Mann-Whitney U-Test, p<0.01), indicating support for H4. Prior literature has already demonstrated that individuals tend to react only little to changes in the size of social or environmental benefits. We are the first to show that the presentation format of information can significantly alter the extent to which peoples' WTP changes in response to an increase in the magnitude of social or environmental impact.

Result 4. The introduction of labels changes investors' sensitivity compared to the no label condition. Furthermore, investors' sensitivity is higher in the strict label condition compared to the loose label condition.

To further support these findings, in Table 2 we categorize investors according to their sensitivity. Negative sensitivity includes participants with a sensitivity below 0, reverse sensitivity with a sensitivity between 0 and 1, insensitive people with a sensitivity of precisely 1 (the same WTP in both rounds), low sensitivity between 1 and 2, and high sensitivity for participants whose WTP more than doubles.

⁹We exclude participants with a negative sensitivity. A few observations indicate a large positive WTP for Fund L but a small negative WTP for Fund S, such as \notin -0.625. This leads to large negative sensitivities for these participants and wide confidence intervals. During the robustness checks in section 3.8, we show that the results are robust to the inclusion of these observations.

Across all treatments, the plurality of participants is insensitive, that is, they have the same WTP for Fund L and Fund S. Taking a closer look at the sensitivity distribution across groups reaffirms that a strict label increases sensitivity. The share of participants with high sensitivity increases by approximately 50 % with the introduction of a strict label compared to the no label treatments, rising from 14 % to 21 % (two-proportions Z-Test, p<0.05). Conversely, the share decreases by about 35 %, to 9 %, in the loose label condition (two-proportions Z-Test, p<0.05). Additionally, the share of participants completely insensitive to changes in impact is significantly lower in the strict label condition than in the other treatments (p<0.05, two-proportions Z-Test). Additionally, no significant differences in the share of participants with reverse sensitivity can be observed. Regarding the low sensitivity group, a significant increase in the share can be observed for both types of labels compared to the no label treatments (p<0.1, two-proportions Z-Test).

	Loose Label	No Label	Strict Label
Negative Sensitivity	0.04	0.07	0.05
Reverse Sensitiity	0.10	0.13	0.09
Insensitive	0.43	0.41	0.33
Low sensitivity	0.33	0.25	0.31
High Sensitivity	0.09	0.14	0.21

Table 2: Share of Retail Investors in each Sensitivity Groups by Label Treatment

3.6 Labels and Warm Glow Utility

Our final two hypotheses concern the mechanism by which labels influence participants' WTP for sustainable investments. First, we analyze if labels influence the positive feeling associated with investing in a sustainable fund. As explained in section 2.3, the variable warm-glow utility is based on participants' responses to the question: "Compared to investing in the conventional fund, investing in the impact fund feels..." Responses were recorded on a 5-point Likert scale from "much worse" (-2) to "much better" (+2).

Similar to earlier analyses, we examine variations both within subjects and between subjects. Within subjects, a higher label score is consistently associated with a higher measure of warmglow utility across all treatments (Wilcoxon signed rank test, p<0.05). Fisher's method for combining p-values indicates joint significance (p<0.01).

As noted before within subject comparisons induce changes not only in the label score but also in the connected impact. To isolate the effect of labels on warm-glow utility, we focus on between subjects comparisons, keeping the impact treatment constant. Higher label scores lead to a significantly larger reported warm-glow utility in the high impact treatments (Mann Whitney U test, p<0.05). For the low impact treatments we find a marginally significant effect of higher label scores on warm-glow utility for fund L (Mann Whitney U test, p<0.1) and no significant effect for comparing the warm-glow utility connected to fund S. Furthermore, testing for joint significance supports the hypothesis that a higher label score is connected to higher warm-glow utility (Fisher's method, p < 0.01). Overall, we find support for H5:

Result 5. A higher label score increases retail investors' warm-glow from investing in a sustainable fund.

	D	Dependent variable: Willingness to Pay (WTP)				
	Fund L	Fund L	Fund S	Fund S	Pooled	Pooled
WG	12.97*** (1.25)	8.41*** (1.82)	11.88*** (1.34)	9.58*** (1.94)	12.49*** (1.14)	8.93*** (1.66)
WG x Label		6.17*** (1.80)		3.25 (1.99)		4.91*** (1.65)
Low Impact	-5.19*** (1.87)	-4.73** (1.86)	-5.42*** (1.81)	-5.19*** (1.82)	-5.32*** (1.72)	-4.96*** (1.72)
Fund L					4.70*** (0.70)	4.66^{***} (0.70)
Controls	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Observations \mathbb{R}^2	1,005 0.12	1,005 0.13	1,013 0.10	1,013 0.10	2,018 0.12	2,018 0.12

Table 3: Regression Results for Climate Impact and Warm-Glow

*p<0.1; **p<0.05; ***p<0.01, standard errors are reported in brackets.

3.7 Warm-glow Utility and WTP for Sustainable Investments

We now analyze the effect of warm-glow utility on the WTP for a sustainable investment. Table 3 presents the results of regressing participants' WTP on our measurement of warm-glow utility (WG). The analysis is conducted individually for Fund L, Fund S, and pooled for all investment decisions in our sample. The pooled regressions include participants twice, accordingly we cluster standard errors on the participant level. All regressions control for the level of impact and include the following socio-demographic variables: age, education level, income, net worth, and gender. In the pooled regressions, we additionally control for whether the observation refers to Fund L or Fund S. Additionally, some regressions include an interaction term between WG and a dummy variable equal to 1 for both the strict and loose label conditions.

In line with the results before, participants' WTP decreases by about €5 in the low impact treatments compared to the high impact treatments. Additionally, the WTP for Fund L is approximately €5 higher than for Fund S. Both effects are significant across all regressions

More importantly, the results show that warm-glow utility is strongly related to retail investors' WTP. A one-point increase on the five-point Likert scale is connected to an increase in the WTP of \notin 8 to \notin 13. This effect is highly significant in all regressions and economically meaningful. Using all investment decisions, as shown in column (5), this effect equals \notin 12.45 per point on the Likert scale, which is about a third of the average WTP or 40 % of a standard deviation. These findings align with the analysis of Heeb et al. (2023) and Gutsche, Wetzler, et al. (2023), showing that retail investors' WTP for sustainable investments is heavily influenced by the positive emotions they associate with the investment.

Furthermore, the interaction effect reveals that this effect of warm-glow utility on the WTP

is particularly strong when a label is present. The interaction term is positive in all regressions, increasing the change in WTP that is connected to a one-point higher warm-glow utility by an additional \notin 3 to \notin 6. The effect is highly significant for the regressions focusing on Fund L and the pooled regressions but just misses statistical significance for the regression related to Fund S (p = 0.102).

Result 6. The WTP of retail investors for a sustainable investment depends strongly on the positive emotions connected to the investment. This effect is even stronger when a label is present.

3.8 Robustness Checks

We conduct several tests to show the robustness of our results, which are all shown in the Appendix. First, we show that result 4 is robust to including observations with a negative sensitivity Figure C1 shows that compared to the no label treatments, the average sensitivity increases in the strict label treatments but decreases in the loose label treatments. Furthermore, sensitivity is significantly higher in the strict label treatments than in the no or loose label treatments (Mann-Whitney U-Test, p < 0.01).

Additionally, we re-run all the tests from above using the complete sample, that is, without excluding participants who answered that climate change is not a serious problem and without winsorizing the WTP. In general, the results and patterns remain consistent with those reported above. The results are depicted in Table B3-B4 and Figure C2 in the Appendix. The only minor deviation is found for the effect of labels on retail investors' sensitivity, that is, H4. Not winsorizing the data results in a few highly influential data points, leading to different findings when only the means are compared (see Appendix Figure C3). However, we find the same significant results as before when considering only non-negative sensitivities (see Appendix Figure C4 and Table B3). Moreover, when focusing on the comparison between the different sensitivity groups, we observe the same patterns and statistical significance as before. Specifically, the introduction of a strict label still increases the proportion of retail investors with high sensitivity.

Additionally, we examine whether income or net worth are significant predictors of WTP (see Appendix Table B6). This is not the case, suggesting that the imbalances across treatment groups, noted in section 3, do not drive the results. Furthermore, we conduct our primary analyses separately for individuals above the median and below or at the median income, reaffirming our findings, see Appendix Table B5.

4 Conclusion

In this paper, we demonstrate that the way investors receive sustainability information significantly influences their investment behavior, particularly in terms of sensitivity to impact and warm-glow utility. Furthermore, we find that both within and between subjects, investors' WTP for a sustainable fund increases as the label score improves. Additionally, we show that presenting the positive impact associated with an investment using a label, rather than plain text, changes investors' sensitivity to impact. Notably, depending on the label standard investors' sensitivity increases with a strict label and decreases with a loose label. Finally, we establish that changes in participants' WTP for sustainable investments are strongly dependent on the warm-glow utility they report, with this effect being more pronounced when a label is present.

Our findings have implications for policymakers who want to regulate label standards, as e.g., the FCA in the UK (Financial Conduct Authority 2024). If policymakers want to ensure that firms face strong incentives to act environmentally friendly, our findings suggest that label standards should not be too loose.

References

- Andor, Mark A., Andreas Gerster, and Stephan Sommer (2020). "Consumer Inattention, Heuristic Thinking and the Role of Energy Labels". In: *The Energy Journal* 41 (1), pp. 83–112.
- Andreoni, James (1990). "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving". In: *The Economic Journal* 100 (401), pp. 464–477.
- Barber, Brad M., Adair Morse, and Ayako Yasuda (2021). "Impact Investing". In: *Journal of Financial Economics* 139.1, pp. 162–185.
- Bassen, Alexander, Katrin Gödker, Florian Lüdeke-Freund, and Josua Oll (2019). "Climate Information in Retail Investors' Decision-Making: Evidence From a Choice Experiment". In: *Organization Environment* 32 (1), pp. 62–82.
- Bauer, Rob, Tobias Ruof, and Paul Smeets (2021). "Get Real! Individuals Prefer More Sustainable Investments". In: *The Review of Financial Studies* 34.8, pp. 3976–4043.
- Becker, Gordon, Morris H. Degroot, and Jacob Marschak (1964). "Measuring utility by a single-response sequential method". In: *Behavioral Science* 9 (3), pp. 226–232.
- Bioy, Hortense, Boya Wang, Noemi Pucci, Arthur Carabia, Shamir Popat, Hunter Beaudoin, Risa Mizuta, and Biddappa A. R. (2024). *Global Sustainable Fund Flows: Q3 2024 in Review*. Morningstar Sustainalytics. uRL: https://www.morningstar.com/lp/global-esg-flows.
- Bonnefon, Jean-François, Augustin Landier, Parinitha Sastry, and David Thesmar (2025). "The moral preferences of investors: Experimental evidence". In: *Journal of Financial Economics* 163, p. 103955.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2013). "Salience and Consumer Choice". In: *Journal of Political Economy* 121.5, pp. 803–843.
- Chen, Daniel L., Martin Schonger, and Chris Wickens (2016). "oTree—An open-source platform for laboratory, online, and field experiments". In: *Journal of Behavioral and Experimental Finance* 9(C), pp. 88–97.
- Døskeland, Trond and Lars Jacob Tynes Pedersen (2016). "Investing with Brain or Heart? A Field Experiment on Responsible Investment". In: *Management Science* 62.6, pp. 1533–1841.
- EU (2019). Regulation (EU) 2019/2088 of the European Parliament and of the Council of 27 November 2019 on sustainability-related disclosures in the financial services sector. URL: https://eur-lex.europa.eu/eli/reg/2019/2088/oj (visited on 10/30/2024).
- Financial Conduct Authority (2024). Finalised non-handbook guidance on the Anti-Greenwashing Rule. URL: https://www.fca.org.uk/publication/finalised-guidance/fg24-3.pdf (visited on 10/15/2024).

- Gutsche, Gunnar, Heike Wetzler, and Andreas Ziegler (2023). "Determinants of individual sustainable investment behavior - A framed field experiments". In: *Journal of Economic Behavior* & Organization 209, pp. 491–508.
- Gutsche, Gunnar and Andreas Ziegler (2019). "Which private investors are willing to pay for sustainable investments? Empirical evidence from stated choice experimentss". In: *Journal of Banking and Finacne* 102, pp. 199–214.
- Hartmann, Patrick, Martin Eisend, Vanessa Apaolaza, and Clare D'Souza (2017). "Warm glow vs. altruistic values: How important is intrinsic emotional reward in proenvironmental behavior?" In: *Journal of Environmental Psychology* 52, pp. 43–55.
- Hartzmark, Samuel M. and Abigail B. Sussman (2019). "Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows". In: *The Journal of Finance* 74.6, pp. 2789–2837.
- Heeb, Florian, Julian F. Kölbel, Falko Paetzold, and Stefan Zeisberger (2023). "Do Investors Care About Impact?" In: *The Review of Financial Studies* 36 (5), pp. 1737–1787.
- Imas, Alex (2014). "Working for the "warm glow": On the benefits and limits of prosocial incentives". In: *The Journal of Public Economics* 114, pp. 14–18.
- Kaufmann, Marc, Peter Andre, and Botond Kőszegi (2024). "Understanding Markets with Socially Responsible Cons". In: *The Quarterly Journal of Economics* 139 (3), pp. 1989–2035.
- Kesternich, M., D. Römer, and F. Flues (2019). "The power of active choice: Field experimental evidence on repeated contribution decisions to a carbon offsetting program". In: *European Economic Review* 114, pp. 76–91.
- Kirchler, Michael, Jürgen Huber, Matthias Stefan, and Matthias Sutter (2015). "Market Design and Moral Behavior". In: *Management Science* 62.9, pp. 2615–2625.
- Klaaßen, Lena and Bjarne Steffen (2023). "Meta-analysis on necessary investment shifts to reach net zero pathways in Europe". In: *Nature Climate Change* 13, pp. 58–66.
- Kölbel, Julian F., Florian Heeb, Falko Paetzold, and Timo Busch (2020). "Can Sustainable Investing Save the World? Reviewing the Mechanisms of Investor Impact". In: *Organization & Environment* 33.4, pp. 554–574.
- Kuilen, Gijs van de and Peter P. Wakker (2011). "The Midweight Method to Measure Attitudes Toward Risk and Ambiguity". In: *Management Science* 57 (3), pp. 582–598.
- Lohmann, Paul M., Elisabeth Gsottbauer, Anya Doherty, and Andreas Kontoleon (2022). "Do carbon footprint labels promote climatarian diets? Evidence from a large-scale field experiment". In: *Journal of Environmental Economics and Management* 114, p. 102693.
- Lopes, Ana Faria and Gorm Kipperberg (2020). "Diagnosing Insensitivity to Scope in Contingent Valuation". In: *Environmental and Resource Economics* 77, pp. 191–216.
- Lyon, Thomas P. and A. Wren Montgomery (2015). "The Means and End of Greenwash". In: *Organization & Environment* 28.2, pp. 223–249.
- Morningstar (2021). Morningstar Sustainability Rating Methodology. URL: https://www.morningstar. com/content/dam/marketing/shared/research/methodology/744156_Morningstar_ Sustainability_Rating_for_Funds_Methodology.pdf (visited on 10/15/2024).

- Pace, Davide D., Taisuke Imai, Peter Schwardmann, and Joël J. van der Weele (2025). "Uncertainty about carbon impact and the willingness to avoid CO2 emissions". In: *Ecological Economics* 227, p. 108401.
- Riedl, Arno and Paul Smeets (2017). "Why Do Investors Hold Socially Responsible Mutual Funds?" In: *Journal of Finance* 72, pp. 2505–2550.
- Rodemeier, Matthias (2024). "Willingness to Pay for Carbon Mitigation: Field Evidence from the Market for Carbon Offsets". In: *IZA Discussion Paper* 15939.
- Rodemeier, Matthias and Andreas Löschel (2024). "Information Nudges, Subsidies, and Crowding Out of Attention: Field Evidence from Energy Efficiency Investments". In: *Journal of the European Economic Association* In press.
- Schulze Tilling, Anna (2024). "Changing consumption behavior with carbon labels: Causal evidence on behavioral channels and effectiveness". In: *Working Paper*.
- Seifert, Marcel, Florian Spitzer, Simone Haeckl, Alexia Gaudeul, Erich Kirchler, Stefan Palan, and Katharina Gangl (2024). "Can information provision and preference elicitation promote ESG investments? Evidence from a large, incentivized online experiment". In: *Journal of Banking & Finance* 161.

Appendix

A Survey

The experiment was conducted in German. Below, you see the same investment decision for people in the three different label conditions as they appeared in the survey. Instructions in English are available here: https://drive.google.com/file/d/1_stQx8zQZjElmkiT3XjERAAH9WHT1ltT.

	Fonds A	Fonds B	Information
Fondskategorie	US Large-Cap BlendEquity	US Large-Cap BlendEquity	Anlageklasse und Marktsegment, die der Fonds investiert
Durchschnittliche Rendite (3 Jahre)	6%	6%	Durchschnittlicher jährlicher Ertra einer Anlage in den Fond
MorningStar Risiko	Durchschnittlich	Durchschnittlich	Bewertet die Variation der monatlichen Renditen eines Fond im Vergleich zu ähnlichen Fonds
Klima-Impact	Eine 1.000C-Investition in einen Fonds mit diesem Label spart keine CO ₂ -Emissionen.	Eine 1.000C-Investition In einen Fonds mit diesem Label spart St CO ₂ , Diese CO ₂ -Reduktion entspricht: • Der CO ₂ -Einsparung durch das Pflanzen von dreifigi Baumen, • Den CO ₂ -Emissionen einer Flugreise von 15.000km. • Der CO ₂ -Emission verursacht durch einen EU-Bürger innerhalb 250 Tagen.	Manche Fonds finanzieren Projek die CO2-Emissionen einsparen Einige Expert*innen sagen, dass di eine entscheidene Moglicheit Investor*innen ist. dem Kinawan entgegerzuwirken. Andere argumentieren, dass die Ablenkung ist und die notwendig Maßnahmen zur Bekämpfung di Kiimawandek (z.B. CO2-Steuer verzögern könnte.



Figure A1: Investment Decision - No Label

Anlageoptionen:					
	Fonds A	Fonds B	Informationen		
Fondskategorie	US Large-Cap BlendEquity	US Large-Cap BlendEquity	Anlageklasse und Marktsegment, in die der Fonds investiert		
Durchschnittliche Rendite (3 Jahre)	6%	6% Durchschnittlicher jährlicher Ertrag einer Anlage in			
MorningStar-Risiko	Durchschnittlich	Durchschnittlich	Bewertet die Variation der monatlichen Renditen eines Fonds im Vergleich zu ähnlichen Fonds		
Klima-Impact	99999	Manche Fonds finanzieren Projekte. die CO2-Emissionen einspare Einige Expert'finnen sagen, dass dies eine entscheidende Moglichkeit für Investor ⁴ innen ist. dem Klimawandel entogenzuwirken. Andere argumentieren, dass dies Ablenkung ist und die notwendigen Maßnahmen zur Bekämpfung des Klimawandes (z.) CO2-Steur) verzögeren könnte.			
⊢ Klicken Sie hier, um die z Bitte geben Sie an, in wel	ugrunde liegenden Grenzw	erte erneut zu sehen: estieren möchten: Investitionsent	scheidung 1:		
Investition von 1.000€ abzüglich einer Gebühr von 10.00€ in Endet B Endet B					

Figure A2: Investment Decision - Loose Label

Anlageoptionen:					
	Fonds A	Fonds B	Informationen		
Fondskategorie	US Large-Cap BlendEquity	US Large-Cap BlendEquity	Anlageklasse und Marktsegment, in die der Fonds investiert		
Durchschnittliche Rendite (3 Jahre)	6%	6%	Durchschnittlicher jährlicher Ertrag einer Anlage in den Fond		
MorningStar-Risiko	Durchschnittlich	Durchschnittlich	Bewertet die Variation der monatlichen Renditen eines Fonds im Vergleich zu ähnlichen Fonds		
Klima-Impact	99999	99999	Manche Fonds finanzieren Projekte, die CO2-Emissionen einsparen. Einige Expertfinnen sagen, dass dies eine entscheidende Möglichkeit für Investorfinnen ist, dem Klimawandel entgegenzuwirken. Andere argumentieren, dass dies Ablenkung ist und die notwendigen Maßnahmer zur Bekämpfung des Klimawandels (z.B. CO2-Steuer) verzögern könnte.		
► Klicken Sie hier, um die zugrunde legenden Grenzwerte erneut zu sehen: Bitte geben Sie an, in welchen Fonds Sie 1.000€ investieren möchten:					
	investrionsentscheidung 1:				
Investition von 1.000€ Investition von 1.000€ abzüglich einer Gebühr von 10.00€ in Fonds A Fonds B					

Figure A3: Investment Decision - Strict Label

B Additional Tables

Table B1: Descriptive Statistics

Variable	Mean	Min	Max	SD
Female	0.37	0.00	1.00	0.48
Education	3.53	1.00	6.00	1.26
Income	6.04	1.00	13.00	2.69
Age	51.51	19.00	88.00	13.37
Climate Awareness	5.88	-10.00	10.00	4.87
Net Worth	3.02	1.00	8.00	2.29

Table B2: Summary of Means and Standard Errors by Treatment

Treatment	Female	Education	Income	Age	Climate Awareness	Net worth
Loose Standard High Impact	0.33 (0.03)	3.56 (0.09)	6.49 (0.20)	51.55 (0.91)	6.40 (0.35)	2.89 (0.15)
Loose Standard Low Impact	0.37 (0.03)	3.75 (0.09)	6.84 (0.21)	51.94 (0.91)	5.81 (0.32)	3.42 (0.17)
No Standard High Impact	0.37 (0.03)	3.44 (0.09)	5.68(0.18)	53.45 (0.88)	5.55 (0.37)	2.88 (0.16)
No Standard Low Impact	0.39 (0.03)	3.50 (0.08)	5.93 (0.18)	50.81 (0.98)	5.64 (0.35)	2.80 (0.15)
Strict Standard High Impact	0.42 (0.03)	3.44 (0.09)	5.79 (0.18)	50.25 (0.93)	6.28 (0.30)	3.15 (0.17)
Strict Standard Low Impact	0.35 (0.03)	3.48 (0.09)	5.49 (0.17)	51.19 (1.00)	5.61 (0.36)	2.97 (0.17)
Kruskal–Wallis (p-values)	0.568	0.101	0.001***	0.251	0.154	0.045**

Table B3: Results Robustness Tests - Con	nplete Sample not Winsorized
--	------------------------------

Н	Test	Comparison	Variable 1	Variable 2	p-value
H1a	Wilcoxon signed rank test	Within subject	No Label Low Impact Fund S	No Label Low Impact Fund L	<0.01***
H1a	Wilcoxon signed rank test	Within subject	No Label High Impact Fund S	No Label High Impact Fund L	0.012**
H1b	Mann-Whitney U test	Across subjects	No Label Low Impact Fund S	No Label High Impact Fund S	0.7943
H1b	Mann-Whitney U test	Across subjects	No Label Low Impact Fund L	No Label High Impact Fund L	0.038**
H2a	Wilcoxon signed rank test	Within subject	Strict Standard Low Impact Fund S	Strict Standard Low Impact Fund L	< 0.01***
H2a	Wilcoxon signed rank test	Within subject	Strict Standard High Impact Fund S	Strict Standard High Impact Fund L	< 0.01***
H2a	Wilcoxon signed rank test	Within subject	Loose Standard Low Impact Fund S	Loose Standard Low Impact Fund L	< 0.01***
H2a	Wilcoxon signed rank test	Within subject	Loose Standard High Impact Fund S	Loose Standard High Impact Fund L	< 0.01***
H2b	Mann-Whitney U test	Across subjects	Strict Standard High Impact Fund S	Loose Standard High Impact Fund S	0.277
H2b	Mann-Whitney U test	Across subjects	Strict Standard Low Impact Fund S	Loose Standard Low Impact Fund S	0.205
H2b	Mann-Whitney U test	Across subjects	Strict Standard High Impact Fund L	Loose Standard High Impact Fund L	< 0.01***
H2b	Mann-Whitney U test	Across subjects	Strict Standard Low Impact Fund L	Loose Standard Low Impact Fund L	< 0.01***
H2b	Fisher's method test		*	*	< 0.01***
H3	Mann-Whitney U test	Across subjects	Strict Standard High Impact Fund S	Strict Standard Low Impact Fund S	0.011**
H3	Mann-Whitney U test	Across subjects	Strict Standard High Impact Fund L	Strict Standard Low Impact Fund L	0.030**
H3	Mann-Whitney U test	Across subjects	Loose Standard High Impact Fund S	Loose Standard Low Impact Fund S	0.015**
H3	Mann-Whitney U test	Across subjects	Loose Standard High Impact Fund L	Loose Standard Low Impact Fund L	0.038**
H4a	Mann-Whitney U test	Across subjects	Sensitivity Strict Standard	Sensitivity No Label	< 0.01***
H4a	Mann-Whitney U test	Across subjects	Sensitivity Loose Standard	Sensitivity No Label	0.368
H4b	Mann-Whitney U test	Across subjects	Sensitivity Strict Standard	Sensitivity Loose Standard	< 0.01***
H4a	Mann-Whitney U test	Across subjects	non-negative Sensitivity Strict Standard	non-negative Sensitivity No Label	< 0.01***
H4a	Mann-Whitney U test	Across subjects	non-negative Sensitivity Loose Standard	non-negative Sensitivity No Label	0.92
H4b	Mann-Whitney U test	Across subjects	non-negative Sensitivity Strict Standard	non-negative Sensitivity Loose Standard	< 0.01***
H4a	Two proportions Z test	Across subjects	share high Sensitivity Strict Standard	share high Sensitivity No Label	0.010**
H4a	Two proportions Z test	Across subjects	share high Sensitivity Loose Standard	share high Sensitivity No Label	0.043**
H4b	Two proportions Z test	Across subjects	share high Sensitivity Strict Standard	share high Sensitivity Loose Label	p<0.01***
H5a	Wilcoxon signed rank test	Within subject	WG Strict Standard Low Impact Fund S	WG Strict Standard Low Impact Fund L	0.014**
H5a	Wilcoxon signed rank test	Within subject	WG Strict Standard High Impact Fund S	WG Strict Standard High Impact Fund L	< 0.01***
H5a	Wilcoxon signed rank test	Within subject	WG Loose Standard Low Impact Fund S	WG Loose Standard Low Impact Fund L	< 0.01***
H5a	Wilcoxon signed rank test	Within subject	WG Loose Standard High Impact Fund S	WG Loose Standard High Impact Fund L	0.033**
H5b	Mann-Whitney U test	Across subjects	WG Strict Standard High Impact Fund S	WG Loose Standard High Impact Fund S	0.189
H5b	Mann-Whitney U test	Across subjects	WG Strict Standard Low Impact Fund S	WG Loose Standard Low Impact Fund S	0.063*
H5b	Mann-Whitney U test	Across subjects	WG Strict Standard High Impact Fund L	WG Loose Standard High Impact Fund L	0.061*
H5b	Mann-Whitney U test	Across subjects	WG Strict Standard Low Impact Fund L	WG Loose Standard Low Impact Fund L	0.157
H5b	Fisher's method		-	-	0.020**

If not specified otherwise the variables refer to the WTP. WG refers to the measurement of warm-glow utility. Results for H6 can be found in the regression table below *p<0.1; **p<0.05; ***p<0.01

	Dependent variable: Willingness to Pay (WTP)					
	(1)	(2)	(3)	(4)	(5)	(6)
WG	24.04*** (4.75)	14.00* (7.30)	26.53*** (5.07)	20.29*** (7.74)	25.18*** (3.96)	16.89*** (4.57)
WG x Label		13.54* (7.47)		8.76 (8.19)		11.37** (5.09)
Low Impact	-20.52*** (7.54)	-19.53*** (7.56)	-21.51*** (7.34)	-20.89*** (7.36)	-21.00*** (6.89)	-20.17*** (6.92)
Fund L					5.96** (2.79)	5.89** (2.79)
Controls	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Observations R ²	1,091 0.04	1,091 0.04	1,102 0.04	1,102 0.04	2,193 0.04	2,193 0.04

Table B4: Regression Results for Impact and Warm-Glow - complete sample not Winsorized

*p<0.1; **p<0.05; ***p<0.01, Standard errors are reported in brackets.

Table B5: Robustness Tests	- Low and High Income
----------------------------	-----------------------

Н	Test	Comparison	Variable 1	Variable 2	p-value Low Income	p-value High Income
H1a	Wilcoxon signed rank test	Within subject	No Label Low Impact Fund S	No Label Low Impact Fund L	< 0.01***	< 0.01***
H1a	Wilcoxon signed rank test	Within subject	No Label High Impact Fund S	No Label High Impact Fund L	0.163	0.013**
H1b	Mann-Whitney U test	Across subjects	No Label Low Impact Fund S	No Label High Impact Fund S	0.581	0.688
H1b	Mann-Whitney U test	Across subjects	No Label Low Impact Fund L	No Label High Impact Fund L	0.416	0.252
H2a	Wilcoxon signed rank test	Within subject	Strict Standard Low Impact Fund S	Strict Standard Low Impact Fund L	< 0.01***	< 0.01***
H2a	Wilcoxon signed rank test	Within subject	Strict Standard High Impact Fund S	Strict Standard High Impact Fund L	< 0.01***	< 0.01***
H2a	Wilcoxon signed rank test	Within subject	Loose Standard Low Impact Fund S	Loose Standard Low Impact Fund L	< 0.01***	0.011**
H2a	Wilcoxon signed rank test	Within subject	Loose Standard High Impact Fund S	Loose Standard High Impact Fund L	< 0.01***	< 0.01***
H2b	Mann-Whitney U test	Across subjects	Strict Standard High Impact Fund S	Loose Standard High Impact Fund S	0.252	0.758
H2b	Mann-Whitney U test	Across subjects	Strict Standard Low Impact Fund S	Loose Standard Low Impact Fund S	0.073*	0.975
H2b	Mann-Whitney U test	Across subjects	Strict Standard High Impact Fund L	Loose Standard High Impact Fund L	< 0.01***	0.091*
H2b	Mann-Whitney U test	Across subjects	Strict Standard Low Impact Fund L	Loose Standard Low Impact Fund L	0.031**	0.221
H2b	Fisher's method	-	-	-	< 0.01***	0.394
H3	Mann-Whitney U test	Across subjects	Strict Standard High Impact Fund S	Strict Standard Low Impact Fund S	0.295	< 0.01***
H3	Mann-Whitney U test	Across subjects	Strict Standard High Impact Fund L	Strict Standard Low Impact Fund L	0.420	0.043**
H3	Mann-Whitney U test	Across subjects	Loose Standard High Impact Fund S	Loose Standard Low Impact Fund S	0.765	< 0.01***
H3	Mann-Whitney U test	Across subjects	Loose Standard High Impact Fund L	Loose Standard Low Impact Fund L	0.326	0.016**
H4a	Mann-Whitney U test	Across subjects	Sensitivity Strict Standard	Sensitivity No Label	< 0.01***	< 0.01***
H4a	Mann-Whitney U test	Across subjects	Sensitivity Loose Standard	Sensitivity No Label	0.38	0.651
H4b	Mann-Whitney U test	Across subjects	Sensitivity Strict Standard	Sensitivity Loose Standard	0.037**	< 0.01***
H5a	Wilcoxon signed rank test	Within subject	WG Strict Standard Low Impact Fund S	WG Strict Standard Low Impact Fund L	0.276	0.019**
H5a	Wilcoxon signed rank test	Within subject	WG Strict Standard High Impact Fund S	WG Strict Standard High Impact Fund L	< 0.01***	0.025**
H5a	Wilcoxon signed rank test	Within subject	WG Loose Standard Low Impact Fund S	WG Loose Standard Low Impact Fund L	0.268	0.011**
H5a	Wilcoxon signed rank test	Within subject	WG Loose Standard High Impact Fund S	WG Loose Standard High Impact Fund L	< 0.01***	< 0.01***
H5b	Mann-Whitney U test	Across subjects	WG Strict Standard High Impact Fund S	WG Loose Standard High Impact Fund S	0.296	0.534
H5b	Mann-Whitney U test	Across subjects	WG Strict Standard Low Impact Fund S	WG Loose Standard Low Impact Fund S	0.122	0.126
H5b	Mann-Whitney U test	Across subjects	WG Strict Standard High Impact Fund L	WG Loose Standard High Impact Fund L	0.084*	0.533
H5b	Mann-Whitney U test	Across subjects	WG Strict Standard Low Impact Fund L	WG Loose Standard Low Impact Fund L	0.102	0.208
H5b	Fisher's method	,		•	0.040**	0.279

If not specified otherwise the variables refer to the WTP. WG refers to the measurement of warm-glow utility. *p<0.1; **p<0.05; **p<0.01

	Dependent variable: Willingness to Pay (WTP)						
	(1)	(2)	(3)	(4)	(5)	(6)	
Income	-0.49 (0.54)	-0.46 (0.53)	-0.03 (0.52)	$0.05 \\ (0.51)$	-0.26 (0.37)	-0.21 (0.37)	
Net worth	0.13 (0.41)	-0.03 (0.41)	0.07 (0.39)	-0.20 (0.39)	0.10 (0.28)	-0.12 (0.28)	
Fund L					6.28*** (1.34)	6.28*** (1.33)	
Treatment FE		\checkmark		\checkmark		\checkmark	
Observations R ²	1,033 0.01	1,033 0.05	1,033 0.01	1,033 0.04	2,066 0.02	2,066 0.05	

Table B6: Regression Results for Income and Net Worth - Winsorized and filtered sample

Column (1) and (2) refer to Fund S, (3) and (4) to Fund L, and (5) and (6) pool the data. p<0.1; p<0.05; p<0.05; p<0.01, Standard errors are reported in brackets.

C Additional Figures



Figure C1: Average Sensitivity by Label Treatment



Figure C2: Average WTP by Treatment Group - Complete Sample not Winsorized



Figure C3: Average Sensitivity by Label Treatment - Complete Sample not winsorized



Figure C4: Average Sensitivity by Label Treatment- Complete Sample not winsorized excluding negative values