Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of the Environment,
Transport, Energy and Communications DETEC

Swiss Federal Office of Energy
Energy Research and Cleantech

Final report of 22 May 2026

INFINEED

The Interplay of Feedback and Incentive Effects on Electricity Demand

---

Source: own design, 2026.

Publisher:
Swiss Federal Office of Energy SFOE
Energy Research and Cleantech
CH-3003 Berne
www.energy-research.ch

Subsidy recipients:
University of Neuchâtel (UniNE)
CH-2000 Neuchâtel
www.unine.ch

University of Applied Sciences and Arts Western Switzerland (HES-SO)
CH-1227 Carouge
www.hesge.ch

Authors:
Adrian Holzer, UniNE, adrian.holzer@unine.ch
Sylvain Weber, HES-SO, sylvain.weber@hesge.ch
Mehdi Farsi, UniNE, mehdi.farsi@unine.ch
Bruno Kocher, UniNE, bruno.kocher@unine.ch
Manon Berney, UniNE, manon.berney@unine.ch
Valentin Favre-Bulle, HES-SO, valentin.favre-bulle@hesge.ch
Fabien Giauque, UniNE, fabien.giauque@unine.ch
Vladimir Macko, UniNE, vladimir.macko@unine.ch
Michael Puntiroli, UniNE, michael.puntiroli@unine.ch

SFOE project coordinators:
Anne-Kathrin Faust, Anne-Kathrin.Faust@bfe.admin.ch

SFOE contract number: SI/502530-01

Summary

This project studies how individual feedback and social comparisons combined with financial incentives influence residential electricity consumption. To study the impact of interventions, we rely on a field experiment conducted on a sample of 201 households from private clients of an energy provider in Bernese Jura. We develop a digital interface and a mobile app fed by smart meter consumption data to provide feedback on the household's realized savings, which are carefully computed, based on comparable user profiles from a reference group of about 6,000 households. Econometric models are used to identify the moderating effects of a selection of individual characteristics. We also analyze the participants' engagement using 25 follow-up interviews. Finally, we conduct five online experiments on large national samples varying from 291 to 1,199 respondents to identify the preferences of various household types under a selection of different treatments. These experiments allow us to study relative tendencies among different sustainable actions, and the effect of ambiguity in choices and outcomes.

The field experiment data provide little evidence of any statistically significant effect on realized savings. While being statistically insignificant, the results point to the possibility of counterproductive effects of interventions in participating households equipped with photovoltaic (PV) production. However, the results suggest a significant engagement in saving and a high appreciation of information among the participants. The stated-preference experiments confirm the respondents' readiness to save electricity and suggest a significant effect of information feedback such as social comparisons and climate change awareness. The information on financial savings appears to have little effect on saving intentions and could cause a crowding-out effect for intrinsically motivated individuals. The findings point to a strong heterogeneity of responses among different population segments and different sustainable actions namely conservation, efficiency investment, and purchase of green electricity. Finally, the experiments provide evidence on the negative effect of ambiguity on sustainable behavior thus highlighting the importance of information and energy literacy.

Zusammenfassung

Dieses Projekt untersucht, wie individuelles Feedback und soziale Vergleiche in Verbindung mit finanziellen Anreizen den Stromverbrauch von Haushalten beeinflussen. Um die Auswirkungen der Massnahmen zu untersuchen, stützen wir uns auf ein Feldversuch mit einer Stichprobe von 201 Haushalten, die Privatkunden eines Energieversorgers im Berner Jura sind. Wir entwickeln eine digitale Schnittstelle und eine mobile App, die mit Verbrauchsdaten von Smart Metern gespeist werden, um Feedback zu den realisierten Einsparungen des Haushalts zu geben, die sorgfältig auf der Grundlage vergleichbarer Nutzerprofile aus einer Referenzgruppe von etwa 6.000 Haushalten berechnet werden. Die verwendeten ökonometrischen Modelle dienen dazu, die moderierenden Effekte einer Auswahl individueller Merkmale zu identifizieren. Außerdem analysieren wir das Engagement der Teilnehmer anhand von 25 Folgeinterviews. Schließlich führen wir fünf verschiedene Online-Experimente mit großen nationalen Stichproben von 291 bis 1.199 Befragten durch, um die Präferenzen verschiedener Haushaltstypen unter einer Auswahl unterschiedlicher Behandlungsmethoden zu ermitteln. Diese Experimente ermöglichen es uns, relative Tendenzen zwischen verschiedenen nachhaltigen Maßnahmen sowie die Auswirkungen von Unklarheiten bei Entscheidungen und Ergebnissen zu untersuchen.

Die Daten aus den Feldversuchen liefern kaum Hinweise auf einen statistisch signifikanten Effekt auf die erzielten Einsparungen. Obwohl statistisch nicht signifikant, deuten die Ergebnisse auf die Möglichkeit kontraproduktiver Auswirkungen von Interventionen in teilnehmenden Haushalten hin, die mit einer PV-Anlage ausgestattet sind. Die Ergebnisse deuten jedoch auf ein erhebliches Engagement für das Sparen und eine hohe Wertschätzung von Informationen unter den Teilnehmern hin. Die Experimente zur erklärten Präferenz bestätigen die Bereitschaft der Befragten, Strom zu sparen, und deuten auf einen signifikanten Effekt von Informationsrückmeldungen wie sozialen Vergleichen und Klimawandelbewusstsein hin. Die Informationen über finanzielle Einsparungen scheinen wenig Einfluss auf die Sparabsichten zu haben und könnten bei intrinsisch motivierten Personen einen Verdrängungseffekt hervorrufen. Die Ergebnisse deuten auf eine starke Heterogenität der Reaktionen verschiedener Bevölkerungsgruppen und verschiedener nachhaltiger Maßnahmen hin, nämlich Einsparungen, Investitionen in Effizienz und Kauf von Ökostrom. Schließlich liefern die Experimente Belege für die negativen Auswirkungen von Unklarheiten auf nachhaltiges Verhalten und unterstreichen damit die Bedeutung von Informationen und Energiekompetenz.

Résumé

Ce projet étudie comment les informations individualisées et les comparaisons sociales combinées à des incitations financières influencent la consommation d’électricité des ménages. Pour étudier l’impact des interventions, nous nous appuyons sur une expérience de terrain menée auprès d’un échantillon de 201 ménages privés clients d’un fournisseur d’énergie du Jura bernois. Nous développons une interface numérique et une application mobile alimentées par les données de consommation provenant de compteurs intelligents afin de fournir des informations sur les économies réalisées par les ménages, estimées à partir de profils d’utilisateurs comparables issus d’un groupe de référence d’environ 6 000 ménages. Des modèles économétriques sont utilisés pour identifier les effets modérateurs d’une sélection de caractéristiques individuelles. Nous analysons également l’engagement des participants ’au travers de 25 entretiens qualitatifs. Enfin, nous menons cinq expériences en ligne sur de larges échantillons nationaux allant de 291 à 1 199 répondants afin d’identifier les préférences de ménages soumis à différents traitements. Ces expériences nous permettent d’étudier les préférences entre différentes actions durables, ainsi que l’effet de l’ambiguïté dans les choix et les réactions.

Les données issues des expériences de terrain fournissent peu de résultats statistiquement significatifs sur les déterminants des économies réalisées. Cependant, les résultats laissent penser qu’il existe des effets contre-productifs des interventions pour les ménages équipés d’installations photovoltaïques. De manière générale, on note un engagement significatif en faveur des économies d’énergie et une grande appréciation des informations parmi les participants. Les expériences portant sur les préférences déclarées confirment la volonté des répondants d’économiser l’électricité et suggèrent un effet significatif des informations telles que les comparaisons sociales et la sensibilisation au changement climatique. Les informations sur les économies financières semblent au contraire avoir peu d’effet sur les intentions d’économiser et pourraient entraîner un effet d’éviction pour les personnes intrinsèquement motivées. On constate encore une forte hétérogénéité des réponses entre les différents segments de la population et les différentes actions durables, à savoir la sobriété, les investissements dans l’efficacité énergétique et l’achat d’électricité verte. Finalement, les expériences suggèrent un effet négatif de l’ambiguïté sur les comportements durables, soulignant ainsi l’importance de l’information et des connaissances dans le domaine énergétique.

Main findings («Take-Home Messages»)

- We observe no detectable realized savings that can be linked in a statistically significant manner to any of our interventions (social-comparison feedback or financial incentives). While being statistically insignificant, the estimated treatment effects point to a possibly counterproductive effect of social feedback and financial incentives for PV owners.
- Normative messages such as social-comparison feedback can be used to motivate individuals' engagement for electricity conservation. However, uniform feedback strategies and consumers' heterogeneous responses could induce boomerang effects, especially for intrinsically motivated people and those with financial constraints.
- Online experiments with large and representative samples highlight the detrimental effects of uncertainty and ambiguity in engaging in sustainable behaviors such as electricity conservation. Clear information and simple messages are likely to be more effective than those outlining multiple benefits and factual details.
- Estimation of both potential and realized savings at the individual level poses important challenges. Reasonably accurate estimates require a large sample of participants, thus highlighting the importance of legal barriers for accessing confidential consumption data.

Contents

List of abbreviations

1D-CNN	One-dimensional Convolutional Neural Network
AI	Artificial Intelligence
CoSi	Co-Evolution and Coordinated Simulation of the Swiss Energy System and Society
CNN	Convolutional Neural Network
DCE	Discrete Choice Experiment
DCE-2023	DCE conducted in 2023 (integrated in SHEDS 2023)
DCE-2024	DCE conducted in 2024 (prior to field experiments)
DiD	Difference-in-Differences
DSRM	Design Science Research Methodology
FEx	Field Experiment
HES-SO	University of Applied Sciences and Arts Western Switzerland
HH	Household
INFINEED	Project: "*The Interplay of Feedback and Incentive Effects on Electricity Demand*"
IRENE	Institute of Economic Research, at University of Neuchâtel
La Goule	Société des Forces Électriques de La Goule
LSTM	Long Short-Term Memory
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
PV	Photovoltaic
RP	Revealed Preferences
S+F Group	The experimental group treated with a combination social-comparison feedback and financial incentives
SEP	Swiss Energy Par
SFOE	Swiss Federal Office of Energy
SHEDS	Swiss Household Energy Demand Survey
SM	Smart Meters
SMART	Project: "*The Smart Meter Rollout as an Experimental Setting: Promoting Energy-Saving Behaviors with Target Group Specific Interventions*"
SP	Stated Preferences
SWEET	Swiss Energy research for the Energy Transition
UniNE	University of Neuchâtel
WP	Work Package

1. Introduction

1.1. Context and motivation

Despite a large and growing literature on non-price interventions such as social comparison feedback, there are significant research gaps in identifying causal effects and in distinguishing the mechanisms underlying specific responses (Andor & Fels, 2018; Schwartz et al, 2019). For instance, in the case of commitment devices (present-day choices that restrict future options), most studies use a combination of treatments, failing to disentangle the specific effect of commitment. Moreover, many studies examine hypothetical responses in discrete choice experiments (DCE) rather than in real setups. The data on Swiss households' conservation effort is not conclusive. Using historical data over five years, Mari et al. (2025) suggest that national mobilization has been effective in increasing conservation efforts, but the realized saving has been limited to periods (such as winter) when consumers have more saving opportunities. Similarly, Cellina et al. (2024) report a significant but temporary effect of their energy saving app on a sample of 55 self-selected treated households. In both cases, the realized electricity saving is below 5% and limited to a few months.

Helping people achieve sustained energy conservation behaviors through the use of nudges has received mixed results (Caballero & Della Valle, 2021), prompting researchers from different disciplines to find explanations. One such explanation points to psychological values modulating long term commitment and consumption habits (Puntiroli et al., 2022). Recognizing the importance of peer effects, many studies identify the observation of other community members and normative social influences as a main channel for behavioral change (e.g., Wolske et al., 2020). It has been shown that financially constrained consumers are more likely to disengage from comparison-based information that highlights their relative disadvantage (Hamilton et al., 2019). In addition, social comparison feedback has been shown to produce counterproductive responses among some consumers; for example, in an energy conservation context, households who were already consuming less than the norm increased their energy use after receiving descriptive norm feedback, a phenomenon known as the boomerang effect (Schultz et al., 2007).

Concerning monetary incentives, literature suggests that small but salient price signals could be effective (Wagner, 2017). Comparing monetary incentives with nudges, a few studies (Holladay et al., 2019; Weber et al., 2017) point to potentially significant effects of both measures. There is, however, little research about the effect of policies combining multiple interventions, particularly on the interplay between non-monetary and monetary incentives (Schwartz et al., 2019) and their long-term effects. In line with the Behavioral Reasoning Theory that links actions to psychological motives (Claudy et al., 2013), we expect that awareness of potentials for energy saving and how it can be achieved have a positive effect on sustainable behavior. Monetary incentives might cause a motivation-crowding-out effect by "legitimizing" antisocial behavior (Lanz et al., 2018), hence, undermining the individual's intrinsic motivations. On the other hand, they could provide a crowding-in effect, should they provide a salient price signal prompting the actions in the right direction (Wagner, 2017; Brandon et al., 2018).

1.2. Project objectives

The project has three interdependent objectives, each linked to a specific research question and addressed with its own tailored scientific approach. First, we aim to investigate sustainable actions from different perspectives. We focus on a more complex interplay between various sustainable behaviors and human psychology. In particular, we study conservation behavior, efficiency investment, and the adoption of green electricity. Considering various sustainable actions with different costs and information requirements allows us to accommodate heterogeneous preferences, including for instance low-income groups interested in conservation, as well as affluent households opting for high-cost/high-information actions such as efficiency investments. This novelty helps assess inclusive policies based on monetary and non-monetary incentives and to adapt them to the individuals' own information requirement levels.

Second, we assess the impact of our interventions on actual and long-term behavior. Following the literature (e.g., Andor & Fels, 2018), we focus on a selection of non-monetary interventions that target various cognitive biases. Specifically, we consider information cues (via individual feedback), norm-based interventions (via social comparison), and commitment (via goal setting). Among these three categories, social comparisons are the most investigated, and reportedly the most effective intervention-type, linked to significant reductions in energy consumption. That is followed by the goal-setting interventions that are also considered promising. We assess how repeated interaction with this non-monetary information impacts behavior over time.

Finally, an important objective is to identify the effects of personal characteristics such as energy literacy, information overload as well as pro-environmental values on the type of sustainable behavior (or lack thereof). Our past research (Schubert et al., 2022) suggests that habits and routines play an important role in energy demand, while energy literacy is not necessarily correlated with higher intentions (Farsi et al., 2020). This observation, in conjunction with previous disappointing observations on the intention-behavior gap (Park & Lin, 2020) raises the question about how to best engage consumers in the desired behavior thanks to the information at hand. We argue that two key mechanisms are at play here. The first one concerns consumers' personal level of desired information and literacy. Thanks to our surveys, we are able to test how to best calibrate the quantity and type of information to communicate. The second mechanism concerns consumers' engagement. Our past research (Holzer et al., 2020) demonstrates that properly designed feedback motivates individuals to engage in a desired behavior. We implement a feedback mechanism (smartphone app) to deliver personalized information in order to trigger sustainable action aligned with the consumers' energy consumption goals.

2. Approach, method, results and discussion

We organize this section according to the project's work packages: understanding electricity consumption patterns and saving potential (WP0), the analysis of user experience (WP1), stated preferences and user surveys (WP2), and finally revealed preferences and field experiments (WP3). Each WP has specific methods that are explained below. While WP0 relies mainly on predictive models such as Machine Learning (ML), other WPs test various hypotheses, under two generic treatments, social comparison (alone) or a combination of social comparison and financial rewards. WP1 uses a mixed (qualitative and quantitative) analysis to assess the effectiveness of the goal setting device in users' engagement. WP2 deals with realized savings relying on econometric methods while emphasizing the moderating effects of various behavioral factors. Finally, WP3 provides an econometric analysis of the treatment responses in a series of experimental setups including preferences over different electricity-saving actions, saving goals, and ambiguity, which we define as the presence of multiple plausible interpretations of a single outcome. Below, we provide a summary of main methods and results in each WP.

2.1. WP0: Understanding consumption data and saving potential

The objective of WP0 is to better qualify the problem to be investigated in the project. The main goal is to validate the methodological approach required for providing meaningful feedback to households based on smart meter readings. In particular, we conduct an in-depth analysis of real user data that we obtain from our implementation partner (La Goule). These data include hourly readings from about 5,800 SMs over a period of five years, that are be used in combination with six available waves of household data (SHEDS 2016-2021) to identify distinct household profiles based on demographics as well as electricity consumption patterns. This WP seeks to answer two main questions: First, to what extent do data quality issues risk undermining the accuracy of savings calculations for field experiment participants? Second, can machine learning or econometric models provide a reliable estimate of a household’s electricity saving potential given the limited sample size and available household characteristics?

2.1.1. Understanding consumption data

Understanding consumption data is important for the estimation of both the realized savings relative to a baseline and the saving potential compared to the “best practice”. Smart meter records of electricity consumption pose two issues that could affect the savings estimation. First, missing observations cause gaps in time series. Missing values can be imputed by various methods, ranging from simple interpolation of similar periods to more elaborate machine-learning (ML) models. The second issue is the presence of zeros, which mostly correspond to cases of solar photovoltaic (PV) owners that are self-sufficient during certain periods of time and are therefore genuine zeros. Unfortunately, with the data available to us, we cannot draw clear conclusions about the small share of remaining zeros for non-PV households.

We also point out that the consumption data used in this project is strictly anonymized. For all households, we therefore have electricity consumed every 15 or 60 minutes (depending on the smart meter type). For most households, this is the only information available to us. For the relatively small sample of participants in the online survey, we were able to match their responses with their consumption, as they were informed and gave their consent.

Smart meter readings contain a non-negligible share of missing values. In our initial dataset (available from the beginning of the project), covering three years (2020-2022) with strictly anonymized consumption data measured every 15 or 60 minutes for each household (all private clients of the electricity provider La Goule), the share of missing, negative or zero consumption amounts to about 20%. Aggregating the data to a daily panel reduces the zero consumption values to about 4% of the observations. We use 15-minutes consumption data of a subsample of 212 households from this dataset to assess the performance of various models. Details of data and methodology are reported in John (2023). Overall, our assessment favors long short-term (LSTM) models (with a mean absolute percentage error of 20%) over simple interpolation models (mean absolute percentage error of about 30%). While the missing data might remain an important issue in high-frequency estimates such as hourly consumption, the impact on the estimates of daily consumption is limited to a negligible share of the sample.

In order to confirm that the information communicated to the field experiment participants, namely daily consumption and realized weekly savings, would not be impacted by missing values, we also studied strictly anonymized smart meter readings from about 5,779 households over five years (August 2018 to August 2023) with approximately equal shares of 15-minutes and hourly readings. We observe that approximately 20% of records show no change compared to the previous readout, hence zero consumption during a 15-minute or 1-hour period. Our assessments indicate little effect on the estimation of realized weekly savings and suggest that a small share of daily consumptions could genuinely be zero.

In the ML analysis of saving potential (described below), which has been carried out with a later release of data from January 2023 to June 2025, the small sample size does not allow effective training of LSTM models. This dataset is limited to 162 households with non-missing characteristics collected via a survey before the field experiment. Here, we impute the missing values (about 9% of the raw data) by using the average distribution of values from the preceding and following weeks to allocate the known cumulated consumption over the sequence of missing observations. Missing values are not imputed if the sequence is longer than 48 hours. As for zero values, the ML model does not seem to be affected by their presence, although it does not predict them as being absolutely zero, but usually small values close to zero. More details are provided in **Appendix WP0**.

2.1.2. Saving potential

We use two methodologies for estimating individual saving potential, ML models and econometric models. While focusing on ML models for smart meter data, we use econometric models as complementary analysis to assess the challenges related to segmentation of different household profiles based on characteristics usually available in surveys. The ML results and specifications are reported in Appendix WP0. The econometric approach is reported in **Farsi & Weber (2024)**.

ML models can be developed to provide electricity consumption forecasts that can be used to give information to households on their electricity saving potential. On that ground, we created an ML model to forecast electricity consumption over the coming week at the household level. The strategy followed to construct and train such a model relies on a one-dimensional convolutional neural network (1D-CNN, see Kiranyaz et al., 2021) that takes a multiple time series of household-level features, time-variant context inputs such as weather and climate, and past electricity consumption as inputs. The model learns hidden features from the input data to predict one week of electricity use by minimizing a loss function. Data from 162 households collected over a 2.5-year period (January 2023 to June 2025) have been used to train and test the model. Data could not be used over the whole period for all households because of missing values.

In the present context, a quantile loss function has been used, so that the model is trained to predict quantiles of the output series. The idea behind this choice is that lower quantiles can represent targets for saving electricity. The model that has been trained is hence able to provide forecasts for the 30th, 50th, and 70th quantiles, with the difference between the 30th and 50th quantiles acting as saving potential. Different observation frequencies have been tried for both input and output series: hourly, 3-hour, and 6-hour time steps. High frequency data do not seem to provide significantly better results, in part because of the random nature of some peaks occurring in electricity consumption. Using lower frequency input and output data helps smooths peaks and troughs, hence stabilizing the model. The general outcome is that an ML approach can effectively be used to forecast electricity saving targets at household level. The quality of forecasts depends on the sample size and input variables, especially the household features, used for training the model. Nevertheless, the 1D-CNN that has been trained produces encouraging results, which supports the use of ML to forecast electricity saving potential.

The econometric analysis (detailed in Farsi & Weber, 2024, section 4.4) is based on a series of Quantile Regression models applied to household data from seven waves of the Swiss Household Energy Demand Survey (SHEDS) from 2016 to 2023. We focus on households' self-reported annual electricity consumption in kWh. After excluding outliers and invalid values, the sample size encompasses 6,302 households. Using three relevant determinants of electricity demand (4 household size categories x 3 heating system categories x apartment-vs-house indicator), we specified 24 household typologies. SHEDS data include many other factors and characteristics that could potentially be used, but our strategy based on 24 typologies ensures a reasonable number of households (ranging from 48 to 1,728) remain in each segment. After running several quantile regressions with various specifications for each of the 24 segments, we converged to a quantile frontier model at 20th percentile (1st quintile) of logarithm of yearly consumption, controlling for location indicators (city agglomeration, countryside, French-speaking region) and dwelling size, as well as household income and respondent's education. This specification gives the best overall goodness-of-fit while assuming 20% of households in each segment are "perfectly efficient" with negligible saving possibility. The saving potential for the remaining households is computed as their consumption difference with the first consumption quintile (the "ideal" household) in their respective segment.

The estimated savings suggest that a reasonable majority of households in each segment could achieve significant savings, with an average saving of 27 to 40 percent, depending on the segment. Part of these reductions might be related to factors beyond the households' control, such as outdoor temperature.

However, according to these estimates, most households (more than 2/3) can be expected to achieve 10 percent savings.

Our results from both econometric and ML models suggest that a reasonably accurate estimation of saving potential for an individual household requires large samples that can allow more training base for ML models, and a greater number of segments in econometric models. While pointing to a relatively important saving potential, the econometric analysis shows that strong heterogeneity exists, even within segments that are relatively similar based on usual survey variables and traditional determinants of electricity consumption. In comparison to econometric models, ML prediction models demonstrate two important advantages. First, ML models prove to be quite efficient with a limited number of household characteristics, thus overcoming data availability and confidentiality issues. Second, ML models can readily be applied to large datasets as well as frequent data emanating from smart meters, without any serious issue caused by missing values.

Given that household characteristics are not available for non-participants, the relatively small sample of field experiment participants does not allow us to create meaningful uniform segments for saving potential estimations. We therefore avoid providing individual feedback about a specific household's saving potential but instead use generic feedback powered by a Generative AI system (as described below in WP1).

2.1.3. WP0's key takeaways

Missing values and zeros in smart meter readings, while substantial in the raw data, do not meaningfully distort daily or weekly consumption estimates and savings calculations. Therefore, daily aggregation is used to deliver reliable participant feedback. The model employed demonstrates that a machine learning approach can forecast a household's saving potential on the basis of (anonymized) consumption history and weather data alone, without requiring detailed survey variables. The econometric analysis suggests that most households have a reasonable ability to reduce consumption: over two-thirds of households can realistically achieve savings of at least 10%, with average saving potential ranging from 27% to 40% across segments. However, the small number of households with full characteristics available prevents the calculation of statistically reliable individual saving targets, justifying the use of generic rather than personalized feedback in the field experiment.

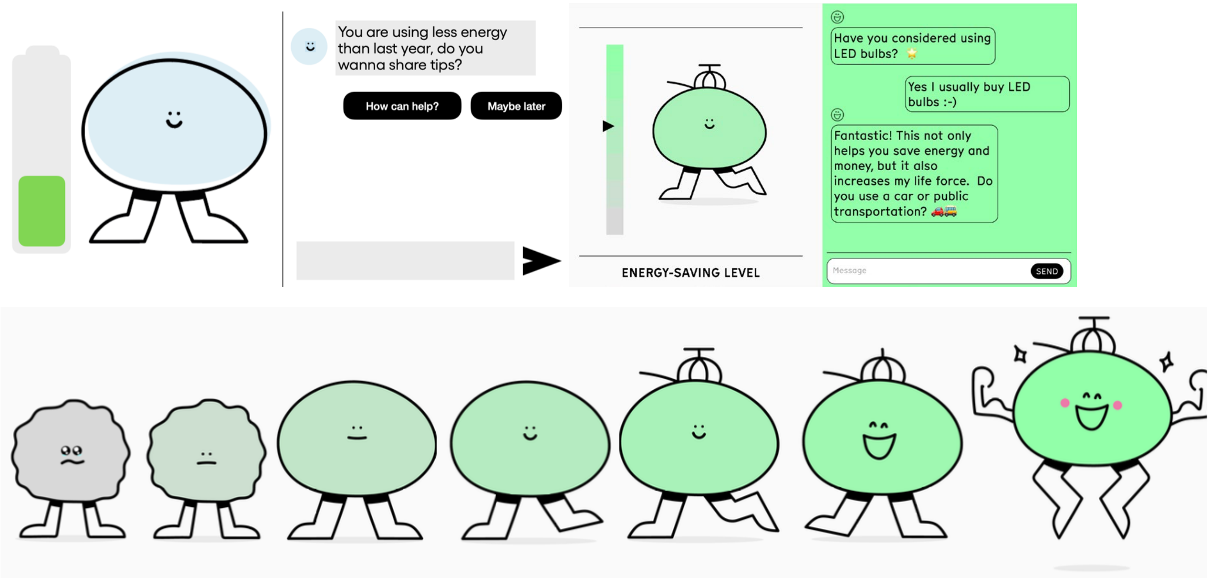
These findings point to two important policy implications for upscaling conservation programs based on individual feedback. First, missing values in smart meter readings do not cause any serious issues insofar as individualized feedback messages are based on daily (or lower frequency) consumptions. Second, accurate estimation of saving potential at individual level presents an important challenge, making it difficult to use as a basis for personalized feedback.

2.2. WP1: User experience analysis

WP1's main objective is the design and implementation of the digital artifact to present smart meter data in a user-friendly way and to provide dynamic feedback to electricity consumers. The focus is on the user experience dimensions of the experiment. The WP also investigates how the design and user experience of a digital eco-feedback interface influence household engagement and electricity-saving behavior. The ultimate goal is to identify actionable design principles that maximize the effectiveness of digital behavioral interventions. The analysis is conducted in two stages: a pre-field experiment analysis aiming at a participatory design of the digital interface, and a post-field ex analysis to measure the impact of user experience on saving outcomes. Specifically, we analyze 1) how participatory design methods can be used to create a feedback system that fosters emotional attachment and sustained user interaction; and 2) to what extent different user experience dimensions and varying household profiles mediate the relationship between social feedback conditions and actual energy savings.

2.2.1. Participatory design and pre-field experiment analysis

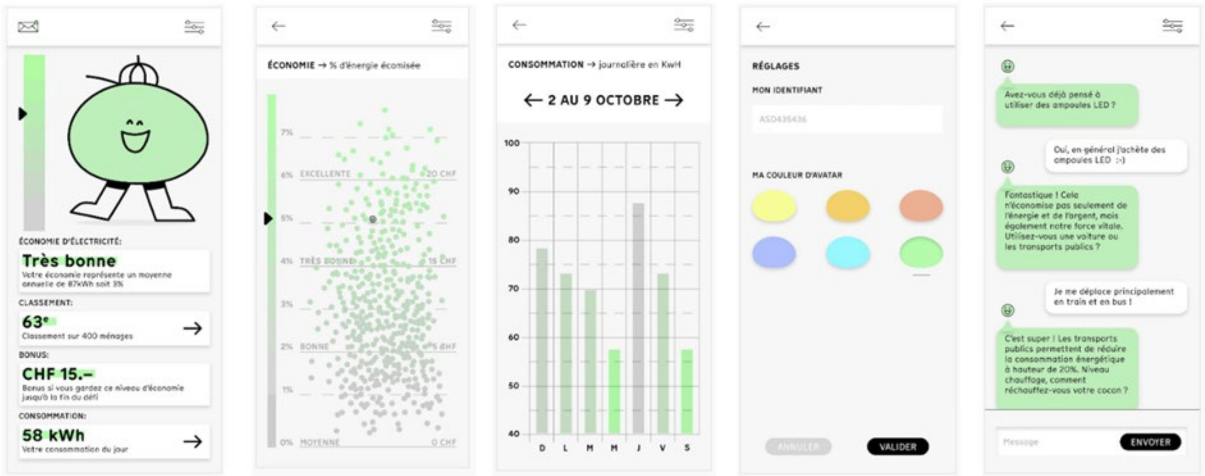
We initiated the participatory design by a first half-day Co-Design workshop to better understand electricity consumption behavior and saving goals, and to collect users' ideas over potential feedback design solutions to help consumers reach these goals. The workshop generated discussions and artifacts (e.g., drawings, voting, post-its), which were collected and analyzed. The results, combined with insights from scientific literature on eco-feedback were used to design a first concept for the INFINEED system. In particular, we chose to pursue a care-based approach, where energy-saving is shown as the lifeforce of a digital avatar, called INFI, similar to a "digital pet", hence a positive metric for reducing consumption while creating some emotional attachment to the feedback system (see Figure below).



The design also includes a conversational interaction feature, powered by a Generative AI system, to convey information to consumers about energy saving tips in a piecemeal fashion rather than through a potentially overwhelming fact sheet. We completed the process with an iterative and incremental agile development approach to validate and improve the design through extensive pilot testing with online participants followed by analyses and discussions with designers and researchers. For instance, we have iterated through online tests with 150 users to adjust the conversational interaction through prompt engineering. We then evaluated how the care-based designs enhanced with conversational interactions can increase emotional attachment and potential energy saving behavior through a larger online test (N=450). Our preliminary findings suggested that our approach enhances emotional attachment and subsequently increases users' willingness to adopt energy-saving actions.

A second participatory design workshop was conducted with 12 stakeholders at La Goule, where we explored various design aspects of the mobile app, including different conditions such as monetary incentives and social scoring. The workshop highlighted that statistical data like the percentage of electricity saved or kWh might be challenging for users to understand. Therefore, it is essential to present qualitative textual information on electricity savings (e.g., "Excellent savings," "Good savings," "Moderate savings," or "No savings"). Nonetheless, it was agreed that including kWh information remains important for raising user awareness about their energy consumption.

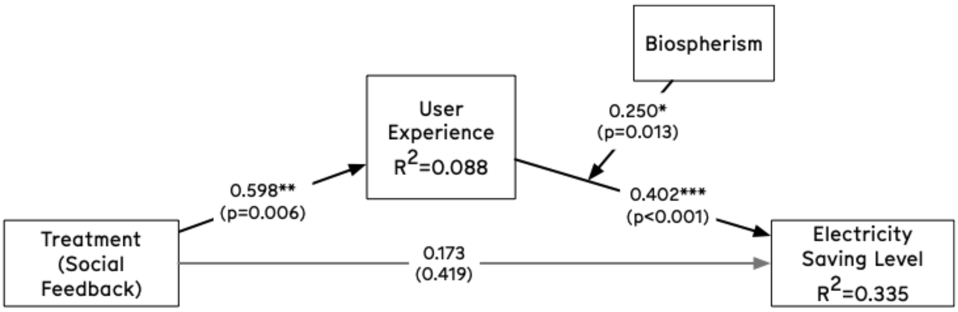
The Figure below shows the final mockups of the mobile app for the full treatment condition (with social and financialincentives,). The home screen (first screen on the left) shows Infi, the user's avatar that represents the level of energy saved. The more energy is saved the more life energy Infi has. Next to Infi is a gauge that shows the level of energy saved in an abstract sense (we removed numbers/percentage from the gauge following feedback from the workshop). This screen also summarizes the level of savings qualitatively (e.g., "Très bonne"). Then it offers a section on social ranking that shows the user's position compared to similar households (e.g., 15th among 52 households). Then a section on monetary rewards is displayed (e.g., CHF 10) to indicate how much reward the user would get should they continue saving as they have done so far. Finally, there is a standard consumption section that shows the number of kWh consumed. When clicking on the first three sections users get to the second screen where the savings are displayed visually on a scale with visualizations of thresholds that combine percentages, qualitative indicators, social comparison and monetary rewards. When clicking on the last section, the user gets to the third screen which shows a standard electricity consumption graph that users expect to have on a smart meter app. In addition to these core features, users can change the color scheme of the app in the settings pane (fourth screen) by pressing the settings button (top right of the home screen) and a chat screen (fifth screen) by pressing the chat button (top left of the home screen).



The digital interface enables users to access their weekly electricity consumption data by linking their app login information to their utility's customer account. For participant updates and communication, we opted to use SMS as the primary channel. Every week, we send batch SMS messages to all participants to inform them of updates regarding their electricity consumption and other relevant information. The participants had also access to a dedicated website (https://infineed.app) with information about the project with various communications.

2.2.3. Post-field experiment analysis

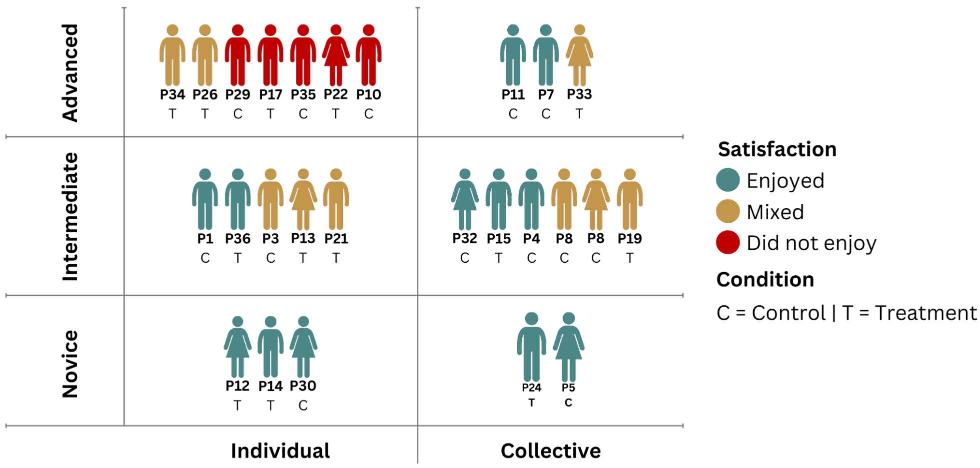
In the post-field experiment analysis, we focus on the first wave (Summer edition) of the experiment, which took place from August 19 to September 30, 2024. We focus on participants who answered the usability questions (N=68), looking at the final energy savings as displayed on the app, the user experience as measured by the Attrakdiff questions (a UI evaluation tool), and the environmental concerns (biospherism) as measured by Universalism–Nature dimension of the Schwartz Value Survey. The Figure below presents the results of our structural path model. To complement this analysis, we conducted semi-structured interviews with 25 participants to gain deeper insight into how they integrated the system into their households and to better understand the factors shaping their experience.



SmartPLS model illustrating the mediation of social feedback effects on energy savings through user experience (UX). Path coefficients are standardized; significance levels: $*p < .05$, $**p < .01$, $***p < .001$.

Our structural path model, run on the SmartPLS software with a bootstrap of 5,000 draws, reveals that social feedback conditions, combining financial incentives and social comparison, improved electricity savings indirectly by enhancing user experience. Rather than affecting behavior directly, social feedback made the system more engaging, motivating, and emotionally meaningful, which in turn led to greater savings. Importantly, these effects operated through hedonic and emotional dimensions of user experience rather than pragmatic aspects such as ease of use or clarity. This finding underscores that how information is presented matters as much as the information itself.

The impact of user experience on savings was amplified by environmental motivation. Users with stronger biospheric values were significantly more likely to translate a positive experience into actual electricity savings, suggesting that eco-feedback systems are particularly effective when they resonate with users' pre-existing environmental concerns. This points to the value of framing energy-saving interventions in terms of environmental benefits rather than purely financial terms, particularly for environmentally motivated segments of the population.



Nevertheless, the interviews revealed that environmentally oriented participants represented a heterogenous group and as such a one-size solution does not fit all. Indeed, we observed three distinct user profiles, each with different needs and responses to the system. **Novice households**, those with limited prior engagement in electricity-saving behaviors and no access to consumption feedback, benefited most from the intervention. For these users, simple tips and feedback increased awareness and triggered concrete actions such as installing power strips, using eco-modes, and air-drying clothes. They appreciated the system's simplicity and often engaged with it in a playful manner, treating weekly updates as a household ritual. **Intermediate households**, those already implementing some electricity-saving behaviors but lacking consumption feedback, valued the system primarily as a reminder. These users often knew the advice already but appreciated being nudged to maintain good practices they did not consistently apply. As one participant noted, the system provided *"more discipline and the will to make an effort than actual advice."* **Advanced households**, those already implementing electricity-saving behaviors and with access to consumption feedback through existing monitoring systems, presented the most polarized responses. While some maintained a constructive orientation, some faced a "saving ceiling" and became frustrated when rankings highlighted limits beyond their control. These users found the basic tips patronizing and felt the system failed to acknowledge their existing efforts.

The divergent reactions to information presentation deserve attention. Participants valued diverse features for accessing information: graphs for social comparison, weekly tips, and the avatar visualization. However, the same presentation style elicited different reactions depending on user expertise. Novice users appreciated simple, easy-to-understand information and actionable tips that felt empowering. Advanced users found the same simplicity patronizing, perceiving that the system treated them as uninformed and did not respect their existing knowledge. This tension suggests that adaptive interfaces, offering different levels of detail and sophistication based on user profiles, could better serve diverse populations.

Our analysis of the interviews also revealed two different household dynamics in how users engage with eco-feedback systems. **Collective use**, discussing feedback within the household, was associated with overall positive engagement across households from novices to experts. In contrast, **individual use**, where sole "energy managers" (often men in traditional household configurations) was associated with lower levels of motivation in more advanced households - particularly in expert households.

These findings carry implications for practitioners designing eco-feedback systems. First, user experience appears to be an important aspect of the system, possibly contributing to motivation for behavior change. User experience leverages two main components, a utilitarian aspect and a hedonic aspect. Both these components appear to be crucial. Our pre-field experiment analysis (described above) shows that emotional attachment through an avatar could enhance energy saving intentions. Here we show that adding social feedback further improves the user experience. Nevertheless, the utilitarian facet of user experience should not be forgotten. In particular, our interviews revealed that when it comes to energy savings, households at different levels of engagement might have different information needs. A tiered approach to information presentation would better serve diverse users: simple tips and encouraging feedback for novices, regular reminders and progress tracking for intermediary users, and granular data with contextual benchmarks for advanced users who need acknowledgment of their existing efforts rather than basic advice.

Finally, the system we designed focused on a single user through a mobile app. However, our interviews revealed that collective use of the system was associated with a particularly positive experience. As such, systems could be designed to further explore how collective household dynamics could be fostered.

2.2.4. WP1's key takeaways

When building digital feedback apps for energy saving, designers should focus on emotional user interface features like avatars and encouraging words rather than just accurate charts: feeling engaged might matter more for savings than ease of use. A one size fits all solution might not be adequate. Novice users appreciate getting simple practical tips while advanced users need to receive more detailed advice on fine-tuning their consumption. To foster shared household discussions and engagement, the app could include specific feature such as a shared screen, with group challenges that encourage everyone in the household to participate together. Finally, for the highest savers, replace competitive rankings with supportive messages to maintain motivation without the frustration of hitting a savings ceiling. Finally, users expect real time feedback where the app reacts instantly when an appliance is turned on or off.

From a policy perspective, the findings point to important challenges for upscaling promotion programs based on digital feedback. These challenges stem from the strong heterogeneity of individuals in their emotional response and knowledge. Therefore, one-size-fit-all solutions can be counter-productive for some user segments, hence, should be avoided.

2.3. WP2: Stated preferences and online experiments

The main objective of WP2 is to implement a discrete choice experiment to identify the households' tendency to sustainable behaviors under two main policy treatments (social nudges and financial information) and their interaction. The policy treatments follow the intervention features designed in WP1. The survey provides the basic household characteristic inputs to the saving potential algorithm developed in WP0. Finally, an important secondary objective is to identify the mediating effect of personal characteristics such as energy literacy, perceptions of information overload as well as pro-environmental values on the type of sustainable behavior or lack thereof.

In this WP, we examine how different forms of normative interventions influence households' electricity-saving preferences and prosocial behaviors and aim to understand how ambiguity weakens these effects. Specifically, we ask: Which types of social norm messages are most effective at motivating energy conservation, and for which consumer segments? How do dynamic norms framed around energy security differ from those framed around climate change in shaping saving goals? And to what extent do different forms of ambiguity reduce engagement?

Two definitions are warranted here:

- **Dynamic social norms:** refer to an evolving trend in attitudes and behaviors in a group. In contrast to static norms that stem from what others do now, dynamic norms refer to emerging shifts and highlight that a positive change is underway.
- **Ambiguity:** In line with psychological literature, we define ambiguity as situations, stimuli, or information that lack a single clear meaning, and are open to multiple interpretations. This definition is related, but not the same as, the definition used in economic literature, that is, situations that involve unknown probabilities.

2.3.1. Experiments design

In WP2, we developed two discrete choice experiments (DCEs) and three incentivized experiments. All experiments were embedded in online surveys. The overall objective is to examine the effectiveness of normative interventions across different contexts related to energy conservation and sufficiency. The first discrete choice experiment (DCE-2023) was conducted on a subsample of 621 respondents in SHEDS 2023. The experiment design aims at identifying the effects of social comparisons and financial information on preferences for three sustainable actions: conservation, efficiency investment, and purchase of green electricity. The treatments include two social comparison feedbacks: one framed in terms of electricity consumption (kWh) and the other combining social comparison with financial information expressed in monetary savings (SFr).

The second discrete choice experiment (DCE-2024) was conducted on 291 customers of *La Goule*, who agreed to participate in the upcoming field experiment. DCE-2024 was implemented before the two waves of the experiment that attracted about two thirds of the DCE respondents. Informed by the results of DCE-2023, we adopted a simplified design focusing on electricity conservation. Drawing upon an emerging literature, this experiment investigates the role of dynamic norms (i.e., normative information about how behaviors or attitudes are changing over time), which emphasize a transition toward greater commitment to electricity savings, rather than describing the current average behavior. Two drivers of this evolution are examined: growing concerns about (i) energy supply security and (ii) climate change.

The three incentivized experiments focus on the role of ambiguity, which refers to as the presence of multiple plausible interpretations of the experimental stimulus, or the provided information about the respondent's action. Ambiguity is likely to be a key factor moderating both individuals' engagement in sustainable behaviors and the effectiveness of normative interventions. The objective is to assess how different forms of ambiguity affect prosocial behaviors in contexts relevant to energy transition. The external validity is addressed by linking the responses to global electricity consumption and/or carbon emissions, thus representing a realistic, albeit small, impact. The incentivized experiments also investigate whether social norms can help mitigate the detrimental effects of ambiguity and uncertainty in different aspects of sustainable behavior. Specifically, we explore ambiguity regarding (i) non-financial motives, (ii) tradeoff between private and social benefits, and (iii) private costs and individual budget constraints. These experiments were carried out on representative samples of the Swiss population, with 793, 843, and 1,199 respondents.

The methods and results of the two DCEs are respectively reported in Giauque et al. (2025) and Giauque and Farsi (2026). The incentivized experiments are presented in Appendix WP2 with an extensive discussion of study design, collected data and the results of each experiment.

2.3.2. Results of Discrete Choice Experiments

The results from DCE-2023 indicate that the effectiveness of social comparisons is limited to conservation among households with below-average electricity consumption. Additionally, the results suggest a potential crowding-out effect of financial information. First, adding monetary savings to social comparisons adversely affects the below-average consumers away from sustainable behavior. Secondly, individuals who perceive conserving energy as important react negatively to financial information, possibly because such information conflicts with their intrinsic or moral motivations. For households with above-average consumption—those with the greatest saving potential—both types of social comparisons increase preferences for taking no action, though the inclusion of financial information slightly mitigates this negative response.

Overall, DCE-2023 illustrates that incentivizing households' behavior could be challenging, with a limited effect for social comparisons and a risk of crowding-out effect for financial information. These results prompted us to explore alternative interventions that leverage the power of social norms by placing greater emphasis on social change rather than static comparisons.

DCE-2024 shows that the effectiveness of dynamic norms depends on how they are framed. Framing growing commitment to electricity saving as a response to energy security concerns enhances preferences for saving goals and effort on average. In contrast, framing it as a response to climate change does not increase average saving intentions. However, the climate framing affects behavior in a different way. It changes how individuals are distributed across distinct behavioral profiles (idealist, ambitious, engaged, and minimalist). In other words, it influences which types of preference patterns are more common, rather than changing preferences within each group.

Another key observation from DCE-2024 is the participants' fairly high stated saving targets with a median of 5% and average of about 8% reduction in electricity consumption. About two thirds of these respondents have later participated in the field experiment. As we will see later, these results do not provide any evidence that these saving targets can be realized in practice.

Finally, the results show the asymmetric impact of dynamic norms in the energy-security framing: it activates more ambitious (>10%) saving goals primarily among individuals theoretically identified as less motivated and encourages less ambitious (<10%) saving goals for those identified as more concerned. From a policy perspective, this asymmetry is crucial. Dynamic norms framed around energy security appear effective in mobilizing less-engaged households, thereby broadening participation. However, they are less effective at increasing ambition among already committed individuals. As such, these interventions may be well suited to expanding the base of participation in electricity saving, but complementary measures may be required to sustain or deepen ambition among the most motivated households. These findings could translate into a useful segmentation strategy, favoring security-framed messages for broad outreach campaigns to attract new audiences, while shifting to peer challenges, to maintain higher savings targets in follow-up communications for existing participants.

2.3.3. Results of ambiguity experiments

The findings from the incentivized experiments indicate that ambiguity undermines engagement in sustainable behaviors. While consistently confirming an average effect across three experiments, the data suggest that the magnitude of the effect varies substantially depending on the source of uncertainty the individual faces. Moreover, the ability of normative interventions to counteract the ambiguity effect varies substantially across studies.

The first study (793 participants) examines ambiguity about the moral meaning of individual actions, that is, situations in which it is unclear whether a choice can be considered fair or socially appropriate. When this link is blurred, prosocial behavior declines markedly. Clear and explicit normative messages emphasizing fairness substantially reduce this effect, restoring a large share of the lost engagement. The results further show that ambiguity is particularly influential among individuals who either strongly recognize fairness norms or hold strongly self-interested preferences. Normative interventions are most effective for individuals with low to moderate self-interest, while highly egoistic individuals remain relatively unresponsive.

The second study (843 participants) focuses on ambiguity surrounding the social benefits of a sustainable action versus private benefits of inaction, in a context where participants decide whether to offset emissions associated with AI use. This form of ambiguity produces the largest reduction in prosocial behavior. While normative messages partially mitigate this effect, their overall impact remains limited. Responses to ambiguity and norms also differ strongly across individuals. For those who do not believe that others would act prosocially, ambiguity substantially discourages engagement. By contrast, individuals who already believe that others contribute are largely protected from the negative effects of ambiguity. Normative messages can even backfire among intrinsically motivated individuals, while they tend to be more effective for individuals with more self-interested preferences.

The third study (1,199 participants) examines ambiguity about private costs, such as uncertainty about the budget that can be allocated to prosocial action. This form of ambiguity has a relatively small average effect on behavior. However, in this context, social comparison messages are highly effective. Providing information about others' behavior not only offsets the negative effect of ambiguity but fully compensates for it. Ambiguity also increases positive emotions among those who act prosocially, such as pride and a sense of belonging. These emotional responses are reduced when normative information is provided, suggesting that norms can limit the additional emotional rewards associated with acting under uncertainty.

Taken together, these findings point to a clear pattern. Ambiguity is most harmful when it obscures the moral trade-off between private and social benefits, and less harmful when it mainly affects affordability or feasibility. The effectiveness of normative interventions follows the opposite pattern: norms are most powerful when ambiguity leaves room for interpretation, rather than when it fundamentally weakens moral evaluation. Importantly, ambiguity and norms do not affect all individuals in the same way. Strong normative beliefs can buffer against ambiguity, while weak norms and egoistic preferences amplify its negative effects. Overall, the results indicate that ambiguity represents a robust barrier to prosocial engagement in environmental and energy-related contexts, but that its effects can be partially mitigated through carefully designed normative interventions. From a policy perspective, these results carry direct implications for designing communication strategies, suggesting that effective strategies should first seek to reduce ambiguity wherever possible and, when ambiguity cannot be eliminated, provide the relevant type of normative message.

2.3.4. WP2's key takeaways

The experimental results obtained in WP2 point to several policy implications that can be helpful in designing social norm messages to encourage electricity saving. These conclusions can be summarized as follows. It is however important to note that these conclusions are based on small-scale experiments and should be considered with caution:

- Combining consumption comparisons with monetary savings figures could be subject to a crow-out effect that may demotivate environmentally conscious users.
- Dynamic social norm messages that highlight a growing trend of people committing to save energy (e.g., "more and more households are reducing their reliance on the grid") are effective at bringing less-engaged households on board and setting them more ambitious saving targets.
- Stated saving targets in surveys often fail to translate into real-world results, and programs should treat high self-reported ambitions with caution.
- Ambiguity regarding the extent to which voluntary contributions benefits society greatly weakens individuals' willingness to act. Communications strategies should therefore make the social and environmental impacts of individual actions as concrete as possible.
- When ambiguity cannot be removed, explicit normative messages (like "most people in your community consider saving fair and important") are powerful for resolving confusion. Practitioners should match the type of normative message to the source of uncertainty and be aware that highly self-interested individuals may remain unresponsive.

2.4. WP3: Revealed preferences and field experiment

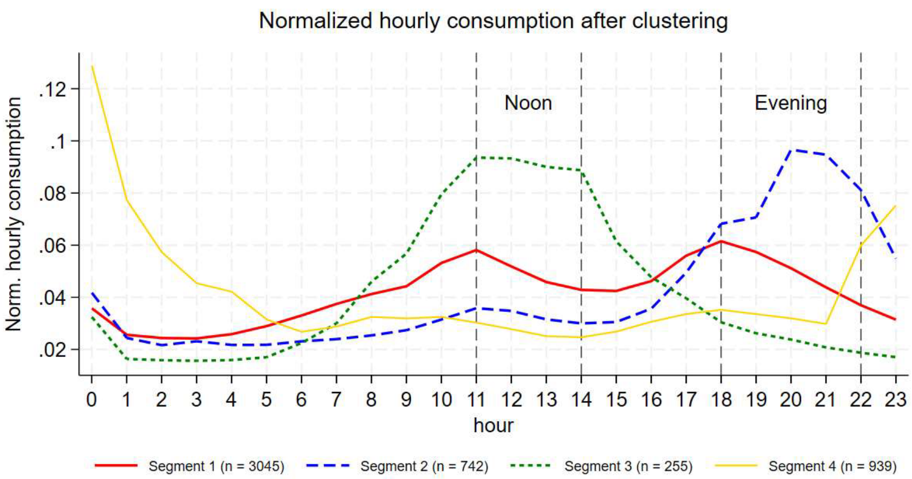
WP3 implements a field experiment to assess the potential and relative effect of individual feedback and social nudges (peer effects) on households' electricity conservation behavior. We target simple visualized feedback with little cognitive burden. The outcome is the achieved reduction (both immediate and long-term) in electricity consumption, measured via smart meters throughout the study period. We also measure the respondents' engagement by tracking user actions and the time spent on the smartphone app developed in WP1.

The objective is to test whether providing social comparison feedback via a mobile app, leads to real electricity savings in a field experiment. It also aims to uncover why certain interventions do or do not work by examining how personal circumstances alter people's responses. The goal is to replace assumptions with evidence about what changes household energy use in practice. The field experiment design and data analysis with all methodological details and alternative specifications are reported in **Favre-Bulle and Weber (2026b)**. Details of the computation of realized savings, based on a cluster analysis of load profiles, are also provided in Favre-Bulle and Weber (2026b). Further details on a similar but more sophisticated cluster analysis are available in **Favre-Bulle and Weber (2026a)**. Here, after a brief presentation of the cluster analysis, we provide a summary of the experiment design and main findings, as well as the results of a complementary analysis of heterogeneity and moderating factors.

2.4.1. Load profiles

Seasonal variations and other external factors affect electricity consumption. Hence, electricity savings cannot be computed as simple variations from one period to the next in the field experiment; rather, savings must be estimated based on a statistical methodology that corrects for such "natural" time variability. Each participating household is therefore compared to a reference group of similar non-participating households. We use hourly electricity consumption data from 4,981 households over a one-year period immediately preceding the experiment (August 2023 to July 2024) to cluster households into comparable segments. The model includes several variables based on households' load curves to measure seasonal and intra-day variations. We use the k-means clustering method and identify four distinct household segments.

The Figure below shows the average consumption for each segment identified in the cluster analysis. The consumption is normalized (i.e., measured in proportion of consumption in each hour) to make the load profiles directly comparable across segments and households regardless of their absolute consumption (measure in kWh). We observe substantial differences in patterns across segments. The largest segment (more than 60% of households) shows a typical camel-shaped load curve with a consumption spike in the middle of the day and another one in the evening. Segment 2 also has two spikes, but the magnitude of the evening one is considerably larger, whereas segment 3 displays an important and single spike in the middle of the day. Segment 4 has a radically different consumption profile, with a very important night consumption.



The realized saving in each household for a given week is then approximated by comparing their consumption reduction with a counterfactual baseline reduction in a comparable reference group. This is done by double-differencing with respect to a reference week (last week before the experiment) and a reference group (all non-participating households in the same cluster segment), using the following four-step algorithm:

1. We compute the household's consumption change compared to their consumption in the reference week (denoted Diff1).
2. For each segment, we compute the median (or 3rd quartile) consumption change using all households from the reference group (denoted Diff0). This median (or 3rd quartile) change is used as the counterfactual, i.e., the change that would have been observed in absence of

treatment, for each participating HH in the same segment. We used median for the 1st wave, but given the small estimated savings achieved by participants in that wave, we decided to use the third quartile in the second wave.

3. The HH's realized saving is then estimated as the difference Diff0-Diff1, i.e., the counterfactual change minus the actual change of consumption in the HH.
4. We compute cumulative savings over the weeks since the beginning of the experiment in relative terms (%) and in absolute values (kWh).

It is important to note that this algorithm provides an approximate indicator of savings for utilization in the field experiment. Because the counterfactual is an aggregate measure (median or 3rd quartile) of reductions in a reference group, it is subject to potential errors due to within-segment variations. However, this approximation gives a more realistic picture of achieved savings than a simple differencing of one household's own consumption, which is subject to external and seasonal factors.

2.4.2. Field Experiment: study design and treatment effects

We study the effects of social and financial incentives, as well as their interaction, on household electricity savings through a field experiment in which participants used the INFINEED app with weekly updates during one or two six-week waves. In total, 201 distinct households took part in at least one of the two waves, with 198 participating in the first wave and 92 in the second wave.

Participants were randomly assigned to one of the following three groups. The control group received information on their daily and weekly consumption, estimated savings from the previous week, estimated cumulated savings since the start of the experiment, and personalized energy-saving tips via an interactive chat. Two treatment groups received the same information, augmented with the following elements.

The second group (Social treatment) received comparisons of their estimated saving rates and cumulative savings against those of other participants from the same segment. These comparisons were presented through both rankings and graphs. The third group (Social + Financial treatment) received the same information as the Social group, but were moreover entitled to a monetary reward based on estimated savings: 5, 10, 20, 30, or 50 CHF for savings of respectively 0–2%, 2–4%, 4–6%, 6–8%, or above 8% at the end of each wave. This design allowed us to assess the effect of social incentives and the combined effect of social and financial incentives.

The following Table presents descriptive results from the field experiment at the end of each wave. First, it describes the distribution of participants across the three groups. Second, it reports the share of participants in different categories of estimated savings. The intervals used to present the distribution of savings are calibrated to correspond to the thresholds used to define financial rewards received by participants in the Social + Financial treatment group. The average reward obtained by participants in this group is also displayed at the bottom of the Table.

	Wave 1				Wave 2			
	Control	Social	S+F	Total	Control	Social	S+F	Total
Number of participants	71	62	65	198	35	27	30	92
	Estimated cumulative savings / Financial reward*							
	Control	Social	S+F	Total	Control	Social	S+F	Total
$< 0\,\%$ (No savings) / 0 CHF	67.6%	59.7%	58.3%	62.1%	25.6%	25.9%	10%	20.7%
$0 - 2\,\%$ / 5 CHF	1.4%	6.5%	4.6%	4.0%	14.3%	0%	6.7%	7.6%
$2 - 4\,\%$ / 10 CHF	5.6%	9.7%	4.6%	6.6%	2.9%	3.7%	6.7%	4.3%
$4 - 6\,\%$ / 20 CHF	8.5%	3.2%	4.6%	5.6%	5.7%	7.4%	10.0%	7.6%
$6 - 8\,\%$ / 30 CHF	1.4%	1.6%	3.1%	2.0%	2.9%	7.4%	13.3%	7.6%
$> 8\,\%$ / 50 CHF	15.5%	19.3%	24.8%	19.7%	48.6%	55.6%	53.3%	52.2%
Median savings (%)	-4.9	-4.05	-5.5	-4.75	8	10.1	9.2	8.85
Mean financial reward (CHF)	–	–	14.85	–	–	–	34.83	–

** In wave 1, savings were estimated using the median of the reference group as the counterfactual, while in wave 2 the third quartile was used, which partly explains the higher savings observed in wave 2. Note that, by definition, the proportions of the reference group in the "no saving" category would be 50% in wave 1 and 25% in wave 2. S + F = Social + Financial treatment.*

While being based on sheer approximations, two interesting patterns emerge from these distributions. First, the share of households in the "no savings" category is in fact higher among the experiment participants than in the reference group in wave 1 (62% vs 50% considering that the median of the reference group was used), but lower in wave 2 (21% vs 25% considering that the 3rd quartile of the reference group was used). Therefore, descriptive evidence does not suggest that participation is associated with greater savings. Second, a substantial share of participants have achieved savings above 8% with apparent differences across treatments. We use an elaborate econometric methodology to estimate these potential treatment effects.

The difference-in-differences (DiD) approach is particularly adapted to identifying the treatment effects. We implement it using a series of regressions explaining individual electricity consumption to determine how participants in treatment groups reacted compared to those in the control group. In other words, the DiD method uses pre-treatment electricity consumption to account for initial differences between groups and compares changes in consumption over time and across groups after the intervention, thereby delivering estimates of the causal impact of the treatments. The analysis is conducted at the weekly and daily levels, and, for a subsample of 146 households, at the hourly level in order to investigate potential treatment effects at a more granular temporal scale.

The main results seem to suggest that the Social treatment led to an *increase* in electricity consumption (contrary to what was expected), with a positive but statistically insignificant effect of about 6%. The Social + Financial treatment shows coefficients close to zero or slightly negative, yet without being statistically significant. Overall, our treatments therefore failed to trigger significant reductions in electricity consumption among the experiment's participants. We also find no evidence of treatment effects at specific hours of the day, across days of the week, or in their evolution over the course of the experiment.

Even though both treatment effects are not statistically significant, the results suggest different patterns across treatments. The Social treatment tends to be associated with higher electricity consumption, whereas the Social + Financial treatment shows no effect or a slight decrease relative to the control group. As the difference between the two treatment coefficients is sometimes economically meaningful (up to 10 percentage points), we formally test this difference using F-tests. In some specifications, particularly for median effects, this difference is statistically significant at the 10% level. This pattern is consistent with the hypothesis that adding a financial incentive to social incentives may mitigate the increase in electricity consumption observed under social incentives alone. Thus, while remaining limited and inconclusive, the results do not show evidence of crowding-out effects.

Households equipped with PV panels obviously have specific (net) electricity load profiles and they face different incentives for reducing their electricity consumption. We therefore hypothesize that households with or without PV could react differently to the treatments and conduct a specific analysis to explore this dimension. The treatment effects remain statistically insignificant. However, for households with PV, both treatments are associated with large positive coefficients, suggesting that none of the treatments is effective in reducing consumption in this group. In contrast, the results for households without PV are similar to those for the full sample, with a positive effect for the Social treatment and a slightly negative effect for the Social + Financial treatment. The differences between the two treatment coefficients for households with or without PV are sometimes significant although not always robust. Households without PV may therefore react more favorably to the Social + Financial treatment. While lacking statistical significance, the estimated treatment effects point to opposite effects among households with or without PV panels, suggesting a possibly negative moderating effect of PV ownership. Social feedback and financial incentives might be counterproductive for households with PV. Confirming these patterns would require further investigation with greater statistical power, hence a larger sample with a sufficient sample size for both groups.

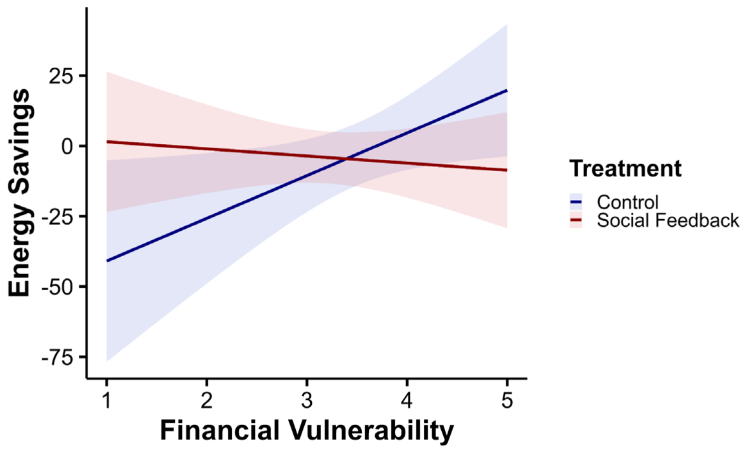
Given the small sample of participants (participation rate of about 3.5%), we examined the possibility that the participants might be a non-random selection of a population that could be more interested in electricity conservation. Our descriptive analysis indicates that compared to non-participating households, the field experiment participants are significantly more likely to live in a house (as opposed to an apartment) and to have a PV or heat pump. The difference is especially important regarding PV ownership. The share of PV-owners among participant HHs (one third) is about four times higher than among non-participating households. These patterns suggest a higher interest in energy-related investments and possibly a higher awareness of energy consumption among the participants. However, the participants' electricity consumption is not statistically different from that of non-participating households (see Favre-Bulle and Weber, 2026a), which suggests that while being a specific selection of households, the participants have a similar saving potential to comparable non-participant HHs. Note also that determining in which direction the treatment effects are biased in presence of self-selection is not obvious. On the one hand, people more interested in the electricity domain could devote more effort and therefore react more, so that our treatment effects would be overestimated. But on the other hand, these interested people are more likely to have already implemented actions to decrease their consumption, leaving fewer opportunities to further decrease their consumption, in which case our estimates would be under-estimations of the true treatment effects.

Finally, we re-examined the possible effect of participation in the field experiment, which could occur independently of the treatments. The single fact of having access to the app might indeed already affect the households' behavior. In order to test this hypothesis, we use propensity score matching (PSM) and create a sample where participating and non-participating households are comparable in terms of the following variables: hourly consumption (similar to the cluster analysis described above), indicators for flat (vs house), presence of PV (yes/no) and heating system. The matched sample is composed of 148 participating households and 270 non-participating households. The results of this analysis suggest that participation in the experiment is associated with a 1.4% reduction in electricity consumption at the sample median. However, the mean difference is not statistically significant, pointing to high variation and possible presence of outliers. Overall, there is no evidence that participation in the field experiment has reduced consumption in any significant manner for all participants.

2.4.3. Field Experiment: moderating effects and heterogeneity

Using the field experiment outcomes in combination with data collected through the online survey allows investigating whether personal characteristics moderate the treatment effects. We apply a series of complementary regressions to explore multiple hypotheses on such moderating effects. Two interesting patterns emerge from these analyses: financial vulnerability has a limiting effect on social comparison and perceived ambiguity has a moderating effect on user's engagement.

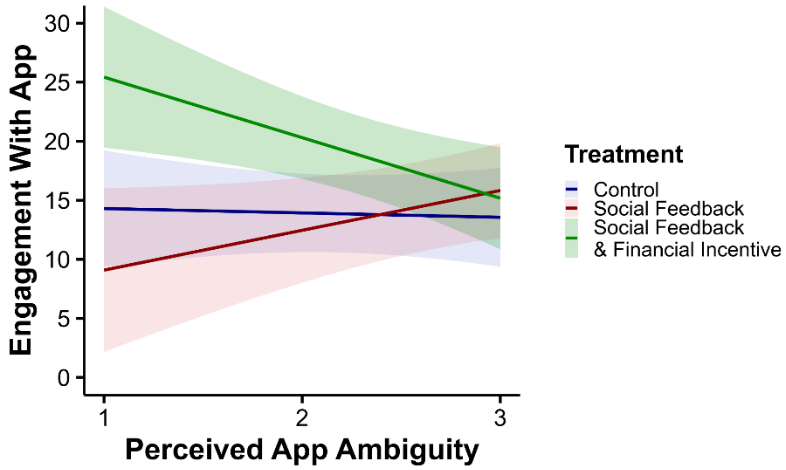
Prior research provides indirect support for the idea that consumer vulnerability – defined as a lack of access to the marketplace because of particular constraints, such as financial, physical or psychological issues – may reduce the impact of social comparisons. Our results align with this view: While less financially vulnerable households reduce energy consumption when exposed to social comparisons, more financially vulnerable households show weaker or even adverse responses, suggesting that comparative feedback is less motivating when financial constraints are stronger. In contrast, financially secure households may be better positioned to interpret social comparisons as attainable benchmarks. This moderating effect of financial vulnerability on the effectiveness of social comparison feedback is illustrated in the Figure below.



From a policy perspective, these findings suggest that uniform feedback strategies risk overlooking important heterogeneity in consumer responses. For financially vulnerable consumers (who could be identified, for instance, through a short onboarding question about income or difficulty paying bills), non-comparative and individualized feedback, focused on personal consumption levels or changes over time, may represent a more appropriate and effective approach (e.g., "You used 7% less than your typical week"). Tailoring feedback to socio-economic conditions could therefore improve both the effectiveness and inclusiveness of energy conservation interventions. These results should be interpreted with caution and warrant further investigation into the mechanisms at play. Nevertheless, they highlight the importance of designing behavioral interventions that take into consideration consumer vulnerability rather than assuming social comparison-based messaging may work for everyone.

Furthermore, our results indicate that perceived app ambiguity – defined as the extent to which users found the app's features and functionality open to interpretation or confusing – shapes how different feedback strategies affect engagement. When perceived ambiguity is low, users in the Social + Financial treatment group show the highest levels of engagement. However, as perceived app ambiguity increases, engagement in this group declines sharply, eventually converging toward the other conditions. In contrast, the Social treatment induces more engagement as perceived app ambiguity increases, while engagement in the control group remains largely stable across ambiguity levels.

The Figure below illustrates this moderating effect of perceived ambiguity on the relationship between feedback type and user engagement.



This pattern suggests that when users are uncertain about how the app works or how to interpret its features, simple social comparison cues may be easier to process and act upon than feedback that combines social comparisons with financial incentives, which may introduce additional complexity.

From a policy perspective, these findings suggest that financially framed incentives are most effective in low-ambiguity digital environments, whereas social comparison feedback appears more robust when users face uncertainty or confusion. Reducing perceived ambiguity or tailoring feedback strategies to users' understanding of the system may therefore be critical for sustaining engagement in digital energy interventions.

2.4.4. WP3's key takeaways

The field experiment's results point to no meaningful impact on energy savings from either social comparisons alone or from a combination of social comparisons and financial incentives. Although insignificant, the results point to some evidence suggesting a counterproductive effect for PV owners. Moreover, financially vulnerable households reacted negatively to being compared with others, suggesting they should receive only personal, non-comparative feedback instead of rankings. When the app was perceived as confusing, simple social comparisons were easier to grasp than combination of social messages and financial incentives. Therefore, in unclear digital environments, the message should be kept as straightforward as possible.

From a policy perspective, the results obtained in this WP can be considered as providing arguments against upscaling promotion campaigns based on individualized digital feedback. However, lack of evidence does not necessarily mean evidence of no effect. There is some evidence that the field experiment participants are relatively motivated and their lack of response might partly stem from having already exhausted some of the saving possibilities. Moreover, the statistically insignificant treatment effects could be at least partially explained by strong heterogeneity in individual responses and saving potentials, which is observed in our experiments and surveys, and in smart meter readings, and highlighted by moderating effects such as PV ownership.

3. Conclusions and outlook

Using a field experiment based on a digital interface and a series of online experiments, we study the effectiveness of two types of interventions and their interplay on sustainable behavior. The objective is to reduce electricity consumption through financial incentives and individualized social-comparison feedback. We also analyzed the impact of social norms and various moderating factors on different actions including conservation, efficiency investment and purchase of green electricity. The main conclusions are summarized below.

The analysis of potential and realized savings based on smart records and survey data suggest that:

- **Estimation of both potential and realized savings at the individual level poses an important challenge**. Reasonably accurate estimates require a large sample of participants and various personal characteristics, which is usually complicated because of legal barriers for accessing confidential consumption data combined with further information.
- Lacking sufficiently large and detailed samples, reasonable estimates of savings can be obtained by comparing the evolution of consumption of specific households with that of a comparable reference group, which plays the role of a counterfactual. The comparable reference groups were identified through a cluster analysis of smart meter consumption profiles and a few household characteristics.
- Missing values in smart meter data do not cause major issues for estimating savings but are not negligible for hourly estimates especially for shifting consumption away from peak hours.

Revealed preferences

Our field experiment was conducted with 201 volunteer households (out of about 6,000 private customers from the utility La Goule). Our extensive analysis of this experiment indicates that:

- Despite incentive payments and generous lottery rewards, the **participation rate in the field experiment is limited to about 3%**. The low participation rate can substantially decrease the statistical power of the analyses although it is relatively standard for similar experiments. This detrimental effect is exacerbated when responses are heterogeneous among different segments of the population.
- **The participant households do not show a detectable realized saving that can be linked in a statistically significant manner to any of our interventions** (social comparison feedback or financial incentives). However, there is a significant difference between households with and without solar panels. Social feedback combined with financial incentives appears to have opposite effects for these two groups, pointing to a possible counterproductive effect among owners of solar panels. On the other hand, households without photovoltaic could react as desired.
- **No evidence of significant effect should not be interpreted as evidence of zero effect**. In fact, social-comparison feedback appears to improve saving outcomes for some households although it has a counterproductive effect on others. These differences cannot be easily linked to observed characteristics.
- **Comparative feedback may not resonate with consumers facing tight financial constraints.** This explains why many participants among relatively low-income groups might disengage or dismiss comparisons with others.
- **User experience mediates saving outcomes.** Social feedback improves user experience, making the system more engaging, motivating, and emotionally meaningful. Social feedback

shows a significant and positive effect on participants' usage of the app and their engagement in electricity saving actions.

- **Perceived ambiguity emerges as an important barrier in user engagement**. Social comparison feedback appears to be helpful in mitigating these perceptions, hence effective in increasing engagement.
- **The user's expectation varies depending on energy knowledge and experience.** While novice participants appreciate information about concrete saving strategies, advanced users might show frustration about saving targets and estimations.
- **There is no one-size-fits-all solution** for an effective social comparison feedback incentivizing households' sustainable behavior. Overlooking heterogeneous responses could lead to counter-productive effects.

Stated preferences

The stated-preference data collected from four experimental studies of relatively large and representative samples, varying from about 300 to 1,200 Swiss respondents, provide suggestive evidence of the following behavioral responses:

- **Information about financial savings could have a crowding-out effect.** In particular, adding this information to social comparisons adversely affects the relatively low-consumption respondents as well as those who perceive conserving energy as important.
- **The effect of messages on dynamic norms** (societal transition toward greater commitment to electricity savings) **depends on their framing**. While climate-change framing appears to be more effective in changing the respondent's preferences toward more sustainable behavior, the security-of-supply framing is more effective in mobilizing less-engaged households, thereby broadening participation.
- **Ambiguity and uncertainty in sustainable choices and their consequences are used as a justification for inaction**, hence producing a detrimental effect on individuals' engagement in sustainable behavior. The resulting effect varies depending on the source of uncertainty.
- **Normative intervention messages could mitigate - to some extent - the ambiguity effects**. Their effectiveness appears to be relatively low when ambiguity concerns the benefits of pro-social action.
- There is **strong heterogeneity** across individuals with respect to their responses to normative messages, especially regarding their choice between three sustainable actions regarding electricity namely, conservation, efficiency investment and purchase of green electricity.

Outlook

Considering the project's results, and given the acquired experience in the implementation of the field experiment and the development of digital interface, we contend that future studies should pay attention to the following points:

Quantifying intervention effects and the potential crowding-out on realized savings require field experiments on relatively large samples of users across multiple regions.

- Given the low participation rate in similar field experiments, there is an important need for further marketing and communication research for a better understanding of participation barriers and drivers.

- Legal difficulties for accessing confidential consumption data, in particular when combined with personal characteristics, are an important barrier for future research especially for achieving reasonably accurate estimates of realized and potential savings at an individual level.
- Choice experiments play an important role in understanding the interplay of multiple interventions. These experiments can also be used as cost-effective tools to improve the design of field experiments.

4. National and international cooperation

We have developed two collaboration channels. The first one is a collaboration with a research team from University of Bern, working on another EES project SMART. The exchange between the two research teams allowed us to share experience regarding targeted interventions. The second channel is a cooperation with SWEET CoSi in order to use the SHEDS survey. We have integrated the DCE 2023 in SHEDS. We also used the SHEDS data to elaborate the models for saving potentials.

5. Publications and other communications

The project's results have been used for four IRENE working papers and one master dissertation that are available online (listed below). Moreover, appendices to this report (listed in section 7 below) will be used for five separate publications that will be submitted to international journals by January 2027. The working papers and the material in appendices will be used as separate PhD chapters of M. Berney, F. Giauque and V. Favre-Bulle. Finally, the project's results have been presented at four conferences listed below.

5.1. Publications available online:

Farsi, M. and S. Weber (2024). Swiss Household Energy Demand Survey: Past experiences and new perspectives (section 4.4). IRENE working paper 24-06, Institute of Economic Research.

Favre-Bulle, V. and S. Weber (2026a). Residential Electricity Consumption Patterns in Northwestern Switzerland: Evidence from the INFINEED Project. IRENE Working paper 26-03, Institute of Economic Research.

Favre-Bulle, V. and S. Weber (2026b). Financial and Non-Financial Incentives, and the Crowding-Out Effect: Evidence from a Field Experiment on Residential Electricity Consumption in Switzerland. IRENE Working paper 26-04, Institute of Economic Research.

Giauque, F. and M. Farsi (2026). Securing the grid or preserving the planet? The impact of dynamic norms on electricity sufficiency., IRENE Working Papers 26-01, Institute of Economic Research.

Giauque, F., M. Farsi, S. Weber and M. Puntiroli (2025). Killing the bill: The interplay of social comparisons and financial information on preferences for electricity-saving behaviors, IRENE Working Papers 25-02, Institute of Economic Research.

John, A. (2023). Addressing Missing Smart Meter Data in Electricity Consumption Using Machine Learning. *Master Thesis*, Faculty of Economics and Business, University of Neuchâtel. https://libra.unine.ch/handle/20.500.14713/99829

5.2. Conference participation:

The results reported in appendix WP2 (Berney et al.) were presented at the CHI Conference on Human Factors in Computing Systems (2024).

The results reported in the working paper by Giauque and Farsi (2026) were presented at the EAERE Annual Conference 2025 in Bergen, anatin the SABE 2025 Annual Conference in Trento.

The results reported in the working paper by Giauque et al. (2025) were presented in SSES Annual Congress 2024 in Lucerne.

The findings reported in the working paper by Favre-Bulle and Weber (2026) were presented in ENERDAY 2024 - 18th International Conference on Energy Economics and Technology.

6. References

Andor, M. A., & Fels, K. M. (2018). Behavioral economics and energy conservation-a systematic review of non- price interventions and their causal effects. *Ecological Economics*, 148, 178–210. https://doi.org/10.1016/j.ecolecon.2018.01.018

Brandon, a., List, J. A., Metcalfe, R.,Price, M. K., & Rundhammerd, F. (2018). Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. PNAS 116 (12) 5293-5298, https://doi.org/10.1073/pnas.1802874115

Caballero, N., & Della Valle, N. (2021). Tackling energy poverty through behavioral change: A pilot study on social comparison interventions in social housing districts. *Frontiers in Sustainable Citie*s, 2, 1-20. https://doi.org/10.3389/frsc.2020.601095

Cellina, F., Fraternali, P., Gonzalez, S. L. H., Novak, J., Gui, M., & Rizzoli, A. E. (2024). Significant but transient: The impact of an energy saving app targeting Swiss households. *Applied Energy*, *355*, 122280. https://doi.org/10.1016/j.apenergy.2023.122280

Claudy, M.C., Peterson, M. & O'Driscoll, A. (2013). Understanding the attitude-behavior gap for renewable energy systems using behavioral reasoning theory. *J. of Macromarketing*, 33(4), 273-287.

Farsi, L. Ott & Weber S. (2020). Les intentions contradictoires des Suisses vis-à-vis de leur consommation d'énergie, *Social Change in Switzerland N° 21.* https://doi.org/10.22019/SC-2020-00001

Farsi, M. and S. Weber (2024). Swiss Household Energy Demand Survey: Past experiences and new perspectives. IRENE working paper 24-06, Institute of Economic Research.

Favre-Bulle, V and S. Weber (2026a). Residential Electricity Consumption Patterns in Northwestern Switzerland: Evidence from the INFINEED Project. IRENE Working paper 26-03, Institute of Economic Research.

Favre-Bulle, V and S. Weber (2026b). Financial and Non-Financial Incentives, and the Crowding-Out Effect: Evidence from a Field Experiment on Residential Electricity Consumption in Switzerland. IRENE Working paper 26-04, Institute of Economic Research.

Giauque, F. and M. Farsi (2026). Securing the grid or preserving the planet? The impact of dynamic norms on electricity sufficiency., IRENE Working Papers 26-01, Institute of Economic Research.

Giauque, F., M. Farsi, S. Weber and M. Puntiroli (2025). Killing the bill: The interplay of social comparisons and financial information on preferences for electricity-saving behaviors, [IRENE Working Papers 25-02](), IRENE Institute of Economic Research.

Hamilton, R., Thompson, D., Bone, S., Chaplin, L. N., Griskevicius, V., Goldsmith, K., Hill, R., John, D. R., Mittal, C., O'Guinn, T., Piff, P., Roux, C., Shah, A., & Zhu, M. (2019). The effects of scarcity on consumer decision journeys. *Journal of the Academy of Marketing Science*, *47*(3), 532–550. https://doi.org/10.1007/s11747-018-0604-7

Holzer, A., Kocher, B., Bendahan, S., Vonèche Cardia, I., Mazuze, J., & Gillet, D. (2020). Gamifying knowledge sharing in humanitarian organisations: A design science journey. *European Journal of Information Systems*, *29*(2), 153-171. https://doi.org/10.1080/0960085X.2020.1718009

John, A. (2023). Addressing Missing Smart Meter Data in Electricity Consumption Using Machine Learning. *Master Thesis*, Faculty of Economics and Business, University of Neuchâtel. https://libra.unine.ch/handle/20.500.14713/99829

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2021). 1D convolutional neural networks and applications: A survey. *Mechanical systems and signal processing*, *151*, 107398.

Lanz, B., Wurlod, J., Panzone, L., and Swanson T. (2018) "The behavioral effect of Pigovian regulation: Evidence from a field experiment", *Journal of Environmental Economics and Management*, 87, 190-205. https://doi.org/10.1016/j.jeem.2017.06.005

Mari, A., Remlinger, C., Castello, R., Obozinski, G., Quarteroni, S., Heymann, F., & Galus, M. (2025). Real-time estimates of Swiss electricity savings using streamed smart meter data. *Applied Energy*, *377*, 124537. https://doi.org/10.1016/j.apenergy.2024.124537

Park, H. J., & Lin, L. M. (2020). Exploring attitude–behavior gap in sustainable consumption: Comparison of recycled and upcycled fashion products. Journal of Business Research, 117, 623-628.

Puntiroli, M., Moussaoui, L. S., & Bezençon, V. (2022). Are consumers consistent in their sustainable behaviours? A longitudinal study on consistency and spillover. *Journal of Business Research*, 144, 322–335. https://doi.org/10.1016/j.jbusres.2022.01.075

Schubert, I., Weber, S. Martinez-Cruz, A.L, Burger, P., Farsi, M. (2022). Structural equation modelling as a route to inform sustainable policies: the case of private transportation. *Frontiers in Sustainability*,3, 1-16. https://doi.org/10.3389/frsus.2022.837427

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). Research Article: The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*, *18*(5), 429. https://doi.org/10.1111/j.1467-9280.2007.01917.x

Schwartz, D., Milfont, T. L., & Hilton, D. (2019). The interplay between intrinsic motivation, financial incentives and nudges in sustainable consumption. In A research agenda for economic psychology. Edward Elgar Publishing.

Wagner, T. P. (2017). Reducing single-use plastic shopping bags in the USA. *Waste Management*, 70, 3-12.

Weber, S., Puddu, S., & Pacheco Barzallo, D. (2017). Move it! How an electric contest motivates households to shift their load profile. *Energy Economics*, 68, 255-270. https://doi.org/10.1016/j.eneco.2017.10.010

Wolske, K.S., Gillingham, K.T. & Schultz, P.W (2020). Peer influence on household energy behaviours. *Nature Energy*, 5, 202-212. https://doi.org/10.1038/s41560-019-0541-9

Appendix

This report has appendices listed below.

Appendix WP0: Forecasting Electricity Saving Potential among Jura Mountains' Residents: Results from a One-dimensional Convolutional Neural Network (author: Ott, L.)

Appendix WP2: Prosocial Behavior under Ambiguity: The Role of Social Norms and Beliefs (authors: F. Giauque, M. Farsi and M. Puntiroli)

These texts will be used for two separate publications by members of the research team, that we expect to submit to international journals by January 2027. Appendix WP2 will be a chapter of Fabien Giauque's PhD dissertation to be completed (defense planned for October 2026). An additional working paper is planned to further investigate hypotheses emerging from the results of WP3. It is expected to form a chapter in Valentin Favre-Bulle's PhD thesis to be completed by the end of 2026 (defense expected in early 2027).

h e g
Haute école de gestion
Genève

unine
Université de Neuchâtel
Faculty of Economics
and Business

Appendix WP0 – Forecasting Electricity Saving Potential among Jura Mountains' Residents: Results from a One-dimensional Convolutional Neural Network*

Laurent Ott†

January 21, 2026

Abstract

This paper presents a one-dimensional convolutional neural network trained to predict quantiles of weekly electricity consumption at household level for a small sample of residents of the Jura mountains. Its goal is to provide consumers with a forecast of their electricity use over the coming week, with a saving potential target represented by the prediction of the 30^{th} quantile of their consumption. Past electricity use, as well as household-level fixed variables and contextual time series on weather, climate, and time of the year, are used as inputs. The hidden features learned by the convolutional layers provide useful information to forecast quantiles of future electricity use, although the lack of household-level data limits the capacity of the model to learn much from them. A larger dataset would be necessary to improve the model's accuracy and extend its capacities.

*This research has been funded by the Swiss Federal Office for Energy as part of the INFINEED project (no. SI/502530). Special thanks go to Mehdi Farsi, Sylvain Weber, Laurent Raeber, Vladimir Macko, and Valentin Favre-Bulle for their help and contributions.

†University of Applied Sciences and Arts – Western Switzerland, Haute école de gestion de Genève. Contact: laurent.ott@hesge.ch.

1 Introduction

Incentivizing households to lower their electricity consumption is a major task for policy makers. As electrification grows, there is an urgent need for a more careful and parsimonious use of electricity by end users. The INFINEED (*Interplay of Feedback and Incentive Effects on Electricity Demand*) project, funded by the Swiss Federal Office for Energy, aims at tackling this issue from a consumer behavior perspective by studying the impact of individual feedback and social comparisons on residential electricity consumption. This paper contributes to solving one of the issues faced when trying to provide individual feedback to consumers: What is their electricity saving potential, given their profile?

This paper proposes to use a one-dimensional convolutional neural network (CNN) trained on past electricity consumption data, dwelling characteristics, and known future contextual inputs to forecast quantiles of one-week ahead electricity demand from the grid. This approach should provide an interval of potential consumptions to each household around the median consumption that could be expected from it, given its characteristics, its recent electricity consumption history, and external contextual factors such as weather or time of the year. These forecasts could be then used to provide feedback to households on their electricity use as well as their saving potential to nudge them into lowering their electricity consumption.

Using deep learning methods to forecast electricity use has already be done by many research teams (see e.g. Klyuev et al., 2022; Nti et al., 2020; Chae et al., 2016). At household level, existing studies include for instance Yildiz et al. (2018) in New South Wales, Australia, Marvuglia and Messineo (2012) in Palermo, Italy, or Kiprijanovska et al. (2020) with the Pecan Street dataset, which mainly covers the city of Austin, USA. The novelty of the present study is the task of the model, which is not necessarily to provide the most accurate prediction, but to predict quantiles of electricity use that can be interpreted as a saving potential target for each household. The prediction could for example be used to provide feedback through an application, with the intend to nudge consumers to lower their electricity use.

The rest of the paper is structured as follows. Section 2 explains the modeling strategy with respect to existing literature and presents the architecture of the model. Section 3 describes the dataset, its structure, as well as the treatment of missing values. Section 4 presents the training of the model and comments the results that have been obtained. Finally, section 5 concludes by summarizing the main outcomes and their limits, and formulates some recommendations regarding the use of the model.

2 Modeling strategy

2.1 Convolutional neural networks and time series

CNNs (see e.g. Li et al., 2021) are a family of neural networks largely used in different fields of research and applications, especially to process two-dimensional (2D) data like images. Their principle is to extract useful features from structured data by applying a filter called a *kernel* to them. This kernel is moved across input data and subsequent hidden layers, and optimized to extract signals by computing the dot product of the kernel and the input—the *convolution* step. The output of the convolutions are then fed into the next components of the model, including an activation function that determines which signal is transmitted to the next layer or not.

The development and applications of CNNs to one-dimensional (1D) data have been reviewed by Kiranyaz et al. (2021). The authors highlight how the use of CNNs has expanded from image treatment to data such as text or time series, with use-cases like anomaly detections in electrocardiograms, or real-time speech-recognition. The main difference with 2D-CNNs is that 1D-CNNs are less complex, and thus easier to train, because they generally have a lower number of parameters, due to their 1D nature. Otherwise, the same logic applies to both types of model, as is described more extensively below.

The present paper focuses on a particular type of 1D data: time series, and a specific task: forecasting. Time series usually consist in repeated measurements over time of a single object at a regular frequency. Traditionally, time series have been analyzed with different flavors of autoregressive models (see Hamilton, 1994), including in forecasting exercises. Machine learning algorithm have offered new options to solve prediction tasks, with recurrent neural network (RNN) models like long short-term memory (LSTM, see Hochreiter and Schmidhuber, 1997) or gated recurrent unit (GRU, see Cho et al., 2014), or even transformer-based models (Vaswani et al., 2017) like the temporal fusion transformer (TFT, see Lim et al., 2021).

1D-CNNs are another way of treating time series in the world of deep learning. Their basic structure is the following, in the case of a multiple time series (MTS).

1. The input MTS is formatted as a $T_{in} \times 1 \times K_{in}$ object, where T_{in} indicates the number of time steps—or observations—it contains, and K_{in} the number of series—or channels—of the MTS.

2. The first hidden convolutional layer is created. To do so, a convolution kernel of dimensions $\tau \times 1 \times K_{in}$ is applied to the multiple-channel input series, with $\tau < T_{in}$, and moved across it by strides of length s. The dot product of the kernel and the sequence of the MTS to which it is applied is computed, summed across channels, and goes through an activation function $f(\cdot)$. The output is stored in a $L_1 \times 1$ vector. To ensure that all observations appear equally across convolutions, 0-padding of length p is added before the first observation and after the last, so that $L_1 = (T_{in} + 2p - \tau)/s + 1$. This operation is

performed for each of the H_1 channels of the first convolutional layer. Note that if $s = 1$ and $p = 0.5(\tau - 1)$, then $L_1 = T_{in}$, which is often the case in practice.

3. The H_1 channels go through a pooling layer, which is meant to reduce the spatial size of the hidden representations that have been learned in the convolutional layer. The pooling operation usually consists in taking the maximum or the mean value obtained during each convolution. This reduces the dimensions of the hidden layer from $L_1 \times 1 \times H_1$ to $M_1 \times 1 \times H_1$, with $M_1 < L_1$, which is especially useful when $L_1 = T_{in}$. Note that the pooling layer is not strictly required; another option would be to use $s > 1$ to reduce the dimensions of the hidden layer instead, so that $M_1 = L_1 < T_{in}$.

4. A new set of convolutions can be performed again on the H_1 channels with a kernel of dimensions $\tau \times 1 \times H_1$, resulting in a second hidden convolutional layer that contains H_2 channels of dimensions $L_2 \times 1$, where $L_2 = (M_1 + 2p - \tau)/s + 1$. A pooling layer can then be added again to reduce the dimensions of the hidden layer channels from $L_2 \times 1 \times H_2$ to $M_2 \times 1 \times H_2$, if needed. Further hidden convolutional and pooling layers can be added after each others to create a deeper network of Q hidden combined convolutional-pooling layers.

5. The hidden representations coming from the last convolutional and pooling layers are flattened as a $(M_Q \cdot H_Q) \times 1$ vector and fed into a set of hidden fully connected layers, shrinking dimensions to match those of the output the analyst wants to predict with the model, i.e. $T_{out} \times 1 \times K_{out}$.

The whole model is optimized with an algorithm by backpropagation to find the parameters that minimize the chosen loss function.

Many variations of 1D-CNNs for time series exist, with different flavors and combinations of convolutional layers—see e.g. Liu et al. (2019) and Oh et al. (2024). These models can show good performances: for instance, Liu et al. (2019) find that their CNN outperform many other models in the prognostics and health management 2015 time series classification challenge.

2.2 A model to predict electricity saving potential

2.2.1 Prediction task and loss function

In this paper, the prediction task to solve is to forecast an electricity saving potential that could be provided to consumers to inform them on how much they could reduce their electricity demand, as part of an incentive scheme. Practically, the model should be able to provide an interval of potential consumption levels while using only minimal personal data, so that it could easily be used by e.g. electricity providers to provide feedback to their new customers while not necessitating to large a dataset, in particular regarding the number of different households needed for training. The target is therefore a single time series of electricity consumption over a period of time.

Given the nature of the prediction task, a quantile loss function is used. For any quantile γ, the loss is:

$$Q(y_i, \hat{y}_i) = \begin{cases} \gamma(y_i - \hat{y}_i) & \text{if } \hat{y}_i < y_i \\ (1-\gamma)(\hat{y}_i - y_i) & \text{if } \hat{y}_i \geq y_i \end{cases} \tag{1}$$

where y_i is the target value and $\hat{y}_i$ the prediction for observation i. The goal of this loss is to give different weights to predictions depending on whether they are above or below a certain quantile. For instance, in the case where $\gamma = 0.2$ and $y_i = 1$, then $Q(y_i = 1, \hat{y}_i = 1.1) = 0.08$ and $Q(y_i = 1, \hat{y}_i = 0.9) = 0.02$. In this example, it means that a prediction exceeding the true value by 0.1 is penalized four times more than a prediction that underestimates it by a same amount. By doing so, the loss gives a different importance to over- and underestimations, so that predictions tend to converge to the desired quantile of the distribution of y_i. Note that when $\gamma = 0.5$, the model predicts the median of the distribution.

To obtain an interval of values around the median, three quantiles are considered here: $\gamma \in \{0.3, 0.5, 0.7\}$. One thus gets predictions for the 30$^{\text{th}}$, 50$^{\text{th}}$, and 70$^{\text{th}}$ quantiles of the distribution of each household's electricity consumption. The idea is that the 30$^{\text{th}}$ quantile indicates the electricity saving potential, in relation to the median. On the other hand, the 70$^{\text{th}}$ quantile shows how one could perform worse in terms of electricity consumption.

2.2.2 Model architecture

The 1D-CNN used to solve the aforementioned prediction task is relatively simple in its architecture. It consists in a set of hidden 1D convolutional layers—the feature extraction phase—followed by a set of hidden fully connected layers—the regression head. A description of the blocks entering into these different phases is provided in Table 1.

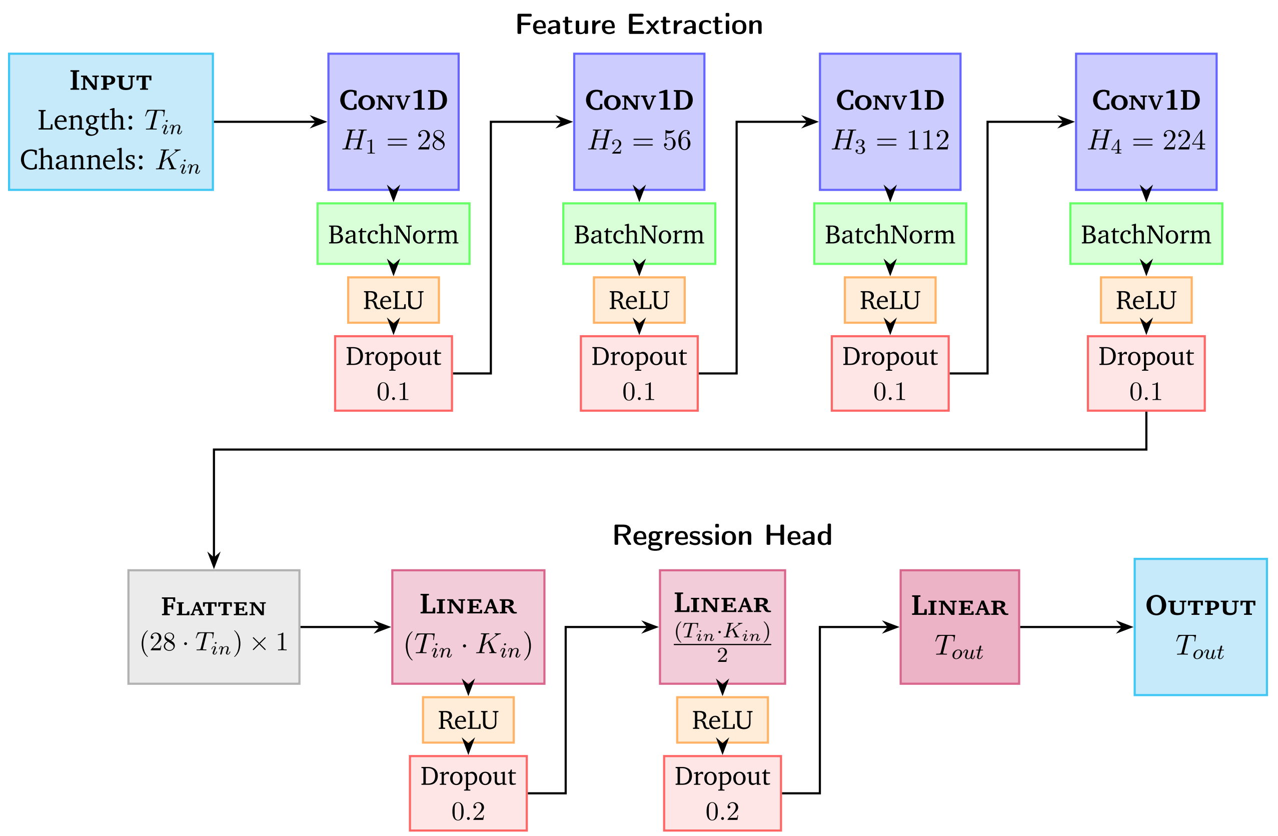
Table 1: Phases and blocks of the model architecture

Phase	Block	Description
Feature extraction	1D-convolution	Kernel size: $3 \times 1 \times$ input channels Padding: 1 Stride: 1 for the first layer, 2 for the others
	Batch normalization	–
	Activation function	ReLU: $f(x) = \max(0, x)$
	Dropout	10% random dropout
Flattening	1D-flattening	Transformation of the output of the last convolutional layer to a 1D vector
Regression head	Linear transformation	Affine linear transformation of the input
	Activation function	ReLU: $f(x) = \max(0, x)$
	Dropout	20% random dropout

A particularity of the model is that it does not feature any pooling layer, but relies on a stride of length 2 for the convolutional layers, starting from the second one, to reduce the dimensions of the hidden layers during the feature extraction phase. This choice has been made after a phase of exploration, where the performance of models with maximum or average pooling would be slightly worse than models with a stride larger than 1 in terms of accuracy. It also makes sense to do so with the data used here to solve the prediction task: The goal is not necessarily to only extract the most salient features from the strongest signals that would indicate the presence of specific events, but to make predictions from the whole input data for a long vector of output values.

The general structure of the model consists in four convolutional layers, one flattening layer, and three fully connected layers. The whole model architecture is displayed in Figure 1. It first learns hidden representations of the input MTS during the feature extraction phase, then flattens them into a single 1D vector before feeding it to three fully connected layers in the regression head in order to map learned representations to the output vector they aim to predict.

Figure 1: Model architecture



2.2.3 Input and target variables

Target variable consists in electricity consumption over one week at a specific frequency—hourly, 3-hour, and 6-hour frequencies have been tested, so the output vector length ranges from 28 to 168 time steps, depending on the model.

Input data contains three categories of variables. The first category is past electricity consumption over the last four weeks, recorded at hourly frequency. The second category comprises external contextual inputs, such as weather and climate, time of the year, or public holidays, during the period of observation of the target variable. Finally, the third category contains relevant household-level dwelling characteristics that could influence electricity use during the period of observation of the target variable.

Input variables entering the model are shaped as a MTS of dimensions $168 \times 1 \times K$, where K indicates the number of channels. Each channel consists in a separate series entering the MTS, be it a dwelling characteristic, a contextual input, or past electricity consumption. The choice has been made to format all channels as vectors containing one observation per hour over a week. It means that past electricity consumption takes the form of four one-week vectors.

The idea behind this choice of formatting for the input variables is that the 1D-CNN will slide over periods of the week that match between all input series and the target, so that the model should learn from past records at the same moment of the week as well as current context, with dwelling characteristics allowing for complex interactions of the different input series with each household's specific features regarding their home.

3 Data

3.1 Electricity consumption

The main variable of interest is household electricity consumption. It has been collected by smart meters (SMs) that have recorded electricity use from the grid at fixed intervals for a sample of households living the Jura mountains region who are customers of La Goule, a local electricity provider[1] that acts as partner of the INFINEED project. The SMs automatically send to the provider the cumulated amount of electricity consumed since their installation; by taking the difference between the previous and the current amounts, it is possible to compute the amount of electricity used in between.

Two generations of SMs record consumption at different frequencies: the older ones hourly, the newer ones every 15 minutes. To ensure homogeneity, 15-minute observations have been aggregated to obtain hourly data for all households.

The overall observation period ranges from Monday, January 2, 2023, to Sunday, June 22, 2025. However, data over the whole period are not available for all households, for which only shorter series are available.

As SMs only record electricity consumption from the grid, they cannot directly provide information on self-consumption for solar photovoltaic (PV) panels owners. During times when a

[1] See https://www.lagoule.ch/.

household is self-sufficient thanks to the electricity generated by its PV panels, and also potentially stored in batteries, its consumption form the grid is recorded as 0. Nevertheless, it can also happen that 0 consumption from the grid is recorded for other households without PV, although this case is unlikely, given that some appliances are generally always turned on, like fridges, freezers, or appliances staying in stand-by mode. These unusual cases make up only a minority of recorded 0s, though.

3.2 Context variables

Some variables providing information on external context are also included as inputs, given that electricity demand varies over time depending on daylight duration, season, or weather conditions. Seven context variables have been selected: total rainfall; sunshine duration; average outside air temperature 2m above the ground; daylight duration; daylight saving time (DST); school breaks; public holidays.

Weather and climate can impact electricity use through the need for artificial light and heating. Sunshine duration also directly influence the generation capacity of PV, as is the length of the day. DST can also be expected to influence electricity use because it shifts many activities by one hour—i.e., activities usually carried out at 7am UTC are shifted to 8am UTC because of DST. On the other hand, school breaks and public holidays increase the risk of lower consumption due to an absence from home, or different electricity use patterns if more people stay at home during the day than usual.

Data on weather and climate have been obtained from the Federal Office of Meteorology and Climatology (MeteoSwiss). MeteoSwiss runs automatic ground-based data collection stations that gather high frequency information on the current situation at their location. The closest station from the region where studied households reside is located in La Chaux-de-Fonds; all measurements hence come from this station, although local conditions may sometimes differ over time. It is thus assumed that local conditions are correlated with those recorded in La Chaux-de-Fonds, because no high frequency measurements would be available from MeteoSwiss otherwise.

It should be noted that *recorded* weather and climate data are used as input, whereas if an electricity provider would want to provide an electricity saving potential forecast to a customer for the forthcoming week, only weather and climate *forecasts* would be available. This choice has been made to train the model on data as close to reality as possible, so that electricity consumption predictions that could be made with the trained model using weather and climate forecasts would be the best possible at the current state of knowledge.

3.3 Household-level characteristics

Although the collection period for SM data is fairly long, the series are only available for a limited number of households. In total, the sample contains data on 162 households. This implies that only few household-level characteristics can be used, otherwise there would not be enough observations for all combination of variable values.

After some tests, only three binary variables concerning the dwelling have been included in the model as inputs: presence of solar PV panels; having an electric system for heating, including heat pumps; having an electric system only for hot water, but not for heating. As there are only very few records of households that have an electric system only for heating but not for hot water in the data, this case has not been considered separately from the others. Information on the type of dwelling—single family house, apartment, etc.—has not been used, because it is too strongly correlated with the other variables, especially the presence of PV. A larger dataset containing more households would be required to include more inputs of that type.

All household-level data come from a survey conducted between July and November, 2024. They thus represent the situation at that moment of time. It therefore cannot formally be excluded that some information that has been collected is not valid during other periods, although no household from the sample has moved to another home during the time period considered here. The accuracy of provided information also fully depends on recorded answers; the eventual presence of errors cannot be excluded.

3.4 Data imputation

A major issue of SM data is that missing observations caused by technical discrepancies are relatively common, introducing gaps in times series that can hinder model training and cause significant data loss. In particular, because of the proximity of the French border to studied region and the use of the mobile phone network to transmit data, some recordings have been lost due to network connection issues. Other types of data transmission failures could also have occurred. However, in all cases, as SMs transmit the cumulated amount of electricity used since they have been installed, the total amount of electricity consumed from the grid during a period of failure is available at the next successful data transmission event.

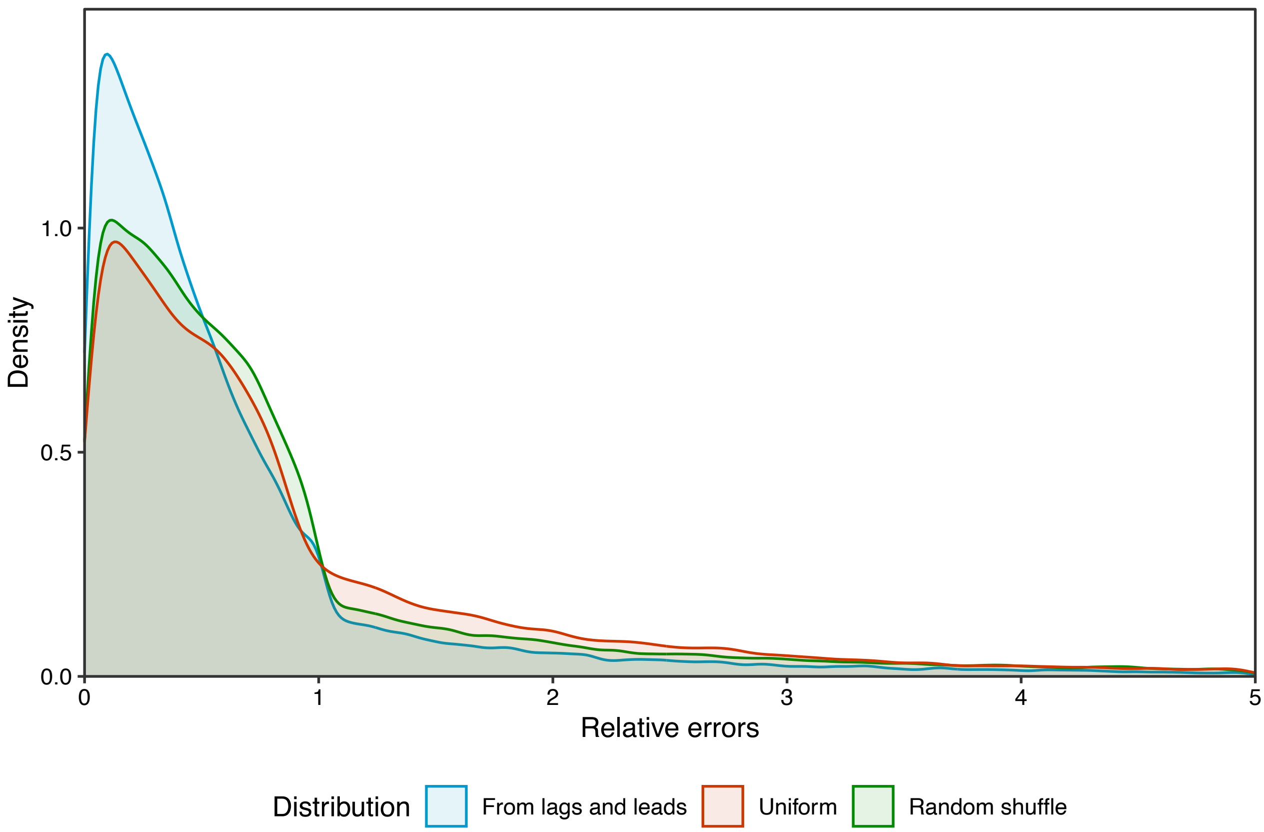
In order to have as many complete electricity consumption series as possible to train and test the model, a data imputation procedure has been set up to fill some of the gaps. Because the length of some series of missing observations can be long, it has been decided to only impute data for periods of up to 48 hours. This ensure that weekly series should not contain too many imputed data, which could be a problem for the model.

Data imputation has been performed in a two-step procedure. First, the average of measured values for the same hour of the same day from the previous and the following weeks is computed for a series of consecutive missing observations. Then, these average values are divided by their sum over the series of missing recordings, which gives the shares of total electricity consumption over the imputation period that should be allocated to each hour. Then, in a second step, the recorded cumulated consumption over the whole period of missing observations is allocated to missing hourly records according to the shares computed during the first step of the procedure.

The quality of the output of the procedure is assessed by using series of records without missing observations for which data imputation is performed, so that real and imputed values can be compared. A set of 3,508 24-hour test imputations have been performed by randomly sampling households and days. The output of the imputation method described above is compared to what would be obtained with a uniform allocation of recorded cumulated consumption over each 24-hour period, or if imputed values where randomly shuffled within each period. The distributions of relative imputation errors obtained with each method are compared for correct non-0 values, as are the number of correctly predicted 0 values.

Figure 2 presents kernel density plots of relative imputation errors for the three methods: distribution of values guessed from lags and leads, uniform distribution, and random shuffling. As can be seen, the imputation method from lags and leads performs better than using a uniform distribution and randomly shuffling imputed values. The distribution of missing values learned from lags and leads, although not perfect, seems to provide relatively better results than uniform or random guesses for non-0 values.

Figure 2: Densities of relative imputation errors



Regarding correct imputations of 0s, a comparison of the three methods shows that 0s are correctly imputed as such in 50.9% of cases if using the distribution learned from lags and leads, against 36.6% with a uniform distribution, and 40.2% with a random shuffling of values. Again, although the method based on lags and leads is not perfect, it performs better than the others.

All in all, it seems that this method can be used to impute missing values, even though it probably also adds noise to the data. Nevertheless, as imputed values only represent about 1% of total observations, the overall influence of this noise should remain low.

3.5 Dataset construction

The dataset contains all aforementioned variables and is structured around household-specific electricity consumption time series. Each row of this tabular dataset contains an observation for a household at a specific date and hour of the day. However, given the model architecture and the shape of input features and output series it requires, the dataset has been formatted as an array of household-specific 5-week data blocks—that is, for each household, each available 5-week observation window is formatted as a $T \times K$ matrix tensor, where T indicates the number of time steps and K the total number of input and output columns.

The goal of the model is to forecast one week of electricity consumption from four weeks of past electricity consumption data, time-invariant household-level characteristics, and time-varying contextual information that is available for the week to forecast. Therefore, the columns of each matrix tensor correspond to:

- Electricity consumption during week t: 1 column;
- Electricity consumption from $t-4$ to $t-1$: 4 columns;
- Time-varying inputs during week t: 7 columns;
- Time-invariant household-specific inputs: 3 columns.

With hourly data, each matrix tensor has dimensions 168×15, which contains the 168×14 matrix of input features and the 168×1 output vector. The reference week of each block is t, i.e. the output week.

This dataset structure makes it possible to create a lot of data blocks from limited observations, because the same week of electricity consumption can be used once as output, and up to four times as inputs. For instance, if nine weeks of data are available for a household, five blocks can be created: the first one with weeks 1 to 4 as inputs and week 5 as output, the second one with weeks 2 to 5 as inputs and week 6 as output, and so on.

In total, 15,902 data blocks are available in the dataset, as electricity consumption has not been recorded for all households during the whole time period. However, as the dataset needs to be split into training, validation, and test sets, not all these blocks can be used. The procedure applied to delimit the three sets is the following. First, 85% of households, i.e. 138 of the 162 present in the data, have been randomly set as belonging to the training and validation sets, whereas the remaining 24 have been allocated to the test set. Then, a splitting date has been determined to delimit the training set, which contains blocks located before this date, from the validation and the test sets, which only contains blocks located after this date. The three sets thus contain the following data:

- Training set: observations before July 7, 2024, for 138 households: 7,566 blocks;
- Validation set: observations from July 7, 2024, for 138 households: 5,980 blocks;
- Test set: observations from July 7, 2024, for 24 households: 1,032 blocks.

Households that belong to the training and validation sets are the same, but the periods considered differ, whereas the test set is completely exogenous from the training set. With this procedure, only observations for the 24 households in the test set located before July 7, 2024, are discarded. This ensures that not too many observations are lost, given the limited number of households for whom data are available.

3.6 Descriptive statistics

Table 2 presents some descriptive statistics of the variables by set. Context variables that only depend on the period of the year—daylight duration; DST; school breaks; public holidays—have not been included because summary statistics on such time series are not very informative. For household-level dummies—presence of solar PV panels; having an electric heating system or a heat pump; having an electric system only for hot water—only the mean is presented, as all other statistics from the table are directly derived from it.

Statistics have been computed at each variable's natural level of variation: over time for time-variant contextual inputs; across households for household-level inputs; over time within each household, then across households for electricity consumption. For the latter, following the Law of total variance, standard deviation has been computed as the square root of the average of household-level variances plus the average of squared differences between household-level means and the average household-level mean: $\sigma = \sqrt{\frac{1}{N}\sum_i \sigma_i^2 + \frac{1}{N}\sum_i (\mu_i - \mu)^2}$.

3.7 Value scaling

As not all variables share the same range of values, it is necessary to scale them to limit the risk that too strong signals from one or several variables negatively impact the learning procedure because of their influence. Different methods are applied depending on each variable's range

Table 2: Descriptive statistics, by set

Variable	**Unit**	**Set**	**Mean**	**S.D.**	**Share of 0s**	**Q30**	**Q50**	**Q70**
Electricity consumption	kWh	Training	0.48	0.70	4.23%	0.16	0.29	0.49
		Validation	0.45	0.67	3.77%	0.15	0.26	0.45
		Test	0.45	0.83	4.35%	0.16	0.23	0.39
Hourly rainfall	mm/h	Training	0.17	0.63	80.97%	0.00	0.00	0.00
		Validation	0.14	0.64	85.99%	0.00	0.00	0.00
		Test	0.14	0.64	85.99%	0.00	0.00	0.00
Hourly sunshine	h/h	Training	0.19	0.35	69.07%	0.00	0.00	0.02
		Validation	0.23	0.38	66.11%	0.00	0.00	0.12
		Test	0.23	0.38	66.11%	0.00	0.00	0.12
Average hourly temperature	°C	Training	7.89	7.65	0.38%	3.20	7.30	11.95
		Validation	8.24	7.66	0.28%	3.70	7.90	12.40
		Test	8.24	7.66	0.28%	3.70	7.90	12.40
Solar PV	%	Training	27.21					
		Validation	25.76					
		Test	21.74					
Electric heating or heat pump	%	Training	25.74					
		Validation	25.76					
		Test	13.04					
Electric water heating system only	%	Training	16.18					
		Validation	15.91					
		Test	8.70					

Notes: Statistics computed at each variable's natural level of variation.

of possible values. Binary variables are not scaled, because they only take 0 or 1 as value. Other variables are scaled according to the methods described in Table 3.

Table 3: Scaling methods for non-binary variables

Variable	**Range**	**Scaling method**
Electricity consumption	$\mathbb{R}_{\geq 0}$	$\log(1+x)$
Hourly rainfall	$\mathbb{R}_{\geq 0}$	$\log(1+x)$
Hourly sunshine	$[0,$ period length in hours$)$	$\log(1+x)$
Average temperature	$\mathbb{R}$	z-score normalization
Daylight duration	$[0,$ period length in hours$]$	$\log(1+x)$

The choice of z-score normalization for the average temperature comes from the fact that this variable does not have a strict minimum and maximum that could be used as bounds. The mean and standard deviation over the training period—i.e., before July 7, 2024—are used as reference values.

For the other variables, the choice of the natural logarithm of the value plus 1 comes from the fact that they cannot take negative values. This method is the easiest to use without requiring scaling parameters like a mean, a maximum, or a standard deviation. In particular, for electricity consumption, determining which scaling parameters to use for z-score normalization would be difficult, given that values and distributions are specific for each household.

4 Model training and results

4.1 Setup and hyperparameters

Different versions of the same model have been trained by varying the frequency of input and output series. Although the base frequency is hourly, there is no imperative need for hourly predictions, given the nature of the prediction task. Hourly inputs have been used to train models that predict hourly, 3-hour, and 6-hour electricity consumption. In addition, a model taking 3-hour inputs to predict 3-hour output has been trained. The goal of training these different models is to identify whether accurate and usable forecasts can be obtained with higher or lower frequency data, and whether predictions obtained from lower frequency inputs impact the quality of outputs in a positive, neutral, or negative way. It should also be noted that models based on lower frequency input data contain less parameters, making them lighter to train and use.

For each combination of input and output frequencies, three models are trained: one for the median, one for the 30$^{\text{th}}$ quantile, and one for the 70$^{\text{th}}$ quantile. They provide an interval of potential consumptions, with the lower bound acting as target for the saving potential with respect to the median. The sole differences between the three models is the parameter γ from Equation (1).

A summary of all model settings and hyperparameters is displayed in Table 4. The loss function is a quantile loss with L2 regularization to limit overfitting. The learning rate and the dropout rate have been determined after a phase of exploration.

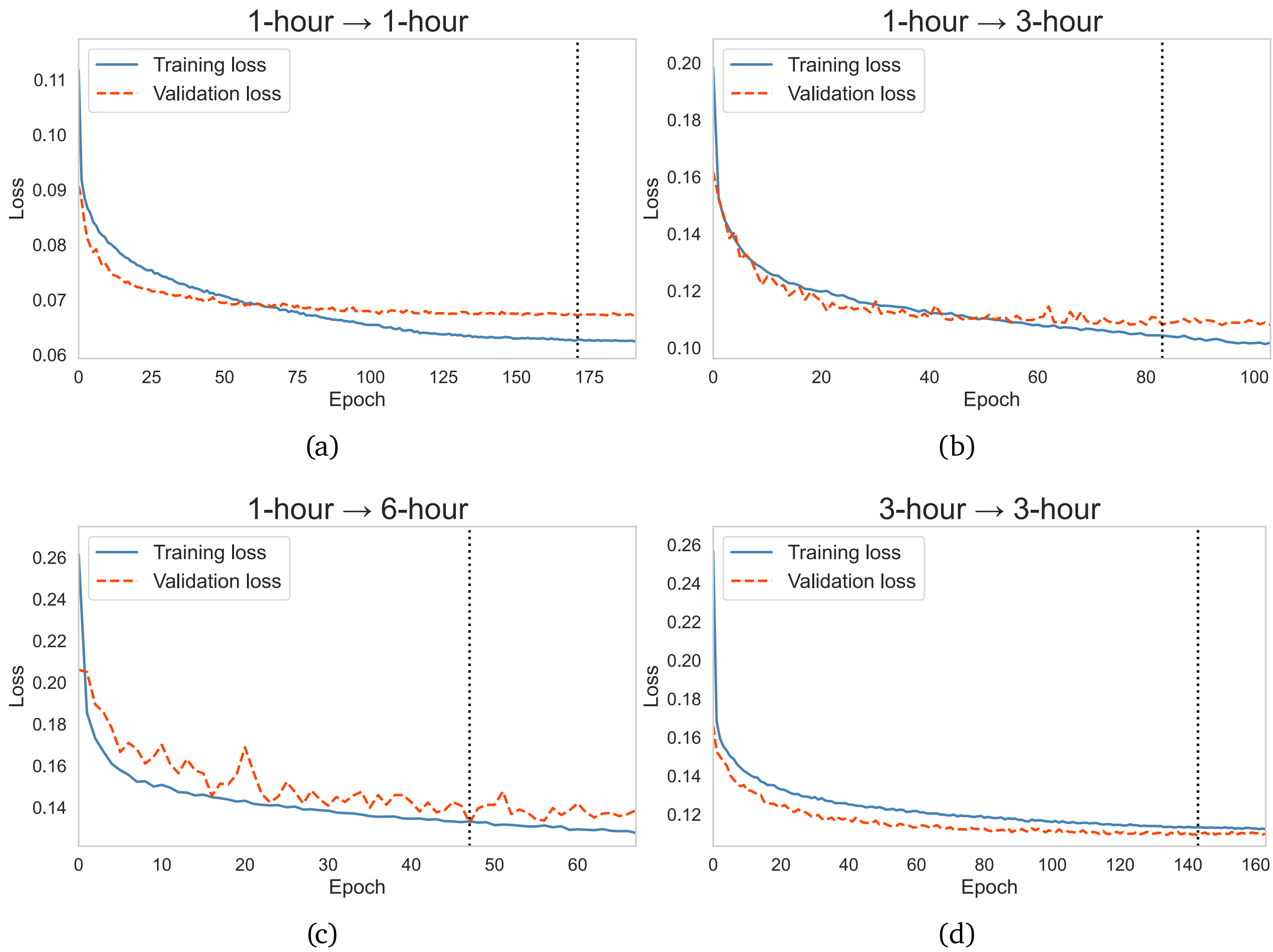
Table 4: Training setup

Setting	**Description**
Input-output frequencies	a) 1-hour $\rightarrow$ 1-hour: 168 time steps $\rightarrow$ 168 time steps b) 1-hour $\rightarrow$ 3-hour: 168 time steps $\rightarrow$ 56 time steps c) 1-hour $\rightarrow$ 6-hour: 168 time steps $\rightarrow$ 28 time steps d) 3-hour $\rightarrow$ 3-hour: 56 time steps $\rightarrow$ 56 time steps
Loss function	$\max\left(\gamma[y - f_\theta(x)], (1-\gamma)[f_\theta(x) - y]\right) + 0.5 \cdot 10^{-5}\|\theta\|_2^2$
Quantiles	$\gamma \in \{0.3, 0.5, 0.7\}$
Optimizer	Adaptive moment estimation (Adam)
Stopping criterion	20 epochs without improvement
Model selection	Model with lowest validation loss before the stopping criterion is met
Learning rate	Starting rate: 10^{-5} Scheduler: Halve learning rate after 10 epochs without improvement
Dropout rate	Feature extraction phase: 0.1 Regression head: 0.2
Batch size	42

4.2 Training and testing

The models have been coded in Python using the Pytorch library. The loss is minimized over the validation set using the parameters learned from the training set. All models have been trained until the stopping criterion was met. Training and validation losses have been recorded and are displayed in Figure 3 for the four models presented in Table 4 with $\gamma = 0.5$—i.e., models for the median.

Figure 3: Training and validation losses



(a) (b) (c) (d)

As can be seen, both training and validation losses tend to decrease over epochs, which is the behavior that is expected with adapted hyperparameters and measures to limit overfitting like random dropout or L2 regularization. However, with models with a 3-hour or 6-hour output frequency and a 1-hour input frequency, validation losses become unstable, which could indicate that trained models might not necessarily converge to an optimum due to overfitting or an overshooting of optimal model parameters. However, the losses for the model with 3-hour frequency for both input and output display a smoother behavior, which tend to support the idea that matching input and output frequencies could provide better models.

An interesting point to highlight here is that models with 1-hour input frequency have more than 14M parameters, whereas the model with 3-hour input and output frequency has only about 1.6M parameters, making it about 90% lighter to store and use—down from 55MB to 6.5MB. It is useful to consider this fact from a practical perspective, if one were to plan to implement it as part of an application to provide information on electricity consumption forecasts to consumers.

Table 5 shows average prediction errors over the three sets for all models that forecast the median. Presented metrics are mean relative average error (MRAE)[2], mean absolute error (MAE), and mean squared error (MSE). MRAE is the preferred metrics, as it does not depend on the scale of forecast units, while the others do. It is therefore easier to compare models with MRAE, whereas it is not possible for models with different output frequencies.

Table 5: Prediction errors, by set

Model (a)			$1h \to 1h$
Error	**Training**	**Validation**	**Test**
MRAE	0.37	0.45	0.48
MAE	0.12	0.13	0.14
MSE	0.04	0.06	0.07

Model (b)			$1h \to 3h$
Error	**Training**	**Validation**	**Test**
MRAE	0.26	0.32	0.35
MAE	0.19	0.22	0.23
MSE	0.09	0.11	0.13

Model (c)			$1h \to 6h$
Error	**Training**	**Validation**	**Test**
MRAE	0.22	0.25	0.28
MAE	0.24	0.27	0.28
MSE	0.12	0.15	0.17

Model (d)			$3h \to 3h$
Error	**Training**	**Validation**	**Test**
MRAE	0.29	0.32	0.35
MAE	0.20	0.22	0.22
MSE	0.10	0.11	0.13

Models with a lower frequency output tend to perform better in terms of MRAE than the others. This result suggests that forecasting higher frequency outputs is more difficult, which seems reasonable in the case of electricity consumption, as it can strongly vary from one period to another if a specific domestic appliance is used or not, for instance. If the period considered is a single hour, an incorrect prediction can cause a large error, whereas if the period is longer, peaks and troughs should be smoothed, potentially lowering the magnitude of prediction errors.

Another interesting element appearing in Table 5 is that using 1-hour frequency input data does not seem to perform better than 3-hour frequency input data when the task is to predict a 3-hour frequency output. If no significant benefit can be obtained from higher frequency input data when the target has a lower frequency, it is useless to use input data with a higher frequency than output data, in the present case.

[2]Due to 0s in the data, relative errors can take values outside $\mathbb{R}$. Relative errors have thus been computed from average errors with respect to the average actual value for each time step, and then averaged over the whole sequence.

4.3 Forecasting saving potential

In this paper, the electricity saving potential for a household during a given week is defined as the forecast obtained for the 30th quantile of its electricity consumption over that week compared to the forecast obtained for the median. The model performance needs not be perfect in terms of accuracy, as its purpose is to provide consumers with information on what their electricity consumption could look like if it were lying in the lower part of its distribution.

An issue is consumption peaks, as they can be hard to accurately predict if they are not happening regularly at the same time of the day or week. Automatic consumption peaks, like a water boiler that starts heating water at fixed times, should be easier to forecast, as long as consumption patterns remain stable over time. More generally, erratic electricity use patterns are too random to be forecast, if they are not correlated with other parameters that could be included as inputs to the model. Therefore, it should not be expected from the model that it can accurately predict such patterns, which are unpredictable by nature.

Examples of forecasts are displayed on Figure 4. Predictions obtained for the 30th, 50th, and 70th quantiles are shown along actual values. Results from the lighter model that uses 3-hour frequency inputs to predict 3-hour frequency outputs have been used to construct the plots, as they are more readable and present more stable results than those from the models with 1-hour frequency input data. Household numbers displayed in the plots' titles correspond to their ID numbers in the original database. They all belong to the test set and their characteristics related to solar PV and heating system are presented in Table 6. Dates of the weeks shown in the six plots are mentioned in their titles.

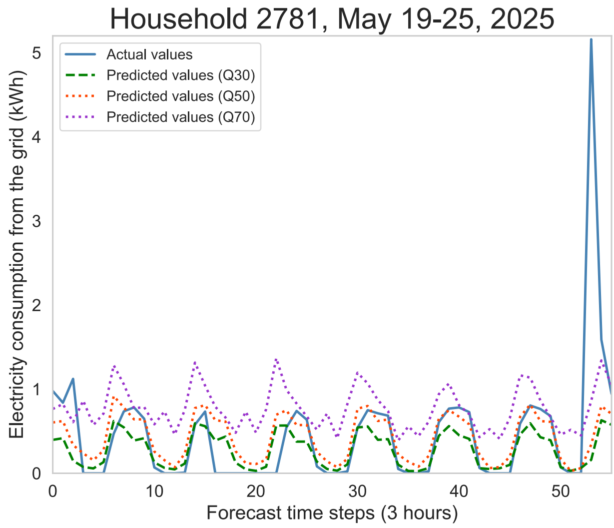
Table 6: Characteristics of example households

Household	Solar PV	Electric heating	Electric boiler only
2781	Yes	No	No
6367	Yes	Yes	No
973	No	No	No
3951	No	Yes	No
3163	No	No	Yes
6434	No	No	No

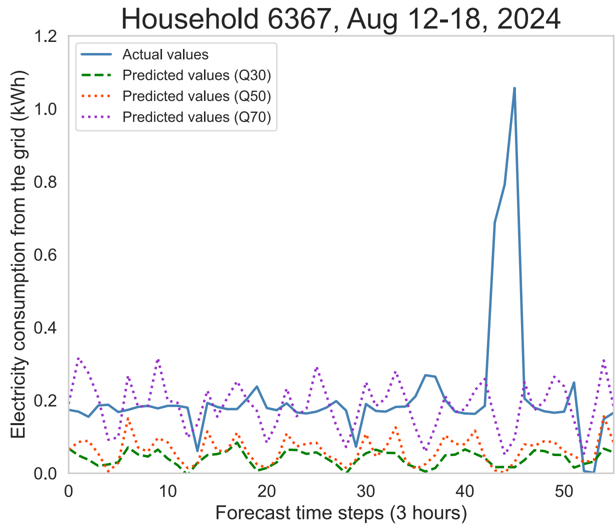
As can be seen, the quality of predictions varies from one household/week to another. Stable patterns are relatively well predicted, whereas irregular consumption peaks are generally not predicted at all. For some households (6367, 3951, and 6434), there might be an issue of bias between the predictions and actual values, the former seeming to have too low an intercept with respect to the latter. It might also be that actual values for these weeks belong to the upper quantiles of the distribution of potential values. It is nevertheless interesting to see that the models provide some useful information on the distribution of potential consumption patterns, especially to inform on the potential for electricity saving by comparing the 30th and 50th quantiles.

A noteworthy element is that forecasts obtained for Household 2781, who has solar PV and

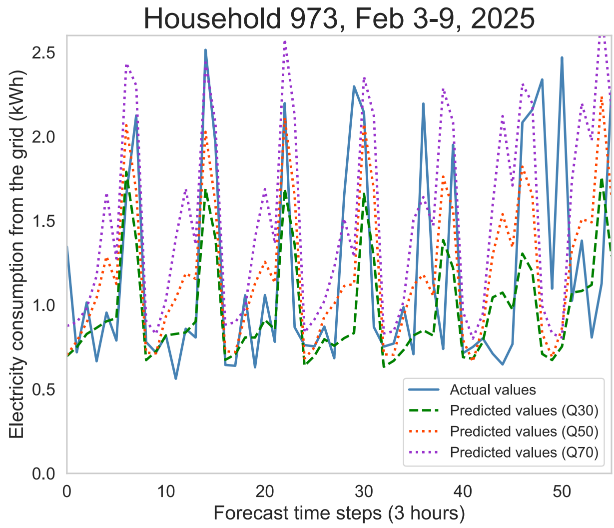
Figure 4: Forecasts and saving potential, 3-hour inputs and outputs



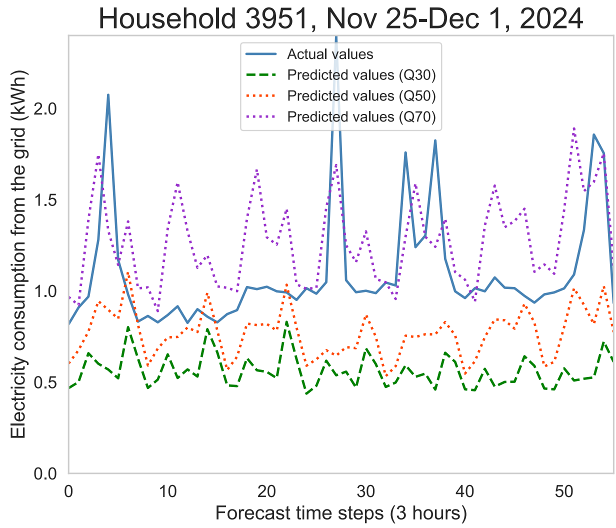
(a)



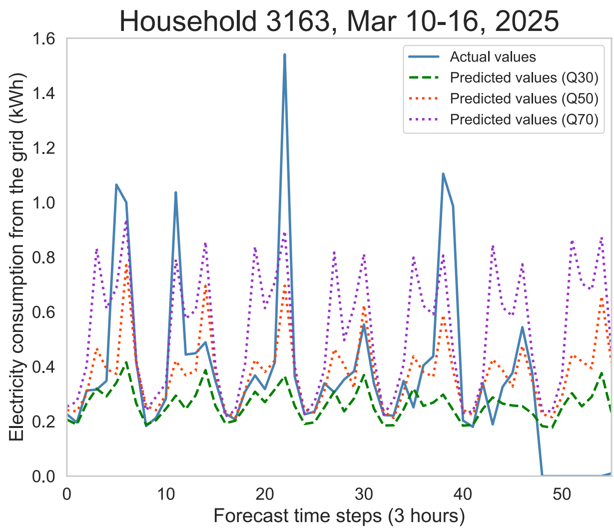
(b)



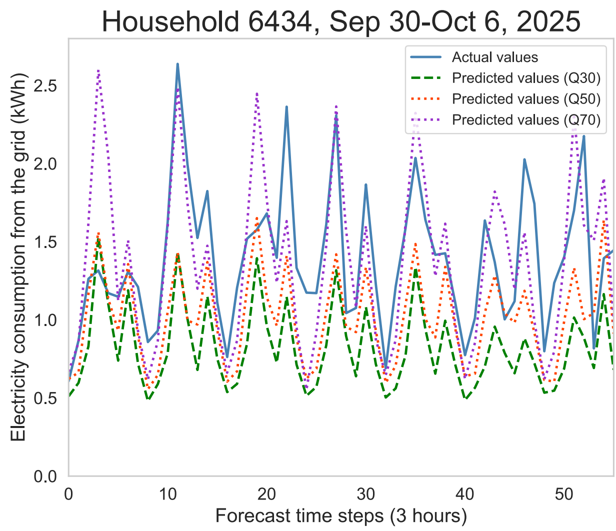
(c)



(d)



(e)



(f)

experienced 0-consumption hours during the week displayed in Figure 4a, are relatively accurate, which tend to show that 0s do not necessarily cause major issues to the model in terms of aberrant values. However, in the total results, no 0 has been correctly predicted as such by any model, which is not necessarily surprising, given the type of loss function used. An option could be to edit the loss function by adding a penalty term for inaccurately predicted 0s, but that would require a significant amount of fine-tuning to determine the optimal level of penalty to apply and to avoid that wrongly predicting 0s becomes less costly than predicting another value. Another option would be to extend the model to first determine whether the value to predict should be a 0 or another value before proceeding to the rest of the training.

The model could also certainly be improved with more data, especially regarding the number of households in the dataset, which is very low. A model with more data could include more input variables, with a better potential to learn more useful features, and hence be more accurate. However, the issue of the randomness of consumption peaks cannot be solved if those are purely random. A more sophisticated model could potentially estimate probabilities for consumption peaks and include them into predictions.

5 Conclusion

5.1 Main outcomes and limits

The 1D-CNN models trained to forecast quantiles of future electricity consumption at household level provide some usable predictions, which show that such an approach can be applied to solve the task of forecasting one-week ahead electricity saving potential. The choice of a 1D-CNN with MTS data is therefore appropriate, in this case. The main outcome from the research carried out here is that the modeling strategy that has been followed could probably be extended to a larger scale to provide customers with information on their electricity saving potential.

The presence of 0s in the data for solar PV owners is not a significant issue, although the models tend not to be able to appropriately forecast those 0s as such. The model could be extended by adding a component that specifically predicts 0s or by adding a penalty term to the loss function.

Using high frequency data does not really improve the quality of predictions. This can be explained by the fact that lower frequency data tend to smooth peaks and troughs, canceling the impact of their randomness within each time step—e.g., if a consumption peak usually happens in the morning between 6am and 9am, the model with 3-hour input frequency will not be affected by the fact that this peak occurs at 7am or 8am, whereas the 1-hour input frequency models will. The imputation of missing values is also less necessary when one does not need very high frequency time series: with electricity consumption data, all meters record cumulative consumption, so if data fails to transmit information during one hour, it is still possible to know how much electricity has been used in the meantime. Lower frequency data are less affected by such random events.

Another argument in favor of lower frequency data is that the model weight significantly decreases, as less parameters must be estimated. This improves the portability of the model and can ease its implementation within e.g. an application that tells consumers about their forecast electricity consumption for the coming week and their saving potential.

Even though the approach to the prediction problem seems satisfactory with respect to the results that have been obtained, several shortcomings would still need to be solved if one were to plan to train and use such a model as part of a program aiming at encouraging electricity use reduction. The main issue is the small size of the dataset, especially regarding the number of households. It means that the model has little between-household variance from which to learn, which limits its capacity to include household-level variables as inputs. Training a better model would certainly require hundreds more households to perform better.

Another data issue could be weather and climate. The data from MeteoSwiss comes from ground stations, which provide data only on what happened at their location. Having weather and climate records at a more local level could also help improve the model.

5.2 Policy implications

The analysis carried out in this paper shows that using a relatively simple machine learning model to forecast quantiles of electricity consumption at household level with limited data is possible. The quality of the model trained at this stage may not be sufficiently good to be directly used as part of program aiming at providing incentives to lower electricity consumption, but the general approach with a 1D-CNN seems to be adapted. The resulting model is light, which is an advantage in terms of portability.

The main limit to the quality of the model is that only limited data are available to train it. With more data, it should be possible to train a model that could be used to show a weekly electricity consumption reduction target to households. In particular, more household-level factors could be included, which should result in more accurate predictions that account for each household's characteristics.

A recommendation to develop this project a step further would be to cover a larger region than the Jura mountains. Various households from different regions of Switzerland could hence contribute to the construction of a more general prediction model for electricity use and saving potential that could be deployed over a wider territory.

References

Chae, Y. T., Horesh, R., Hwang, Y., and Lee, Y. M. (2016). Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. *Energy and Buildings*, 111:184–194.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

Hamilton, J. D. (1994). *Time series analysis*. Princeton university press.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.

Kiprijanovska, I., Stankoski, S., Ilievski, I., Jovanovski, S., Gams, M., and Gjoreski, H. (2020). Houseec: Day-ahead household electrical energy consumption forecasting using deep learning. *Energies*, 13(10):2672.

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., and Inman, D. J. (2021). 1d convolutional neural networks and applications: A survey. *Mechanical systems and signal processing*, 151:107398.

Klyuev, R. V., Morgoev, I. D., Morgoeva, A. D., Gavrina, O. A., Martyushev, N. V., Efremenkov, E. A., and Mengxu, Q. (2022). Methods of forecasting electric energy consumption: A literature review. *Energies*, 15(23):8919.

Li, Z., Liu, F., Yang, W., Peng, S., and Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 33(12):6999–7019.

Lim, B., Arık, S. Ö., Loeff, N., and Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International journal of forecasting*, 37(4):1748–1764.

Liu, C.-L., Hsaio, W.-H., and Tu, Y.-C. (2019). Time series classification with multivariate convolutional neural network. *IEEE Transactions on industrial electronics*, 66(6):4788–4797.

Marvuglia, A. and Messineo, A. (2012). Using recurrent artificial neural networks to forecast household electricity consumption. *Energy Procedia*, 14:45–55.

Nti, I. K., Teimeh, M., Nyarko-Boateng, O., and Adekoya, A. F. (2020). Electricity load forecasting: a systematic review. *Journal of Electrical Systems and Information Technology*, 7(1):13.

Oh, Y., Lim, C., Lee, J., Kim, S., and Kim, S. (2024). Multichannel convolution neural network for gas mixture classification. *Annals of Operations Research*, 339(1):261–295.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

Yildiz, B., Bilbao, J. I., Dore, J., and Sproul, A. B. (2018). Short-term forecasting of individual household electricity loads with investigating impact of data resolution and forecast horizon. *Renewable Energy and Environmental Sustainability*, 3:3.

Appendix WP2

Prosocial Behavior under Ambiguity: The Role of Social Norms and Beliefs

Fabien Giauque* Mehdi Farsi* Michael Puntiroli*

January 24, 2026

Abstract

This paper examines how ambiguity and normative interventions jointly shape prosocial behavior through a series of three incentivized experiments. The data are analyzed using logistic regression models. The studies are conducted in the context of voluntary contributions to the energy transition, a domain in which individuals often face ambiguity regarding costs, consequences, and responsibilities. Across experimental designs, ambiguity consistently reduces prosocial behavior. Norm-based interventions partially offset these effects; however, their effectiveness varies depending on both the type of ambiguity and the form of normative information provided. In addition, substantial heterogeneity in individual responses is identified, particularly with respect to the role of normative beliefs and self-centered value orientations. Finally, emotional responses differ across experimental conditions among donors but not among non-donors, with notable variation in positive emotions such as pride, empathy, and belonging.

Keywords Ambiguity; Normative Information; Incentivized Experiment; Donation; Energy Transition; Artificial Intelligence

JEL Codes D12, D91, Q48,

We acknowledge the financial support of Swiss Federal Office of Energy under grant SI/502530-01 (project INFINEED).

*University of Neuchâtel, Institute of Economic Research, Switzerland.

1 Introduction

Supporting the energy transition requires not only technological progress and public investment, but also sustained individual contributions. Beyond household energy-saving behaviors, citizens can contribute financially to climate and energy goals through voluntary transfers—such as donations to carbon-offsetting initiatives or to organization that fund renewable energy, energy efficiency, or other mitigation projects. Such contributions represent a form of prosocial behavior with a public good dimension: they entail a private cost while generating collective benefits. Understanding the determinants of these voluntary transfers is thus relevant for both behavioral research and policy, especially in a context where governments increasingly rely on complementary private action to close mitigation gaps and finance decarbonization.

A central challenge for prosocial engagement in environmental contexts is that the structure of decisions is often characterized by ambiguity. Environmental actions typically involve uncertain consequences, unclear baselines, and imperfect information about harms and benefits (Baillon et al., 2012). This creates opportunities for individuals to reinterpret the moral meaning of their choices. A large experimental literature shows that when responsibility is blurred, prosocial behavior can decrease because individuals can maintain a positive self-image while acting more selfishly—a mechanism commonly referred to as moral wiggle room (Dana et al., 2006, 2007). Then, ambiguity is defined in the language of economics as a state characterized by unknown probabilities (Ellsberg, 1961). This alters the evaluation of the cost of acting prosocially. In such settings, individuals may adopt more cautious strategies and avoid committing resources when the true budget constraint is not fully known (Klibanoff et al., 2005). These two perspectives suggest distinct but complementary pathways through which environmental prosociality may be undermined.

A potential countervailing mechanism is the provision of social norms (Schultz and Kaiser, 2012). While social norms have been extensively studied as drivers of prosocial behavior, their role in shaping decisions under conditions of ambiguity has received comparatively little attention. Norm-based interventions can reduce the behavioral consequences of ambiguity by clarifying what is socially appropriate (injunctive norms), what others expect (normative beliefs), or what others do (descriptive norms) (Cialdini and Goldstein, 2004). In ambiguous environments, such normative interventions can restore moral clarity and shift attention toward fairness, responsibility, and collective outcomes (Rabin, 1993). At the same time, the effectiveness of norms is likely to be heterogeneous as individuals differ in their values and attitudes, which can shape whether normative information is internalized, ignored, or even resisted (Schultz et al., 2007).

This paper studies how ambiguity and normative messages jointly shape prosocial support for the energy transition in a series of online incentivized experiments. Across three studies, we implement controlled manipulations of ambiguity and normative information while measuring real monetary transfers linked to environmental outcomes. The three studies deliberately vary the source and nature of ambiguity: Study 1 introduces attributional ambiguity in a charitable dictator game; Study 2 examines ambiguity about the personal and environmental benefits; Study 3 studies ambiguity about costs and budget constraint. This multi-study design allows us to assess whether the behavioral impact of ambiguity generalizes across contexts and to identify when and for whom normative interventions can mitigate its effects.

Across studies, our results provide consistent evidence that ambiguity undermines prosocial behavior, while also shaping how individuals respond to normative information. Overall, the findings highlight ambiguity as a key mechanism enabling moral disengagement and misperception, and show that normative interventions can partially—but selectively—counteracting this negative impact. Ambiguity is associated with a robust decline in prosocial choices, ranging from 5% to 16%. Distinct types of ambiguity call distinct mechanisms. On the one hand, ambiguity weakens the perceived link between individual actions and moral consequences, making it easier to justify selfish behavior without threatening one's moral self-image. This results resonates namely with Garcia et al. (2020)'s findings about excuse-driven behavior. On the other hand, ambiguity about personal and societal costs and benefits of prosocial choices allows to interpret the situation in desirable ways (Lind et al., 2019).

The effectiveness of social norms critically depends on the environment. In unambiguous settings, normative messages have little impact on behavior, suggesting that individuals rely primarily on their own assessments when the situation is clear (Van Teunenbroek et al., 2020). By contrast, when ambiguity is present, explicitly highlighting fairness and equal sharing partially offsets the negative effect of ambiguity (Rabin, 1993), while providing normative information in the form of social comparisons entirely countervail the effect. Notably, norm interventions that warns against using ambiguity as a justification prove largely ineffective. Although these results are consistent across experiments, the limited precision of the estimated average effects points to disambiguation through normative interventions as a relevant avenue for further investigation, while highlighting the importance of heterogeneity analysis.

The behavioral effects of ambiguity and norms are highly heterogeneous. Ambiguity is most detrimental for two distinct groups: individuals with strong prosocial normative beliefs and individuals with strongly egoistic values. For the former, ambiguity provides a valuable justification to deviate from fairness norms while preserving a positive self-image. For the latter, ambiguity weakens moral constraints and allows underlying self-interest to translate more freely into behavior. In contrast, ambiguity has weaker effects among individuals with moderate norms or low egoistic values. Normative messages reduce the impact of ambiguity primarily among individuals with at least some prosocial orientation, while strongly egoistic individuals remain largely unresponsive, even when fairness norms are made explicit.

Finally, ambiguity and normative information shape emotional responses to prosocial behavior. Among those who act prosocially, ambiguity amplifies positive self-conscious and social emotions such as pride, empathy, and belonging (Tarditi et al., 2020). Donating under ambiguity appears to be emotionally more rewarding, consistent with the idea that ambiguous contexts allow individuals to perceive their actions as especially virtuous (Tangney et al., 2007). However, this emotional amplification is attenuated when explicit normative guidance is provided, suggesting that norms reduce the perceived exceptionalism of the prosocial act. Notably, no systematic effects are observed for negative emotions such as guilt or shame, nor among individuals who choose not to act prosocially (Trujillo et al., 2021).

Overall, ambiguity alters perceived affordability, weakens moral constraints, and reshapes emotional payoffs in ways that reduce prosocial behavior. Normative messages can counteract these effects, but only when they clearly articulate the relevant moral standard and when individuals retain some openness to prosocial considerations. This highlights both the power and the limits of norm-based interventions in ambiguous decision environments, with important implications for policy design in the context of the energy transition and AI-assisted decision-making

The remainder of the paper is organized as follows. Section 2 presents the theoretical framework underlying the study. Section 3 describes the experimental procedure. Section 5 summarizes the sample characteristics and descriptive statistics. Section 6 presents the empirical results. Finally, Section 7 discusses the implications and concludes the paper.

2 Theoretical framework

Figure 1 presents the theoretical framework underlying the experiment. A central insight from the literature is that ambiguity can undermine prosocial behavior by weakening the perceived link between individual actions and their moral consequences. This interpretive flexibility allows individuals to behave more self-interestedly without incurring psychological or social costs. Experimental evidence shows that introducing ambiguity into charitable or distributive decisions reduces generosity, as individuals can plausibly justify selfish behavior to themselves and others (Dana et al., 2007; Van Der Weele et al., 2014; Fahrenwaldt et al., 2024), or reinterpret the harm when ambiguity applies to costs and benefits of prosocial behavior (Garcia et al., 2020). Additionally, when ambiguity applies on available resources, it creates uncertainty about the true cost of acting prosocially (Ellsberg, 1961). This ambiguity can encourage more self-centered decision-making, as individuals may avoid or delay costly actions when the same contribution could represent either a small or a substantial share of their eventual endowment (Karlan and List, 2007). Ambiguity is therefore expected to reduce prosocial behavior, leading to lower donations compared to an unambiguous control condition. Furthermore, the effect of ambiguity is likely to depend the magnitude of required prosocial effort (Exley, 2016; Grossman and Van Der Weele, 2017; Grossman, 2014).

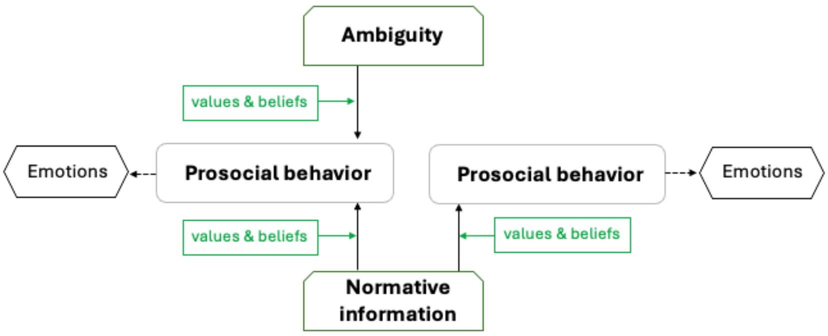


Figure 1: Theoretical framework

Social norms constitute a key mechanism through which behavior can be steered toward socially desirable outcomes (Bicchieri, 2023; Van Teunenbroek et al., 2020). In the proposed framework, normative information enters the decision environment, thereby influencing prosocial behavior. Normative information may mitigate the negative effect of ambiguity by clarifying socially appropriate conduct or by shaping how individuals interpret uncertain situations (De Nardo et al., 2017). To isolate the effect of social norms on mitigating ambiguity, we implement normative information in both ambiguous and unambiguous choice environments. This design allows us to distinguish the extent to which normative information specifically counteracts ambiguity-driven selfish behavior, rather than exerting a general prosocial influence. We therefore hypothesize that the presence of normative information increases prosocial behavior by attenuating the negative effect of ambiguity on donations.

The effects of ambiguity and normative information are unlikely to be uniform across individuals, as both operate through values and beliefs that vary across the population. Two key moderators are therefore expected to shape behavioral responses. First, preexisting normative beliefs about prosociality may play a central role in norm compliance. Individuals who believe that selfish behavior is common may be more likely to interpret ambiguity in ways that legitimize self-interest. Second, egoistic values reflect stable preferences that prioritize self-interest and personal advantage. Within the proposed framework, these values shape how individuals process both ambiguity and normative information. Individuals who strongly endorse egoistic values are more likely to exploit ambiguity to justify selfish behavior. Finally, we also examine the emotional responses associated with prosocial behavior. Emotions such as guilt, shame, pride, empathy, and feelings of belonging are measured as affective outcomes of the choice and provide insight into the psychological mechanisms through which ambiguity and normative information operate. These emotional responses are expected to vary with prosocial behavior and reflect the extent to which individuals experience their choices as morally salient and socially embedded[1].

3 Experimental procedure

The three studies implement similar experimental procedure, consisting of (i) an incentivized task or information stage, (ii) a prosocial decision, and (iii) the elicitation of individual-level moderators. Participants are recruited from the Swiss population using quota-based sampling to ensure representativeness in terms of age, gender, and linguistic region, and are randomly assigned to experimental conditions in between-subjects designs. Study 1 combines a real-effort task and a charitable dictator game (CDG). Participants first complete an incentivized slider task, the performance of which determines a monetary endowment. They are then exposed to experimental conditions before allocating their earned endowment between themselves and a organization active in the energy transition, providing an measure of prosocial preferences.

Study 2 combines an AI-assisted cognitive task and a prosocial decision. Participants complete an incentivized categorization quiz that generates a monetary payoff and then make a binary decision regarding whether to compensate for CO_2 emissions generated during the task[2]. Study 3 similarly focuses on

[1] The pre-registration is available at the OSF: `https://osf.io/2my3s/overview?view_only=f9498ec40b7b442ab719f86128c9f57b`.

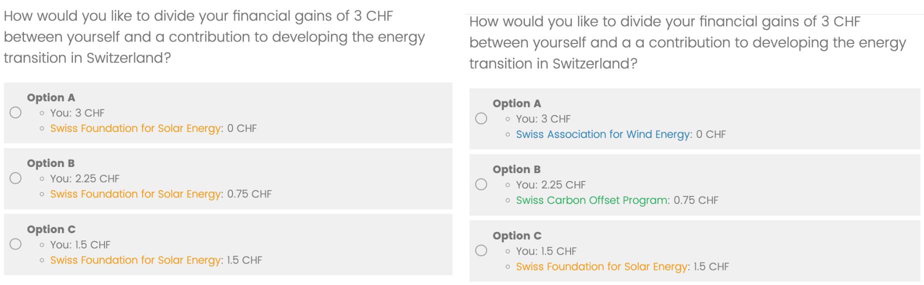
[2] See Figures 5 and 6 in the appendix for illustration

donation behavior related to AI use and emissions[3]. Participants receive standardized information about AI and its environmental impacts and subsequently decide how much of an additional monetary bonus to donate to compensate for associated CO_2 emissions[4]. Across all studies, experimental conditions vary the presence and nature of ambiguity and social norm information. Post-decision questionnaires elicit normative beliefs, values, emotional responses, image concerns, and attitudes toward energy and technology. These measures are used to capture individual heterogeneity and to support the analysis of behavioral responses to the experimental conditions.

4 Experimental condition

4.1 Ambiguity

Ambiguity is likely to be a key factor moderating individuals' engagement. However, the magnitude may vary depending on its nature. Consequently, we explore ambiguity regarding (1) motives, (2) personal and societal benefits, and (3) the costs and budgets faced by individuals making a prosocial decision. Figure 2 illustrates how ambiguity is implemented in the study (1), by varying the identity of the donation recipient across experimental conditions. In the control group (a), the recipient of the donation is identical across all allocation options. Regardless of whether participants choose to keep the full amount, share part of it, or split it equally, the contribution is always directed to the same organization. This design makes the social meaning of each choice transparent: more generous allocations unambiguously signal stronger prosocial motivation, while more selfish choices reflect self-interest.



(a) Control group (b) Treated group

Figure 2: Ambiguity (1)

[3] See Table 7 in the appendix for a comparisons between reported and predicted distributions of AI usage.

[4] See Figures 7 and 8 in the appendix.

In contrast, the treated condition (b) introduces attributional ambiguity by varying the donation recipient across options. While the monetary splits remain identical to those in the control group, each option is associated with a different organization involved in the energy transition. As a result, a participant can choose a more selfish allocation while plausibly attributing this decision to preferences over charities rather than to a lack of prosocial concern. This creates room for acting more selfishly while preserving a positive self-image or social image. To ensure that observed differences in behavior are driven by ambiguity rather than by preexisting preferences for specific organizations, the assignment of charities to options is counterbalanced across participants. This procedure prevents systematic bias due to heterogeneous attitudes toward particular charities and ensures that, on average, no option is intrinsically more attractive than another because of the recipient's identity.

Study (2) introduces ambiguity within a real-effort cognitive task[5]. All participants complete the same categorization quiz, consisting of 10 rounds. In each round, participants have 20 seconds to sort four items into two categories using a drag-and-drop interface. Correct placements are financially incentivized. Participants can either rely on their own judgment or choose to use an AI-generated response, which guarantees a correct answer for that round and therefore ensures the full monetary payoff. The use of AI, however, entails a social cost in the form of CO_2 emissions. In the no-ambiguity condition, the social cost and the financial reward are made fully transparent. Participants are explicitly informed that each use of the AI generates 2 grams of CO_2 emissions and each correct answer earns 0.1 CHF. By providing a precise and quantitative description of the environmental cost and financial gain, this condition leaves little room for interpretation regarding the consequences of using AI. In contrast, in the ambiguity condition, participants are informed that using the AI generates CO_2 emissions, but no quantitative information is provided. The magnitude of the environmental cost and the personal benefits remain unspecified, weakening the clarity of the trade-off between personal financial gain and environmental harm.

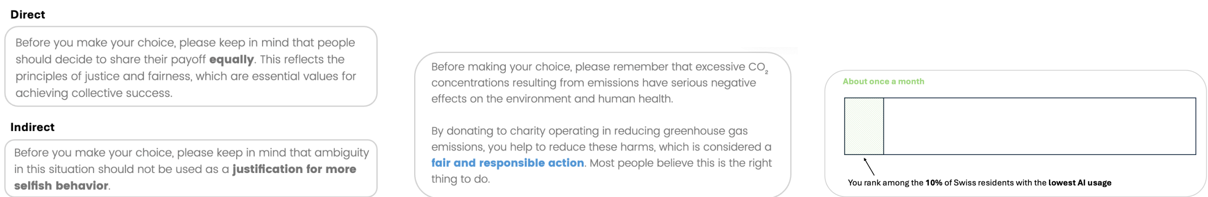
In study (3), ambiguity surrounds the monetary endowment available in the offsetting decision. Participants are randomly assigned to one of two conditions that differ only in the transparency of the bonus amount. In the no-ambiguity condition, participants are informed that they have been selected to receive a fixed bonus of CHF 3.–. This amount is known with certainty at the time of the decision and will be paid after the experiment together with the participation fee. The size of the endowment is therefore fully

[5] See Figure 5 in the appendix.

transparent, allowing participants to form precise expectations about the costs. In the ambiguity condition, participants are told that they will receive a bonus ranging from CHF 1 to CHF 5, but that the exact amount will only be revealed after the experiment is completed. While participants know the bounds of the distribution, they do not know the probability of each possible outcome. This introduces ambiguity about the relative cost of the prosocial option. Importantly, apart from the information regarding the endowment, all other aspects of the experimental environment are held constant across conditions.

4.2 Normative intervention

Figure 3 illustrates the normative messages provided across the experiments. In study (1), participant receive two distinct normative injunction. A ”direct” norm, where participants read an explicit prescriptive message emphasizing what ought to be done. The message highlights equality as the appropriate course of action by stating that people should decide to share their payoff equally, explicitly invoking principles of justice and fairness. An ”indirect” norm, which does not prescribe a specific allocation. Instead, it addresses the decision context itself by cautioning participants against using ambiguity as a justification for more selfish behavior. By explicitly condemning the exploitation of moral wiggle room, this message seeks to limit self-serving reinterpretations of the situation without directly instructing participants to act generously. Prosocial behavior is therefore encouraged indirectly by closing the wiggle room.



(1) (2) (3)

Figure 3: Normative intervention

In study (2), the normative intervention combines elements of injunctive and subjective norms to guide behavior. Participants are reminded that CO_2 emissions have serious negative consequences for the environment and human health, thereby reinforcing the moral relevance of the decision. The message then explicitly frames donating to a charity that reduces greenhouse gas emissions as a “fair and responsible action,” which constitutes a clear injunctive norm by signaling what one ought to do. In addition, the intervention includes a subjective norm component by stating that “most people believe this is the right thing to do.” This element appeals to individuals’ beliefs about others’ moral expectations and conveys social approval of the prosocial choice, without directly reporting observed behavior. By doing so, the message increases the perceived social legitimacy of compensating for emissions.

In study (3), the manipulation unfolds in three steps. First, participants are asked to report how frequently they engage with AI tools. Responses are collected on an ordered scale ranging from "never" to "several times an hour." This self-reported frequency serves as the basis for constructing individualized social comparison feedback. Second, at the beginning of the information section, participants receive a graphical comparison that situates their reported AI use relative to others. Those in the bottom half are informed that they belong to the group that uses AI tools the least, while those in the top half are told that they rank among the highest AI users. This feedback is presented visually through a bar chart, making relative standing salient and easy to interpret. Third, at the end of the information section, participants are reminded of their relative position in the AI-usage distribution. They are then asked to reflect on this comparison by responding to a set of five statements tailored to their assigned group. This two-step exposure—initial feedback followed by a reflective reminder—ensures that the social comparison is both salient and cognitively processed, thereby strengthening the treatment.

5 Data

The empirical analyses draw on three independent samples of participants living in Switzerland (Study 1: N = 793; Study 2: N = 843; Study 3: N = 1,199). Across studies, samples are broadly representative in terms of gender, age, and linguistic region (excluding the Italian-speaking part). Tables 5 report descriptive statistics for sociodemographic characteristics and psychological variables. All studies elicit normative beliefs capturing participants' beliefs about others' prosocial behavior, as well as egoistic values reflecting the importance attached to self-oriented guiding principles such as power, wealth, and ambition. Study 1 further includes perceived ambiguity and attitudes toward the energy transition as controls. Additional measures capture individual heterogeneity in image concerns, trust, emotional responses (guilt, pride, shame, empathy, belonging), values (hedonism, altruism, biospherism) and attitudes toward technology. All attitudinal and emotional measures are elicited after the experimental decision and are used to account for heterogeneity in preferences and responses to the experimental treatments.

Table 6 reports descriptive statistics for study (2) on AI use, prosocial behavior, and monetary outcomes. Overall, 43.8% of participants used AI at least once, with an average of 3.29 uses (out of 10) among users, indicating selective rather than mechanical reliance on AI. Across the full sample, 26.5% of participants made a prosocial choice. For participants who did not use AI, prosociality takes the form of a donation calibrated to the average potential carbon offset, chosen by 20.4% of participants. Among AI users, prosociality corresponds to carbon offsetting, with 34.1% opting to offset their AI-related emissions. This design ensures that prosocial behavior is defined consistently in monetary terms across participants. Participants earned on average 2.93 CHF from the task. Average prosocial contributions amount to 0.08 CHF across the full sample, reflecting the low incidence of prosocial choices, but rise to 0.31 CHF (10.5% of total gains) conditional on contributing. Final payoffs average 2.85 CHF.

6 Results

6.1 Treatment effect

Table 1 reports the effects of introducing ambiguity and normative messages across the three studies, expressed as average marginal effects (AMEs) derived from logistic regression models. Consistent with our theoretical expectations, ambiguity has a statistically significant negative effect on prosocial behavior. Exposure to ambiguity reduces the probability of making a prosocial choice by 10.8, 16, and 5.7%, respectively. These findings provide robust evidence that attributional ambiguity as well as ambiguity regarding societal benefits undermines prosocial behavior, in line with the theoretical expectations: when the link between individual actions and moral outcomes or societal benefits is less transparent, individuals are more likely to behave selfishly. By contrast, the evidence for ambiguity concerning costs is weaker ($p < 0.1$) and the estimated effect is smaller. This form of ambiguity may generate asymmetric expectations among individuals, leading to partially mitigate average effect.

Introducing social norms in unambiguous decision environment yields effects that are close to zero and statistically insignificant, suggesting that in the absence of ambiguity individuals do not rely on normative information when making prosocial choices. By contrast, providing normative messages in the presence of ambiguity can mitigate the adverse effect, although the success of this mitigation differs by the type of normative message. The direct norm intervention, which explicitly emphasizes fairness and equal sharing, has a positive statistically significant effect on prosocial behavior. Although these effects are only weakly significant, they are consistent across contexts, corresponding to increases in donation probabilities of 7.7, 3.8, and 5.8%. These estimates represent a mitigation of the negative effect

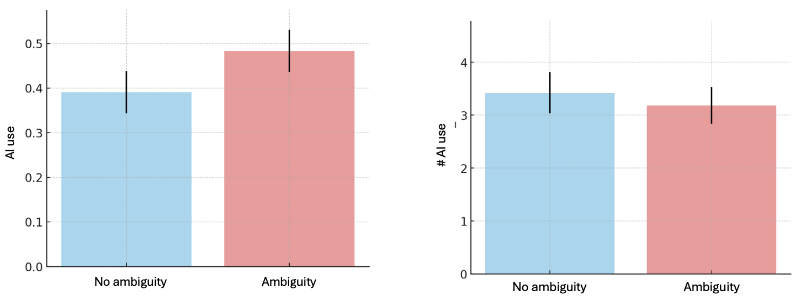
Table 1: Marginal treatment effects

	(1)	(2)	(3)
Ambiguity	-0.108**	-0.160**	-0.057*
	(0.046)	(0.046)	(0.046)
no-ambiguity			
Normative messages		-0.027	-0.017
		(0.045)	(0.045)
ambiguity			
Normative messages	0.077*	0.037*	0.058*
	(0.045)	(0.044)	(0.044)
Indirect	0.018		
	(0.044)		

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Column (1), (2), (3) report results from study (1), (2), and (3). Coefficients report average marginal effects of the experimental conditions estimated from a logistic regression model. The dependent variable is binary and capture prosocial choice. The control group serves as the reference category.

of ambiguity of 71.3, 23.1, and 101.7%, respectively. By contrast, the indirect norm intervention yields a small and statistically insignificant effect. This suggests that warning against the use of ambiguity as a justification for selfish behavior is insufficient to to substantially counteract the impact of ambiguity.

Beyond the impact on donation, it is worth noting that ambiguity leads to increased use of AI-assistant. Figure 4 illustrates the effect of ambiguity on AI use along the extensive margin (whether participants use AI at least once) and the intensive margin (frequency of use, conditional on adoption). Panel (a) shows that ambiguity significantly increases the likelihood of using AI at least once, a result that is statistically significant and illustrates greater reliance on AI when its costs and benefits are less transparent. Panel (b) examines AI use intensity among adopters. In contrast to the extensive margin, ambiguity does not increase usage intensity. Overall, these results suggest that ambiguity primarily affects the decision to adopt AI, rather than the intensity of use once adoption has occurred.



(a) AI use

(b) #AI use

Figure 4: AI use and frequency

6.2 Heterogeneity

We examine whether the effects of ambiguity and normative intervention vary systematically with individual differences in prosocial normative beliefs and egoistic values. Table 2 reports heterogeneous marginal treatment effects by splitting both moderators into low, moderate, and high categories. Across studies, the negative effect of ambiguity is heterogeneous and operates through two main channels. First, ambiguity has a large and statistically significant negative effect among individuals with low egoistic descriptive norms, that is, those who believe that most others would share equally. For these individuals, ambiguity substantially reduces prosocial behavior. Holding strong subjective fairness norms, they face higher psychological costs when deviating from equal sharing under clarity. Ambiguity provides a particularly valuable justification to deviate from the norm while preserving a positive self-image. Second, ambiguity also significantly reduces prosocial behavior among individuals with high egoistic values. For this group, ambiguity weakens the clarity of the moral situation and relaxes social and moral constraints that would otherwise limit self-interested behavior. As a result, underlying preferences are more likely to translate into selfish choices. By contrast, ambiguity has weaker or insignificant effects for individuals with low egoistic values or moderate normative beliefs.

Table 2: Heterogeneous treatment effects

Binary DV: Donation	Normative Beliefs		Egoistic values	
	(1)	(2)	(1)	(2)
Ambiguity				
Low	0.211**	0.254***	0.031	0.024
	(0.090)	(0.420)	(0.078)	(0.363)
Moderate	0.094	0.076**	0.054	0.077**
	(0.053)	(0.385)	(0.080)	(0.380)
High	0.081	0.062*	0.153**	0.067*
	(0.080)	(0.367)	(0.070)	(0.390)
Ambiguity & Social comparisons				
Low	-0.569***	-0.627*	-0.216***	-0.280**
	(0.127)	(0.294)	(0.082)	(0.141)
Moderate	0.016	-0.045**	-0.199***	-0.262*
	(0.054)	(0.210)	(0.077)	(0.155)
High	-0.125	0.093	-0.069	-0.054
	(0.098)	(0.376)	(0.066)	(0.389)

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Coefficients report estimates from interactions between experimental conditions and three levels (low, moderate, high) of normative beliefs and egoistic values, estimated using logistic regression. Standard errors are reported in parentheses. The dependent variable is the binary decision to offset emissions. Columns (1), (2) report estimates from studies (1) , (2). The reference category is the control group (moderate).

Introducing social messages in ambiguous environments also leads to heterogeneous reactions. Among individuals with strong prosocial normative beliefs, adding an explicit fairness norm on top of ambiguity leads to a pronounced increase in prosocial behavior. These individuals appear particularly tempted to exploit ambiguity in the absence of normative guidance, but the explicit norm closes the moral wiggle room by reinforcing the relevant reference norm and limiting opportunities for self-justification. Individuals with low or moderate egoistic values also respond positively to the direct norm: ambiguity allows them to shift toward self-interest, while the fairness message pushes behavior back toward prosociality. In contrast, individuals with high egoistic values are largely unresponsive to direct norms, indicating that strongly self-interested preferences dominate normative considerations even when fairness is made explicit[6].

While these patterns are broadly consistent across studies, one difference is worth noting. In Study 2, individuals with particularly strong prosocial beliefs about others sometimes respond to ambiguity with higher prosocial behavior, suggesting that well-anchored normative expectations can offset—or even dominate—the detrimental effect of ambiguity. In such cases, adding normative messages provides little additional guidance. Taken together, the results indicate that ambiguity is most powerful for individuals who either strongly recognize the fairness norm and therefore need ambiguity to justify deviation, or who are strongly self-interested and benefit from ambiguity relaxing moral constraints. Normative messages can reduce the effect, but primarily for individuals with at least some prosocial orientation, while strongly egoistic individuals remain largely resistant to norm-based persuasion.

Table 3: Heterogeneous donation

Binary DV: CO_2 offset	Technophilia (1)	Tech. openness (2)	Trust science (3)	AI opinion (4)
Strongly disagree	-0.536**	0.296	-1.079**	-0.534**
	(0.267)	(0.657)	(0.433)	(0.215)
Somewhat disagree	-0.342**	-0.843***	0.123	0.042
	(0.171)	(0.276)	(0.221)	(0.166)
Somewhat agree	0.079	0.266*	1.089***	-0.183
	(0.142)	(0.147)	(0.147)	(0.151)
Strongly agree	-0.103	0.150	1.127***	-0.147
	(0.216)	(0.174)	(0.181)	(0.210)

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Coefficients report estimates from self-appreciation related technology (1-2), science (3), and AI (4), estimated using a logistic regression model. The dependent variable is the binary decision to offset emissions. Standard errors are reported in parentheses. The reference category is *neither agree nor disagree*.

Beyond heterogeneity in treatment effect, it is worth noting that prosocial behavior in the energy transition is also moderated by attitudes toward technology and science, as presented in Table 3. We find

[6] Indirect norm interventions exhibit weaker and less robust effects which are therefore not reported here.

that prosocial behavior is lower technophobes and individuals who are less open to technological change. We further find that prosocial behavior is higher among individuals with greater trust in science. Last, lower donation is observed among individuals distrusting science and AI technologies. Overall, these results suggest that reluctance toward technology, rather than technological enthusiasm, is associated with reduced prosocial engagement in this context, while confidence in science and AI is positively related to willingness to offset emissions.

6.3 Emotional responses

We next examine how ambiguity and normative information shape emotional responses, focusing on both self-conscious and social emotions. As presented in Table 4, two clear patterns emerge. First, among donors and in the absence of normative information, ambiguity substantially amplifies emotional responses. Donors exposed to ambiguity report significantly higher levels of pride, empathy, and belonging. This suggests that donations made in ambiguous contexts are accompanied by stronger emotional reactions. Second, this amplification effect is partially constrained once normative information is introduced. This pattern indicates that normative guidance dampens the additional emotional boost associated with donating under ambiguity. One interpretation is that norms reduce the perceived exceptionalism or voluntariness of the prosocial act, thereby limiting the extent to which ambiguity enhances positive self-evaluative and social emotions. In contrast, no statistically significant emotional effects are observed among non-donors, regardless of ambiguity or normative information. Moreover, analyses of negative emotions (guilt and shame)[7] reveal no systematic or statistically significant effects.

7 Conclusion

This study contributes to the ongoing debate on the effectiveness and underlying mechanisms of normative information by examining its interaction with different forms of ambiguity. Its primary novelty lies in incorporating normative nudges into decision-making environments characterized by different types of ambiguity. We identify distinct forms of ambiguity that consistently reduce prosocial behavior. Importantly, the mitigating effect of normative information depends on the nature of the ambiguity: normative messages are approximately twice as effective when ambiguity concerns individuals' motives

[7] Reported in Table 8 in the appendix.

Table 4: Ordered logistic regressions

Categorical DV: Emotional responses	Pride (1)	Empathy (2)	Belonging (3)
Donors			
no-norm / no-ambiguity	0.413*	1.137***	0.582**
	(0.213)	(0.211)	(0.206)
no-norm / ambiguity	0.881***	1.713***	0.787***
	(0.215)	(0.215)	(0.212)
norm / no-ambiguity	0.441**	1.067***	0.558**
	(0.213)	(0.210)	(0.208)
norm / ambiguity	0.664***	1.485***	0.621***
	(0.208)	(0.211)	(0.208)

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Coefficients report estimates from interactions between experimental conditions and the binary decision to offset emissions, estimated using an ordered logistic regression model. Standard errors are reported in parentheses. The dependent variable is a categorical measure of participants' emotional responses following the decision. The reference category is non-donors under the no-norm / no-ambiguity condition. N=1,199. Full estimates reported in Table 8.

rather than the personal and societal benefits of prosocial action. The empirical analysis is grounded in the context of the energy transition and the increasing environmental footprint of artificial intelligence development—two issues of growing relevance. While the study is context-specific, we argue that the findings are likely to generalize beyond this setting and to resonate with the broader literature on prosocial and sustainable behaviors. Future research could test this external validity by replicating similar experimental designs in alternative domains.

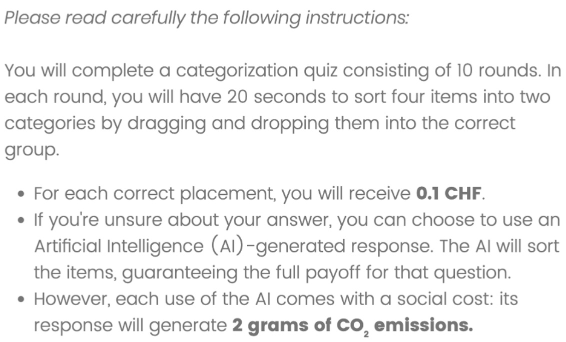
We also acknowledge that the estimated average effects are characterized by limited precision. This likely reflects the fact that both ambiguity and normative messages can elicit opposing behavioral responses across individuals, thereby underscoring the importance of heterogeneity analyses. Ambiguity leaves substantial room for individual interpretation of both the situation and the underlying motives of one's actions, which naturally leads to divergent responses among individuals with different attitudes, prior beliefs, and value orientations. In this respect, our analysis highlights the central role of egoistic values and normative beliefs in shaping behavioral responses. While individual values are relatively stable and difficult to influence directly, beliefs appear more responsive to intervention and therefore represent a potentially important channel for policy action.

8 Appendix

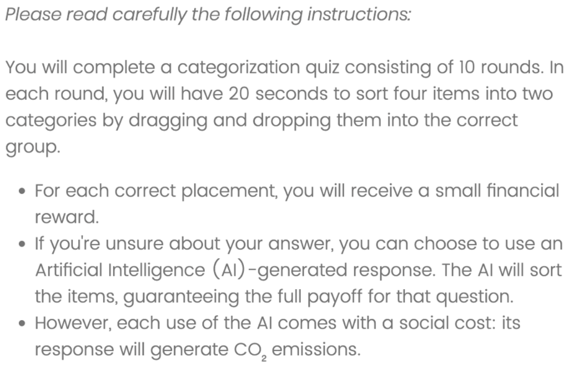
Table 5: Sample characteristics

	Study 1 (N = 793)	Study 2 (N = 843)	Study 3 (N = 1,199)
Binary variables (Share %)			
Female	50.4	49.8	51.5
German-speaking	68.0	72.1	71.6
Age 18–34	26.3	28.2	27.0
Age 35–54	38.0	35.3	37.1
Age 55–79	35.8	36.4	35.9
Compulsory school	6.2	6.3	6.3
Vocational	45.3	45.3	40.0
Baccalaureate	11.1	10.6	11.3
HES	19.9	18.6	19.2
University/ETH	17.5	19.2	23.3
Categorical variables (Mean (SD))			
Perceived ambiguity	3.26 (1.07)		
Energy transition	3.80 (0.73)		
Normative beliefs	3.46 (0.98)	2.89 (0.89)	2.87 (1.14)
Egoistic values	2.58 (0.88)	2.73 (0.80)	2.73 (1.01)
Altruistic values			3.50 (0.96)
Biospheric values			3.63 (0.97)
Hedonistic values			3.74 (0.88)
Guilt		1.86 (0.95)	1.69 (1.09)
Pride		3.35 (0.97)	3.49 (1.21)
Image concerns			2.78 (1.23)
Shame			1.54 (0.96)
Empathy			2.81 (1.30)
Belonging			2.69 (1.21)
Trust science			3.63 (1.00)
Technophilia			3.25 (1.03)
Technology openness			3.84 (0.86)
AI opinion		3.05 (1.12)	3.07 (1.07)

Notes: The table reports descriptive statistics for Studies 1–3. Shares are reported for binary variables and means with standard deviations in parentheses for categorical variables



(a) No ambiguity



(b) Ambiguity

Figure 5: Instructions cognitive task

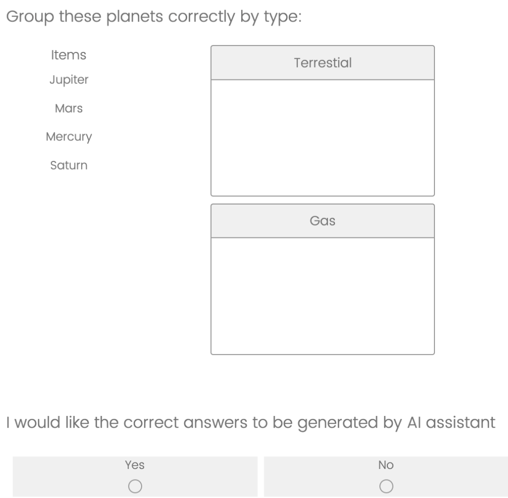
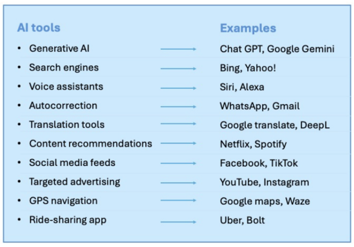


Figure 6: Example cognitive task

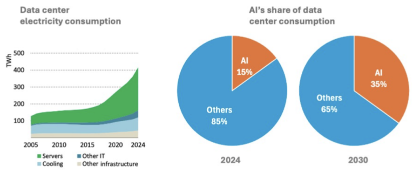
Table 6: Experiment descriptive

Binary variable	Share (%)	
AI help	43.8	
Prosocial choice	26.5	
donation	20.4	
carbon offset	34.1	
Continuous variable	**Mean**	**S.D.**
#AI (if AI = yes)	3.29	2.53
Total gains (CHF)	2.93	0.50
Prosocial (CHF)	0.08	0.17
Final payoffs (CHF)	2.85	0.49
Prosocial (CHF) (if yes)	0.31	0.19

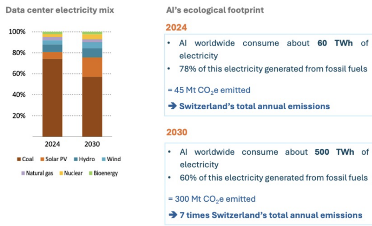
Notes: the table presents descriptive statistics for 843 participants.



(a) AI digital services



(b) AI electricity consumption



(c) AI ecological footprint

Figure 7: Informational messages

Table 7: AI usage frequency

AI usage frequency	Reported (%)	Reported (cum. %)	Predicted (%)	Predicted (cum. %)
Less than once a month	13.9	13.9	10	10
Once a month	7.6	21.5	8	18
Once a week	14.5	36.0	10	28
Several times a week	26.7	62.7	15	43
Once a day	8.0	70.7	14	57
Several times a day	22.9	93.6	10	67
Once an hour	1.4	95.0	21	88
A few times an hour	2.1	97.1	12	96
Several times an hour	2.9	100	4	100

Notes: Reported frequencies are computed from survey responses (N = 1,199). Predicted frequencies were estimated before the start of the survey to design the social comparison.

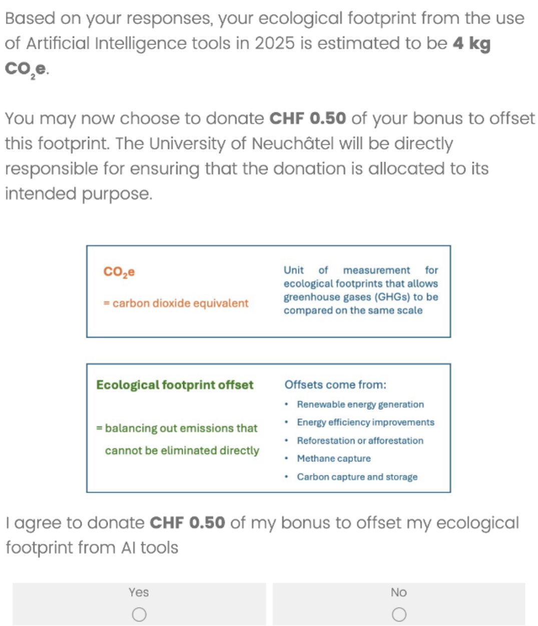


Figure 8: Example of donation decision

Table 8: Ordered logistic regressions

DV	Guilt	Pride	Shame	Empathy	Belonging
Non-donors					
norm / no-ambiguity	0.138	0.256	0.149	-0.067	0.041
	(0.244)	(0.224)	(0.254)	(0.220)	(0.221)
no-norm / ambiguity	0.030	0.075	-0.006	-0.064	-0.284
	(0.246)	(0.223)	(0.258)	(0.220)	(0.218)
norm / ambiguity	0.169	-0.044	0.262	-0.037	0.062
	(0.239)	(0.217)	(0.246)	(0.213)	(0.214)
Donors					
no-norm / no-ambiguity	0.020	0.413*	-0.061	1.137***	0.582**
	(0.233)	(0.213)	(0.245)	(0.211)	(0.206)
norm / no-ambiguity	-0.209	0.441**	-0.222	1.067***	0.558**
	(0.237)	(0.213)	(0.250)	(0.210)	(0.208)
no-norm / ambiguity	-0.160	0.881***	-0.087	1.713***	0.787***
	(0.243)	(0.215)	(0.251)	(0.215)	(0.212)
norm / ambiguity	-0.217	0.664***	-0.086	1.485***	0.621***
	(0.237)	(0.208)	(0.246)	(0.211)	(0.208)
Log likelihood	-1304.91	-1762.18	-1164.63	-1772.43	-1774.09
LR test	5.85	34.79	5.34	180.04	46.22

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Coefficients report estimates from interactions between experimental conditions and the binary decision to offset emissions, estimated using an ordered logistic regression model. Standard errors are reported in parentheses. The dependent variable is a categorical measure of participants' emotional responses following the decision. The reference category is non-donors under the no-norm / no-ambiguity condition. N=1,199.

References

Baillon, A., Cabantous, L., and Wakker, P. P. (2012). Aggregating imprecise or conflicting beliefs: An experimental investigation using modern ambiguity theories. *Journal of Risk and Uncertainty*, 44(2):115–147.

Bicchieri, C. (2023). Norm nudging and twisting preferences. *Behavioural Public Policy*, 7(4):914–923.

Cialdini, R. B. and Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, 55(1):591–621.

Dana, J., Cain, D. M., and Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100(2):193–201.

Dana, J., Weber, R. A., and Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1):67–80.

De Nardo, M., Brooks, J. S., Klinsky, S., and Wilson, C. (2017). Social signals and sustainability: ambiguity about motivations can affect status perceptions of efficiency and curtailment behaviors. *Environment Systems and Decisions*, 37(2):184–197.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4):643.

Exley, C. L. (2016). Excusing Selfishness in Charitable Giving: The Role of Risk. *The Review of Economic Studies*, 83(2):587–628.

Fahrenwaldt, A., Tho Pesch, F., Fiedler, S., and Baumert, A. (2024). What's moral wiggle room? A theory specification. *Judgment and Decision Making*, 19:e17.

Garcia, T., Massoni, S., and Villeval, M. C. (2020). Ambiguity and excuse-driven behavior in charitable giving. *European Economic Review*, 124:103412.

Grossman, Z. (2014). Strategic Ignorance and the Robustness of Social Preferences. *Management Science*, 60(11):2659–2665.

Grossman, Z. and Van Der Weele, J. J. (2017). Self-Image and Willful Ignorance in Social Decisions. *Journal of the European Economic Association*, 15(1):173–217.

Karlan, D. and List, J. A. (2007). Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment. *American Economic Review*, 97(5):1774–1793.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A Smooth Model of Decision Making under Ambiguity. *Econometrica*, 73(6):1849–1892.

Lind, J. T., Nyborg, K., and Pauls, A. (2019). Save the planet or close your eyes? Testing strategic ignorance in a charity context. *Ecological Economics*, 161:9–19.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):1281–1302.

Schultz, P. W. and Kaiser, F. G. (2012). Promoting Pro-Environmental Behavior. In Clayton, S. D., editor, *The Oxford Handbook of Environmental and Conservation Psychology*, pages 556–580. Oxford University Press, 1 edition.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., and Griskevicius, V. (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*, 18(5):429–434.

Tangney, J. P., Stuewig, J., and Mashek, D. J. (2007). Moral Emotions and Moral Behavior. *Annual Review of Psychology*, 58(1):345–372.

Tarditi, C., Hahnel, U. J. J., Jeanmonod, N., Sander, D., and Brosch, T. (2020). Affective Dilemmas: The Impact of Trait Affect and State Emotion on Sustainable Consumption Decisions in a Social Dilemma Task. *Environment and Behavior*, 52(1):33–59.

Trujillo, C. A., Estrada-Mejia, C., and Rosa, J. A. (2021). Norm-focused nudges influence pro-environmental choices and moderate post-choice emotional responses. *PLOS ONE*, 16(3):e0247519.

Van Der Weele, J. J., Kulisa, J., Kosfeld, M., and Friebel, G. (2014). Resisting Moral Wiggle Room: How Robust Is Reciprocal Behavior? *American Economic Journal: Microeconomics*, 6(3):256–264.

Van Teunenbroek, C., Bekkers, R., and Beersma, B. (2020). Look to Others Before You Leap: A Systematic Literature Review of Social Information Effects on Donation Amounts. *Nonprofit and Voluntary Sector Quarterly*, 49(1):53–73.