CER-ETH – Center of Economic Research at ETH Zurich

The Narrative of the Energy Efficiency Gap

S. Houde, T. Wekhof

Working Paper 21/359
July 2021

Economics Working Paper Series

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

The Narrative of the Energy Efficiency Gap

Sébastien Houde[1] and Tobias Wekhof[2]

July 14, 2021

Abstract

For more than forty years analysts have pointed out that society might be too slow in adopting energy efficiency technologies, a phenomenon known as the Energy Efficiency Gap. There are persistent market barriers that impede these efforts. Eliciting these barriers and their heterogeneity is key for policy design. In this paper, we use narratives, a novel approach based on unstructured text answers in surveys, to elicit the barriers and determinants of energy efficiency investments. Using recent advances in Natural Language Processing (NLP), we turn narratives into quantifiable metrics to rank households' barriers and determinants. We find that financial motives are not the primary barriers or determinants of energy efficiency investments. Instead, we find that such investments are highly opportunistic and co-benefits, such as ecological concerns and comfort, also play an important role. Although there is substantial heterogeneity across the population in the type of barriers and determinants, demographics and building characteristics poorly predict heterogeneity patterns. This has important implications for the targeting of policies. Narratives could be a novel and effective way to implement policy targeting.

JEL Classification: Q41, Q50, L15, D12, D83, D91.
Keywords: energy efficiency gap, natural language processing, policy targeting, open-ended questions.

[1]Grenoble Ecole de Management, 38000, Grenoble, France, sebastien.houde@grenoble-em.com.
[2]ETH Zurich, Center of Economic Research, Zurich, Switzerland, twekhof@ethz.ch.

We would like to thank Massimo Filippini, Nina Boogen, Nilkanth Kumar, Adrian Obrist, Elliott Ash, Anne-Kathrin Faust, Claudio Menn, and Andreas Meyer, in addition of numerous seminar participants. The Swiss Federal Office of Energy (SFOE) provided funding for this research under the grant: SI/501886-01. SFOE was not responsible for the study design, the collection, analysis and interpretation of data or in the writing of this paper. The content does not necessarily represent the official views of SFOE. All omissions and remaining errors are our responsibility.

1 Introduction

Encouraging energy efficiency investments is important for policy makers, but also challenging. In the last decade, governments across a wide span of jurisdictions have set ambitious energy efficiency targets as part of their climate change mitigation strategies. However, these targets have been systematically missed (e.g., European Commission 2019). On the one hand, this is not surprising: for more than forty years researchers and analysts have pointed out that society might be too slow in adopting energy efficiency technologies, a phenomenon known as the "Energy Efficiency Gap" (Jaffe and Stavins 1994). On the other hand, it is surprising that after forty years of implementing energy efficiency policies, experimentation, and program evaluation, we have yet to find a successful recipe to reach society's full energy efficiency potential.

The core of the debate on the Energy Efficiency Gap consists in the seeming inability to make use of the true economic potential for energy efficiency, even though the necessary technology is available. This conflict is crucial for the design of policies and relies on one important question: what are the major barriers to energy efficiency that, at the same time, can also be addressed in a cost-effective manner? Answering this question has been particularly challenging because barriers are numerous, interconnected, context specific, and change over time. Deep (whole-home) retrofitting is a case in point that illustrates the complexity of solving the Energy Efficiency Gap. Engineering estimates systematically suggest that deep retrofitting is the most cost-effective way to invest in energy efficiency. However, these calculations do not account for the possible substantial hassle costs of performing such investment, as well as the co-benefits that may occur (e.g., more comfort). These hassle costs and co-benefits are consumer- and building-specific, vary over time, and ultimately are determined by highly heterogeneous consumer preferences and building characteristics. Moreover, several of the factors that may slow down the adoption of energy efficiency technologies are not caused by a dysfunctional market, i.e., classic market failures. In contrast to

market failures, these factors, *market barriers*, are normal components of markets (Sutherland 1991). While market failures may be a rationale for policy interventions, it is less clear for the case of market barriers (Jaffe and Stavins 1994). For this reason, understanding the barriers and determinants of energy efficiency investments is key for the design and the targeting of policies.

In this paper, we analyze two related questions to help better design and target energy efficiency policies in the residential sector. What are the key barriers and determinants to investment in energy efficiency? What policies would households favor and how are these policies linked to barriers and determinants of energy efficiency investments? We analyze these questions with a focus on uncovering systematic heterogeneity, in order to implement policy targeting.

We are not the first to ask those questions. However, we propose a new way to answer them. We use narratives to elicit and rank barriers and determinants households face in the energy efficiency context.

Our approach uses an alternative paradigm in the large and still growing empirical literature on the barriers and determinants of energy efficiency investments (Cagno et al. 2013; Cattaneo 2019; Gerarden et al. 2017; Gillingham et al. 2018; Schleich et al. 2016; Sorrell et al. 2004). This literature follows two broad methodological approaches: non-choice-based and choice-based methods. Under the first approach, analysts will typically construct a survey with well-defined options of barriers and determinants that the analysts think households might face (e.g., Hrovatin and Zorić 2018; Lee 2015; Trotta 2018). With the second method, barriers and determinants are inferred from choices. Hypothetical choice situations can be constructed to have a perfectly controlled environment, which allows to estimate choice models and infer underlying preferences (e.g., Alberini et al. 2013; Banfi et al. 2008; Blasch et al. 2019; Fischbacher et al. 2021; Schleich et al. 2019). Observational or experimental data that are naturally-occurring can also be collected (e.g., Allcott and

Greenstone 2017). These data, together with a model that provides the micro-foundations for the mapping between preferences and observed choices are used to infer preferences, and incidentally particular barriers and determinants. A common problem that arises with non-choice as well as choice-based approaches is that preferences can be highly heterogeneous. For non-choice-based elicitation procedures, this means that the analysts must take a strong stance *ex ante* on which barriers and determinants to focus on. This is prone to induce an elicitation bias. For choice-based methods, highly heterogeneous preferences translate into an identification problem—it might be too difficult to find the right model that explains the data.

To overcome these challenges, we propose an approach that consists to use narratives to elicit preferences. Collecting narratives is deceptively simple and can be exceptionally powerful to understand a wide range of economic phenomena (Shiller 2020). It consists to ask people what they think and let people speak (or write) their mind. In practice, this approach should yield very noisy qualitative data that are hard to interpret. However, recent advances in Natural Language Processing (NLP) allows us to turn narratives into quantifiable metrics to elicit proxy for household preferences. We apply this idea to explain the drivers of important economic decisions and policy-preferences in the energy efficiency context.

The use of open-ended questions has been a subject in social sciences since the 1940s, but has not seen systematic applications on a large scale (Krosnick 1999). Only recently, in the context of advances in NLP, Roberts et al. (2014) suggest to revisit the concept of ope-ended survey questions. In the line of this newly emerging literature, Egami et al. (2018) use open ended questions in an experimental setting to elicit opinions on immigration and Stantcheva (2020) uses open ended survey questions to elicit policy preferences.

We find that narratives are a powerful tool that should be added to the arsenal of a researcher and complement existing choice- and non-choice-based approaches. In our context,

narratives are particularly illuminating to elicit barriers and determinants while capturing heterogeneity across different types of households. They offer two clear advantages over typical closed-ended survey questions. At the respondent-level, narratives elicit a narrower but more precise set of barriers and determinants. Specifically, respondents tend to focus on a few but more important topics explaining their decision-making. Nevertheless, population-wide, a broader set of barriers and determinants are uncovered through narratives compared to closed-ended questions.

Our results suggest that financially-related barriers and determinants are important, but may not be necessarily the dominating ones. The narratives we obtained from the open ended questions tell us that energy efficiency investments are highly opportunistic: households that do not invest in energy efficiency often believe that their house is already too energy-efficient; households that invest do so when a particular building technology becomes obsolete. Heterogeneous vintage of buildings and technologies are thus important. However, this could be a normal market barrier that policymakers may have little leverage to influence.

The role of co-benefits, namely comfort and ecological footprint, is important. It is also the most malleable with respect to the elicitation procedure. Using closed-ended questions, these two co-benefits dominate. In contrast, narratives suggest that those two co-benefits are in fact a much less prevalent determinant of investment. This inconsistency suggests that respondents' top-of-mind awareness combined with the effort to express decision-making in an open-ended question render this elicitation procedure narrower, but also more precise. From a policy targeting standpoint, it has important implications.

With respect to other categories of barriers and determinants, we find some but not overwhelmingly strong support for behavioral barriers, although this explanation has received a lot of attention in the literature recently (Gillingham et al. 2018; Schleich et al. 2016). Our

survey instrument was especially designed to capture some behavioral dimensions of decision-making such as energy-related financial literacy, which is correlated with some determinants.

Finally, we find that narratives are also useful to elicit preferences for policy interventions. There is also a consistent mapping between policy preferences and determinants of energy efficiency investments.

There exists a rich literature investigating the barriers at the source of the Energy Efficiency Gap. However, that literature also shows a lack of consensus about the barriers' relative importance, which is what is ultimately important to guide policymakers. One reason for this lack of consensus is that the methodological paradigms used by empiricists constrain to study one or a few explanations of the Gap in particular contexts. Energy efficiency investments depend on a combination of several factors: local institutions, behaviors, and technologies. Generalizing one explanation to a broader context is not only difficult, but often not possible. Narratives offer a way to efficiently elicit and rank barriers and determinants of energy efficiency. Our proposed method can be easily replicated across markets and used by policymakers in the policy-making process.

The remaining parts of this paper are organized as follows: in Section 2, we revisit the taxonomy used to classify the barriers and determinants of energy efficiency investments and that is at the source of the Energy Efficiency Gap. Next, we present the data and our empirical context in Section 3, and then show in Section 4 the main results about the elicitation of barriers, determinants and policy preferences. We build on these results in Section 5, where we focus on the heterogeneity of barriers/determinants and how to use this heterogeneity for policy targeting, followed by a concluding section.

2 A Taxonomy of Barriers and Determinants

Several frameworks have been proposed to identify and categorize the barriers and determinants of energy efficiency investments. One particularly influential taxonomy has been

proposed by Sorrell et al. (2004), who distinguished between three different perspectives: economic, behavioral, and organizational.

The economic perspective considers rational utility-maximizing agents as the benchmark to understand why or why not agents adopt energy efficiency technologies. The behavioral perspective departs from this purely neo-classical framework and considers different manifestations of bounded rationality, which have also been referred to as behavioral failures in the literature (Gillingham et al. 2009). Finally, the organizational perspective considers the role of institutions agents interact with. Those could be institutions for which governments have very little ability to transform, such as values and culture, or others, that they have considerable influence over, such as fiscal, competition, and regulatory policies.

While this taxonomy has been proven useful to navigate the different explanations at the source of the Energy Efficiency Gap, a more precise categorization is needed for the purpose of policy design. Our taxonomy is motivated by a utility-based model of investment where four types of barriers interact: market, non-market, financial, and behavioral barriers.[1]

Consider the case of a household that decides to invest or not in an energy-saving technology. To fix ideas, we can consider the case of a whole-home (deep) retrofit, where e_{jt} is the quantity of energy consumed at time t if there is an investment, denoted by $j = 1$, or no investment, denoted by $j = 0$. A deep retrofit influences energy usage, but it is also associated with other benefits and so-called "hidden costs". We denote co-benefits (good) by g_j and hidden costs (bad) by b_j. Examples of co-benefits are improvements in indoor air quality, thermal comfort, or aesthetics, to name a few. Hidden costs could be increases in technology complexity (and hence maintenance), or other inconveniences that are caused by a deep retrofit. We consider that g_j and b_j have mostly non-market values, but they are nonetheless important for households. Households trade off the capital cost of the investment, denoted c_j where $c_0 = 0$ (no investment case), with the discounted sum

[1] In our framework, organizational barriers are not explicitly modeled but their impact on households' decisions can be accounted for through "non-market" barriers and determinants that manifest as hassle costs.

of the energy savings: $\sum_t \rho^t \cdot p_t^e \cdot (e_{1t} - e_{0t})$, which is function of the price of energy, p_t^e and the discount factor ρ. Other than energy, households consume a numeraire good, y. Income is function of an hourly salary, which varies as a function of how much leisure time and non-work related activities are allocated by the consumer. We denote t the total time endowment available to do hourly work and l the time allocated to other activities. A deep retrofit takes time to plan and implement. Those are the hassle costs that can be modelled as an increase in l if $j = 1$.[2] Finally, households have other sources of income, denoted I and preferences are quasi-linear with respect to the numeraire good y.

Household's utility for each investment option, $j = \{0, 1\}$ is thus given by

$$V_{jt} = U(e_{jt}, g_j, b_j|\theta) + y \tag{1}$$

where utility is maximized subject to the budget constraint:

$$c_j + \sum_t \rho^t \cdot p_t^e \cdot e_{jt} + y \leq w \cdot (t - l_j) + I \tag{2}$$

The function $U(\cdot)$ is a utility function that varies with the vector of parameters θ. Together, $U(\cdot)$ and θ characterize households' preferences. Finally, the discount factor, ρ can also be considered a household-specific preference.

The decision to invest is thus determined by the following inequality:

$$U(e_{1t}, g_1, b_1|\theta) - U(e_{0t}, g_0, b_0|\theta) \geq c_1 + \sum_t \rho^t \cdot p_t^e \cdot (e_{1t} - e_{0t}) + w \cdot (l_1 - l_0) \tag{3}$$

The expression on the left-hand-side (LHS) corresponds to the net non-market benefits of investing in energy efficiency, which may include comfort, air quality, noise, and other amenities that a deep retrofit may impact. On the right-hand-side (RHS), we have the different

[2]To simplify the notation, we include l only in the budget constraint but not in the utility function. An increase in l takes away wage-related income and thus always decreases utility.

components of the costs. First, we have the capital cost, c_1, the discounted energy savings, $\sum_t \rho^t \cdot p_t^e \cdot (e_{1t} - e_{0t})$, and the hassle costs, $w \cdot (l_1 - l_0)$. A household will invest when the LHS is larger than the RHS.

This framework can used to summarize the important strands of research on the Energy Efficiency Gap and the different economic and behavioral barriers that have been investigated.[3] Based on Sorrell et al. (2004)'s economic, behavioral, and organizational perspectives, our taxonomy distinguishes between market, non-market, financial and behavioral barriers.

Behavioral barriers

For the most part, the empirical research has focused on determining whether households correctly perceive the energy savings component of the net costs of investment. Allcott and Greenstone (2012) use the term γ to scale $\sum_t \rho^t \cdot p_t^e \cdot (e_{1t} - e_{0t})$, where γ corresponds to any type of investment inefficiencies. Some of these inefficiencies may be behavioral, such as inattention and biased beliefs about energy prices, to name a few. These behavioral barriers can, however, be confounded with neo-classic market barriers such as different access to credit or simply time discounting preferences. The γ parameter thus encompasses behavioral and market barriers that are internal to household decision-making (Cagno et al. 2013; Schleich et al. 2016).

Financial barriers

A second set of economic barriers focuses on the role of external factors that are financially-related. The price of energy might be too low, the costs of investment too high, subsidies might not be generous enough, and/or there could exist various financial distortions. In our framework, the role of financially-related barriers operate through the price variables

[3] As discussed above, organizational barriers and determinants can be partly accounted via their impact on hassle costs.

and interact with the marginal utility of money (i.e., sensitivity to price) that is embedded in the vector preference parameters: θ.

Non-market barriers

A third subcategory of economic barriers consists of the non-market components of the investment, the LHS in Equation 3, which includes the various co-benefits and hidden costs of such investments. The literature has pointed out that specific co-benefits could be important and contingent valuation methods have been primarily used to investigate those (Jakob 2006; Ürge-Vorsatz et al. 2014). External organizational constraints factors could impact the hassle costs on the RHS, which we also consider as a form of non-market barrier.

Market barriers

Finally, there are classic market barriers that are normal components of well functioning markets. These barriers could arise because of heterogeneity in building stock, technologies, and/or preferences. For policy makers, understanding this heterogeneity could allow them to target and tag energy efficiency policies to increase their cost-effectiveness (Allcott et al. 2015). This point has been long recognized and several studies have looked at different dimensions of heterogeneity in the decision to adopt energy efficiency technologies (e.g., Jakob et al. 2007).

As we can see, there is a wide range of explanations to the Gap. Table 1 summarizes the four important categories of barriers we consider in this paper.

Table 1: Proposed Taxonomy of Barriers to Energy Efficiency Investments

Type of Barrier	Description	Examples
Market	Normal components of markets that impact decisions	Heterogeneity in building stock, heterogeneity in building/household expected lifetime
Non-market	Non-market goods or bads that impact decisions	Comfort, hassle costs (including costs due to bureaucracy)
Financial	Variables related to prices and costs	Price of energy, interest rate, investment cost
Behavioral	Elements of households' decision-making	Bounded rationality, energy and investment literacy, myopia

3 Data and Environment

To elicit narratives related to energy efficiency investments, we conducted a large survey with homeowners of single-family houses living in the canton of Zurich, Switzerland. The survey was specifically designed to investigate the decision to invest or not in energy efficiency technologies and services.

The survey had several modules. The first module collected information on past and future energy efficiency-related behaviours: whether households performed or intended to perform retrofits, and the type of retrofits. We use these different behaviors to distinguish takers and non-takers of energy efficiency investments. The goal of the remaining modules was then to determine the components of households' decisions that influence these behaviors. One of the most important modules focused on the different types of barriers and determinants. To elicit these components, we used open-ended questions, which provide narratives about specific aspects of the decision-making process. We also used structured closed-ended questions that closely mirror the open-ended questions. Our goal is to provide a benchmark to open-ended questions. Another module focused on preferences for different types of energy efficiency policies. Finally, the other modules elicited household and building

characteristics, including some related to the decision-making process such as financial and energy-related literacy. We use these variables to investigate heterogeneity along several dimensions.

3.1 Implementation

To recruit participants, we collaborated with the Statistical Office of the Swiss Canton of Zurich. We sent personalized invitation letters by mail to a random sample of homeowners. The letter contained a short description of our research project and a link to an online survey. Households had to type the link in a web browser to complete the survey.

We stratified the sample according to the following rules: only single-family homes, year of construction prior to 1990, 50% with renovation permits during the last 5 years; large buckets for age and household size. We also stratified to target homeowners that adopted the main certification for energy-efficient buildings in Switzerland: the Minergie certification.

In the Canton of Zurich, there exist a total of 127,950 single family homes, 10,737 out of which applied for a renovation permit during the past 5 years. The Statistical Office of the Canton Zurich sampled this population and sent out 16,700 letters on Feb. 3, 2020 on our behalf. Households could complete the online survey until March 13, 2020.

Of the 16,700 letters sent on Feb. 3, 2020, the response rate was high: 3,471 respondents started the survey, which translates to a response rate of 20,8%. Furthermore, we reached a completion rate of 82% with an average time spent of 30 minutes.

Although our sampling strategy targeted a population of homeowners of single-family houses, a small number of respondents did not fall into this category. We had 161 tenants and also a small number of respondents living in an apartment (n=23). Those observations are excluded from our analysis.

3.2 Sample Composition: Classifying Household Types

We use past and intended energy efficiency-related behaviors to classify households into two broad segments. First, we distinguish homeowners who adopted the Swiss energy efficiency certification for buildings: Minergie. Our stratified sampling strategy ensured that we observe a large number of those households (n=524). We use these households for a separate study and therefore omit these observations for the main analysis. Second, we distinguish households depending if they performed an energy efficiency retrofit in the past 5 years or plan to do so within the next 5 years. Based on this criterion, households fall into two mutually exclusive categories:

- **Non-Takers**: households who did not perform energy efficiency investments in the past and who are not planning to do one in the future (483 observation, 22% of the sample).
- **Takers**: households who either did perform energy efficiency investments in the past five years or who are planning to do at least one in the next five years (1748 observations, 78% of the sample).

Table 2 shows the different household types and how they differ with respect to key building characteristics, demographics, and psychographics. A detailed description of the variables can be found in the Appendix. With respect to building characteristics, takers and non-takers of retrofits do not differ in building age and floor size. However, non-takers report a higher estimated rental value for their home. Moreover, non-takers have a slightly higher proportion of oil and gas heating, which is, however, the most common form of heating for both groups. Takers, on the other hand, have a higher proportion of heat pumps.
Both groups have roughly the same level of income. However, non-takers tend to be older, have a higher proportion of pensioners and are less likely to be living with children. Regarding the psychographics, most variables do not differ, except for the share of respondents who

took economics classes during their education, which is higher for takers.[4]

Table 2: Summary Statistics by Household Types

Variable	All	Non-Takers	Takers
Building Characteristics			
Building Age (years)	56.35	57.33	56.08
Floor Size (m^2)	168.52	167.54	168.78
# Rooms	5.74	5.60	5.78
% Garden	98.00	97.80	98.10
Rental Value (CHF/month)	3831.23	4050.27	3772.20
% Oil/Gas heating	49.60	52.50	48.80
% Heat Pump	32.50	26.00	34.20
% other heating	18.00	21.50	17.00
% Solar PV	18.60	13.30	20.00
Demographics			
Income	12509.89	12479.80	12517.97
Age	58.81	61.37	58.13
% Male	77.90	79.50	77.50
% Children	48.00	42.40	49.50
% University Degree	60.20	61.90	59.70
% University Degree Spouse	39.20	37.00	39.70
% Employed (fulltime)	33.80	31.10	34.50
Pensioner	34.30	38.30	33.20
Other employment	32.00	30.60	32.30
Spouse: Employed (fulltime)	16.10	15.70	16.20
Spouse: Pensioner	27.70	32.30	26.60
Spouse: other employment	56.20	51.90	57.20
Allergies	19.90	21.10	19.60
Psychographics			
Energy Literacy	3.76	3.72	3.76
% Took Econ	46.70	43.50	47.60
% Math Proficient	45.60	47.30	45.20
Energy Saving Score (/3)	2.29	2.25	2.31
% Donated Enviroment	55.00	53.40	55.40
Happiness Score (/4)	2.11	2.17	2.10

Note: This table presents the summary statistics for the entire survey sample, with exception of respondents who live in a "Minergie" certified building. The sample contains a total of 2231 observations, out of which 483 (22%) are Non-Taker and 1748 (78%) are Takers. Takers are defined as respondents who performed an energy efficiency retrofit either in the past five years or plan to do so within the next five years.

[4]In the subsequent analysis, we do not use all the variables to avoid limiting the sample size too much. Specifically, we do not use the variables for the number of rooms, garden, heating mode and solar PV. Moreover we also do not include information about respondents' spouses because this would not allow us to analyze single-person households. We also reduced the employment measure to the dummy that takes the value of one if the respondent is a pensioner and zero otherwise.

3.3 Eliciting Energy Efficiency Narratives

We used different open-ended questions to infer respondents' reasoning concerning their decision-making process pertaining to an energy efficiency retrofit. We separately elicited barriers from the non-takers and determinants from takers. The structure of the question was similar for both barriers and determinants.

When asking open questions, it is important to provide some context to the participants and indicate why we ask such questions. We thus structured our survey by first presenting the following short introduction explaining the rationale for asking open-ended questions:
"The reasons for energy efficiency retrofits are complex and different for each household. We would like to learn more about why you decided (not) to renovate. What was important to you? Were there alternatives? Your response will help us to better understand how we can support energy efficiency retrofits.".[5]
After providing context, we then asked the following question:

"Describe the reasons why you decided (not) to carry out energy efficiency retrofits. Please write a short text of about 4 sentences."

A key element of our survey design to assess the validity of our elicitation procedure for narratives is that we also asked closed-ended questions at an earlier point in the survey. These questions took the form of multiple choice questions that mirror the topics we expected respondents would state in their answers to the open-ended questions. For barriers, we listed seventeen potential barriers that have been discussed in the literature of the Energy Efficiency Gap. Non-takers were presented with those options and had to select the barriers that were important for them. For the determinants, we established a list of eight potential determinants. Takers had to select which ones were important to them.

[5]The survey was conducted in German. We present in the paper our own translations of the original German questions and answers.

In addition to barriers and determinants, we also extracted narratives about policy preferences at the end of the survey. To do so, we proceeded in a similar fashion. We first presented a short introduction:

"The building sector has one of the greatest potentials for energy savings in Switzerland. One of the goals of our project is to improve public programs for energy-efficient building and renovation."

This introduction was then followed by the open-ended question:

"We would now like to ask for your opinion. What approaches do you think the public sector should promote to encourage energy-efficient construction and renovation for households living in Switzerland?"

Overall, the implementation of the open-ended questions worked very well. By inspecting a large number of responses, we found that respondents provided meaningful answers. The length of the answers to the three open-ended questions varies between 19 to 24 words, on average. The standard deviation is about the size of these averages, and some respondents wrote very long and detailed answers.[6]

The questions on barriers and determinants were mandatory to all non-Minergie participants. As mentioned before, only non-takers were asked about the barriers, and only takers were asked about the determinants. An overview of the summary statistics for the open-ended questions is given in Table 3. For these two open-ended questions, we observe an attrition rate of only 1.5%. This means that upon having to answer one of this particular questions, only 1.5% decided to stop the survey altogether. The question on policy recommendations was not mandatory and was placed at the end of the survey. Furthermore,

[6]The median number of words is between 12 and 21, and the 90% percentile is between 43 and 47, depending on the questions.

all respondents were presented this question. We observe a higher attrition rate but still relatively low: 8.5%. Self selection of respondents is thus not a major concern.

For each question, we extracted the entire text corpus and counted the number of unique words. The question related to determinants of energy efficiency investments has the largest number of unique words, and also the most words that are suitable for topic extraction. The words that we have identified for topic extraction serve as the basis of our main analysis, which we discuss next.

Table 3: Summary Statistics - Open Text Answers

	Barriers	Determinants	Policy Recommendations
# answers	463	1758	2482
% attrition	1.5	1.5	8.5
mean # words	24	21	19
median # words	21	17	12
90 percentile # words	47	44	43
sd # words	17	18	24
total # unique words	1671	3620	2371
total # words used for topics	240	667	492

Note: The questions on barriers and determinants were mandatory to all non-Minergie participants. The question on policy recommendations was open to all respondents but not mandatory to complete the survey. We calculated the attrition for each open-ended question by comparing the number of respondents to this particular question versus the response rate to the last mandatory question that preceded it. For barriers and determinants, all nouns, verbs, adverbs and adjectives with at least four characters were used. For the policy questions, we only selected words that occurred at least twice in the corpus.

4 Uncovering Barriers, Determinants and Policy Preferences

Our analysis proceeds in three steps. First, we extract topics from the open questions using NLP tools. Second, we rank topics based on their relative frequency. Third, we compare rankings obtained with open- versus closed-ended questions.

To elicit topics from narratives in open-ended questions, we first extracted all the words from the respondents' answers and classified them as nouns, adjectives, verbs and adverbs, using the spacy library for python (Honnibal et al. 2020). In a second step, we used word embeddings to map each word to a distance metric. Word embeddings are matrices that

have a column of values for each word that indicate the relative semantic distance between words (e.g., the distance between "heating" and "oil" is smaller than the distance between "heating" and "pencil"). For a pair of words, it is thus possible to calculate their relative distance using the cosine similarity. To construct such a matrix for our corpus of unique words, we mapped all the words present in the answers to the pre-defined German fasttext word embedding vectors (Grave et al. 2018). In a third step, we clustered nouns, adjectives, verbs and adverbs separately using k-means clustering and the measure of relative semantic distance. The reason we performed a semantic clustering was to ease the subsequent topic extraction (we chose the number of clusters such that each cluster contained between 10 and 30 words). In the fourth and final step we extracted topics. We assigned each word, when possible, to one of the existing topics from the corresponding closed-ended question (each word can only belong to a single topic). This step was not automated and was performed manually.[7] During the topic extraction process, we also discovered additional topics, which we then added to the list of predefined topics of the closed-ended question. Of all the unique words, we could assign between 15-20% of the words to a topic. Finally, after assigning words to topics, we labelled the text answers by automatically searching each text answer for the presence of the set of words that define a topic.

Consider for example the following answer to the open-ended question on determinants: *"Oil heating has reached the end of its life-cycle and the introduction of environmentally friendly energy was important for us."* This sentence contains the keyword "life-cycle", which was previously selected as an indicator for the topic "replace broken elements". The keyword

[7]There are various algorithms that can be used to automatically extract topics from text in an unsupervised way, such as the Latent Dirichlet Allocation (LDA) from MALLET (McCallum 2002), Structural Topic Model (Roberts et al. 2019), LDA2Vec (Moody 2016) or Top2Vec (Angelov 2020). We experimented with these various methods and found that they tend to deliver topics that are difficult to interpret with the text answers we obtained for our open-ended questions. One of the shortfalls of those automated, unsupervised, clustering algorithms, is that they cluster the text answers without considering the question we asked the respondents. Moreover, the clusters obtained from those algorithms are sensitive to various model parameters. In comparison, supervised text classification relies on training the algorithm on a training set that usually was classified by humans. The algorithm then "learns" to replicate the human classification. Our approach is therefore more closely related to supervised NLP methods where the initial classification criteria are defined by the researcher.

"environmentally" clusters this answer in addition to the topic "reduce ecological footprint". The same idea was applied to the question with barriers, but with different keywords and topics. To give an example of an answer to the barrier-question: *"Other renovations were more important and urgent. Everything at the same time would not be financially affordable.* This answer contains the two keywords "affordable" and "financially" both of which were selected to classify that answer in the topic "too expensive". We proceeded in the same way for the policy question, e.g. with the following answer: *"good, neutral consulting, subsidies for energy-saving renovations"*. In this example, the respondent used the keyword "consulting", which is classified in the topic "more information" and the keyword "subsidies", which belongs to the subsidy-topic.

Using the approach outlined above, we ranked the barriers and determinants to energy efficiency investments by tabulating topic frequency. We then contrasted the rankings with the ones obtained from the closed-ended questions.

4.1 Barriers

Table 4 presents the results for the barriers. From both the open- and closed-ended questions (Column 1 and Column 2, respectively), the most important barrier for non-takers is the statement that their home is already energy-efficient. Whether this is a belief or a fact about the house respondents live in, we cannot know for sure. Information about building characteristics in Table 2, however, provides us some indications that beliefs, in particular biased beliefs, might be partly at play. We found that two important building characteristics that determine the energy efficiency potential of a house, namely vintage and the type of heating system, are not drastically different between takers and non-takers. The fact that households consider that their house is already energy-efficient is a statement about personal preferences and thus can be considered a normal component of markets (i.e., a market barrier). On the other hand, one could argue that there is also a behavioral component to

this barrier. For instance, if it was the case that biased beliefs were important, information campaigns and subsidized audits could be justified to address those.

Regarding the consistency between the closed- an open-ended questions, we observe that this barrier has a higher frequency when we focus on narratives: 47% of respondents wrote about this, but only 37% selected this option as a potential barrier in the closed-ended question. It remains, nonetheless, the most important barrier for both open- and closed-ended questions.

The cost of a retrofit and the fact that it might be too expensive is the second most important barrier. From the narratives, we observe a slightly higher percentage of people that mention this barrier. Although, financial barriers have been the centered-piece of energy efficiency programs and result in offering generous subsidies, less than a quarter of non-takers mention this as an important barrier. This finding challenges the notion that energy efficiency policies should target primarily this financial type of barrier.

In the narratives, the old age of the respondents is the third most important barrier. This is not a topic that we a priori listed in the closed-ended question. In hindsight, we acknowledge that this can be an important barrier. The fact that households anticipate that the remaining period they will reside in their house is too short to rationalize a long-term investment is also a normal component of markets. That is, heterogeneous life expectancy is a market barrier.

We also observe that several barriers were not mentioned in the answers to the open-ended question, but were selected with a certain frequency with the closed-ended question. There are several potential explanations for this. When provided with a pre-defined list of options, it is almost costless for respondents to select an additional option. A greater diversity of barriers thus emerges from the closed-ended question, but cheap talk might be at play. To the contrary, for an open-question, writing about an additional barrier requires much more efforts. Open-ended questions might then induce more truth-telling in eliciting

the most important barrier(s) each non-taker has faced. For instance, aesthetics and the difficulties associated with renovating old buildings are two barriers that emerge from the closed-ended question with a certain importance, but there is little mentioning of them in the narratives.

Table 4: Barriers to Energy Efficiency Retrofits

	Type of barrier	Open	Closed
The building is already energy-efficient	Market	47.2	37.3
Too expensive	Financial	24.8	21.1
Old age	Market	7.2	0
Too complicated	Non-Market	6.4	9.7
Aesthetics	Market	2.5	7.9
Difficulties due to historic building	Market	2.5	6
Expert recommended against	Behavioral	1.7	1.7
Other priorities	Non-Market	1	0
Difficulties in applying for permits	Non-Market	0.8	3.1
Difficulties in obtaining financing	Financial	0.6	3.9
Craftsman recommended against	Behavioral	0.6	2.3
Hassle	Non-Market	0.4	0
Architect recommended against	Behavioral	0.2	1.7
I did not think of it	Behavioral	0	4.3
Planning to move	Market	0	6.2
Lack of information	Behavioral	0	8.3
The investment too risky	Behavioral	0	2.1
Leaving the house during the renovation	Non-Market	0	5.6
It is difficult to find experts or materials	Non-Market	0	4.3
Bad experiences with previous renovations	Non-Market	0	1.9

Note: Households who did not undertake a retrofit in the past 5 years and did not plan to do so in the next 5 years were asked to choose among several option the reasons why they decided against a retrofit. Later in the survey, we asked the same respondents in an open-ended question for the reasons why they did not do a retrofit. We then classified the text answers into the same categories as the closed answers and added several new topics, such as "old age".

4.2 Determinants

Turning to the determinants of energy efficiency renovations on Table 5, we first observe that the elicitation procedure plays an even more important role. The ranking of the topics from narratives is very different than the one obtained using the closed-ended question.

From the narratives, we learn that the main determinant of an energy efficiency investment is that a particular building technology came to its end of life. Unless obsolescence is strategically manipulated by firms, it is a normal component of the market. Therefore, this

is not a rationale for a policy intervention. The importance of obsolesce as a determinant mirrors the fact that the main barrier to investment is that households consider that fewer energy efficiency improvements can be done. Taken together, this suggests that investments in energy efficiency technologies are opportunistic in nature.

The answers to the closed-ended question suggests that the most important determinants are comfort and reduction of the ecological footprint. Those determinants are also mentioned in the answers to the open-ended question, but with a much lower frequency. Nonetheless, these non-market benefits associated with energy efficiency investments are among the main determinants that were extracted from the narratives.

There is also a large discrepancy between the answers for open- and closed-ended questions regarding the role of financial-related determinants. One quarter of respondents mentioned this determinant in the narratives, but 36.7% selected this option in the closed-ended question. Moreover, the impact of energy efficiency on resale value received little attention in the open text answer, but it was selected by as much as 25.2% of respondents in the closed-ended question.

Overall, the results tell us a clear message with respect to the most important reasons for making an energy efficiency investment: non-market benefits, financial considerations, and obsolescence are all important. The respective importance of each of those determinants is, however, malleable and depends on the elicitation procedure. Again, this has important implications for the targeting of energy efficiency policies. For instance, if we were to leverage the fact that co-benefits, such as comfort and ecological motives, are important in households' decisions, it would be difficult to systematically target those determinants using a single elicitation procedure.

Table 5: Determinants of Energy Efficiency Retrofits

	Type of Determinant	Open	Closed
Replace broken elements	Market	43.9	57.7
Save money	Financial	25.6	36.7
Increase comfort	Non-Market	23.3	68.4
To reduce my ecological footprint	Non-Market	21.7	68.9
Increase resale value	Financial	4.7	25.2
Regulatory	Non-Market	2.9	0
Increase size of home	Market	2.6	0
Recommended by another expert	Behavioral	1.8	6.2
Aesthetics	Non-Market	1.6	0
Safety	Non-Market	1	0
Recommended by an architect	Behavioral	0.4	4.9
Recommended by a craftsman	Behavioral	0.3	4.1

Note: Households who performed a retrofit in the past 5 years or not plan to do so in the next 5 years were asked to choose among several option the reasons why they decided to perform a retrofit. Later in the survey, we asked the same respondents in an open way for the reason why they decided to do a retrofit. We then classified the text answers into the same categories as the closed answers and added several new topics, such as "Regulatory".

4.3 Policy Preferences

In this section, we investigate the mapping between policy-preferences and the most important barriers and determinants of households' energy efficiency investment decisions. Uncovering policy preferences, especially the consistency between such preferences and the most important barriers and determinants, is key to understand how policies should be designed and targeted.

A first set of explanations that we explore is thus the role of past experience with policy measures and how general awareness of the policy landscape might shape preferences.

In our survey, we asked respondents about their awareness of different types of energy efficiency policies [8]. Respondents could choose one of the four following, mutually exclusive, options for each of the policies: "I was not aware of the option", "I am aware of the option", "I have used the option" or "I intend to use the to use the option". Using these answers, we can then construct indices of policy awareness and policy usage. For policy awareness,

[8]The four policies are: discount on mortgage interest rate, tax exemptions or deductions, various subsidies from cantons and municipalities and the so-called "Building-Program" a nationwide subsidy-scheme.

the index is constructed by creating a dummy variable for each policy that takes the value of one if the respondent did not answer "I was not aware of the option". In a second step, the dummies for all policies are added, which means that a respondent can have a maximum score of 4. Hence, the awareness measure gives an indication if a respondent is informed about a policy or made use of it. For policy usage, we proceed similarly, except that the dummy variable for each policy takes the value of one if a respondent answered either "I have used the option" or "I intend to use the to use the option". As for the awareness measure, the policy usage score is constructed by the sum of the dummy variables for all four policies. For awareness, non-takers have a slightly lower mean with 2.69 compared to 2.87 for takers. That difference however is statistically significant with a t-test. Policy use has a big difference between the two groups, with 0.81 for the non-takers compared to 1.47 for the takers. That difference is also significant with a t-test.

In a second step, we rely on narratives to elicit policy preferences.[9] As for the other open questions, we performed the topic extraction by clustering nouns, adjectives, verbs and adverbs using word embeddings and k-means clustering. However, in difference to the open questions on renovation, we only selected words that occurred at least twice in the text corpus. We did this slight change in procedure because we did not have a pre-defined list of topics from a closed-ended question. Moreover, this step considerably facilitates topic extraction.[10]

[9]For the narratives pertaining to policy preferences, we did not use a closed-ended question. Our focus here is to simply extract important topics from the narratives and heterogeneity across the different types of households.

[10]Selecting words with a frequency equal or higher than two significantly reduced the number of words and thus facilitated the topic clustering. Rarely occurring words are mainly important for very precise and small topics. Furthermore, because we did not compare the open policy question to a closed question, this level of precision is not necessary. Working with a corpus with a lower dimensionality also facilitates the initial discovery and definition of topics.

Table 6 presents the results for the open-ended question on policy preferences. A wide range of topics emerged from the narratives. When asked how policies could encourage energy efficiency investments, more generous subsidies came as the top suggestion. This is true for all types of households. A greater focus on solar photovoltaic (PV) technology is the second most popular suggestion. It is interesting to note that in Switzerland, energy efficiency programs and incentives for solar PV are usually not combined. Households, however, would like to have more integration between those measures.

The remaining suggestions are about providing more information, reducing bureaucracy, and favoring standards. Other topics, with smaller shares, also emerged from the narratives. Tax-related measures are discussed but are not a popular topic, especially in comparison to subsidies. Although "subsidy" is the most popular topic, almost 65% of respondents favor other policy measures.

As with barriers and determinants, we classified all policy options proposed by the respondents in three broad categories. The first category consists of market-based instruments. These are policy options related to subsidies and tax. The second category consists of behavioral instruments. In this category, we include instruments that are motivated by behavioral biases. Information provision and standards are the most notable examples in this category. Finally, the third category consists of non-market-based policy instruments, which encompass other types of interventions such as reducing bureaucracy.

Table 6: Policy Preferences from Open Text Answer

	All	Non-Takers	Takers
More subsidy (market)	32.3	36.6	31.1
Focus PV	16.3	17.8	15.8
More information (behavioral)	16.1	17.8	15.7
Less bureaucracy (non-market)	15.8	18.0	15.2
Focus Heating	12.9	13.5	12.8
Standards (behavioral)	9.3	9.5	9.2
Tax deduction (market)	8.7	9.7	8.4
Pollution tax (market)	6.5	7.2	6.3
Focus on new buildings	4.8	4.1	4.9
Focus Insulation	3.6	2.9	3.8
Technology	3.0	2.9	3.1
Property tax (market)	1.5	2.1	1.3
Subsidy threshold	1.1	1.0	1.1
Credit (market)	0.5	0.2	0.6

Note: This table presents policy preferences that were obtained by classifying an open text answer. The classification was done using keywords that are unique to each topic and a response can be part of multiple topics.

5 Heterogeneity and Targeting

Several analysts have pointed out that one way to increase the cost-effectiveness of energy efficiency policies is by implementing policy targeting and tagging (Allcott et al. 2015). In practice, this requires finding dimensions of heterogeneity that are correlated with important barriers and determinants of energy efficiency investments. In this section, we thus take advantage of our rich survey data to uncover heterogeneity patterns.

To analyze how policies can be targeted, we first distinguish between takers and non-takers. In a second step, we analyze the correlation between the main barriers/determinants and the various observables, such as demographics and building characteristics, psychographics, and policy-related variables.

5.1 Heterogeneity: Takers and non-Takers

In Table 7, we present the heterogeneity between takers and non-takers using a linear probability model. The dependent variable is binary and takes the value of one if the respondent undertook a retrofit in the past five years or plans to do so in within the next five years.

Results indicate no statistically significant difference in income, gender and living with children between households. Takers of retrofits are younger than non-takers and tend to hold less frequently a university degree. Housing characteristics such as floor size, house age and rental value are not statistically different between the two groups. For the psychographics, takers tend to have included more frequently economics classes in their education. Apart from that variable, there are no statistically significant differences in psychographics between the two groups. Ceteris paribus, takers have a slightly lower policy awareness compared to non-takers, but a higher policy usage. There are no statistically significant differences for policy preferences between takers and non-takers. These results are in line with the intuition given in Table 2, that the differences in observables between takers and non-takers are mainly due to classical market barriers.

Table 7: Linear Probability Model: Takers vs. Non-Takers

	Dependent variable: *Takers* = 1 / *Non- Takers* = 0	
	Coefficient	s.e.
Building Characteristics		
Building Age	−0.0005	(0.001)
Floor Size	−0.00001	(0.0002)
Log Rental Value	−0.052**	(0.025)
Demographics		
Log Income	0.045	(0.039)
Age	−0.006***	(0.001)
Male	−0.005	(0.027)
Children	−0.009	(0.028)
University Degree	−0.054**	(0.025)
Pensioner	0.086**	(0.037)
Psychographics		
Energy Literacy	0.011	(0.020)
Took Econ	0.058***	(0.022)
Math Proficient	−0.022	(0.022)
Energy Saving Score	0.003	(0.015)
Donated Environment	0.001	(0.021)
Happiness Score	−0.006	(0.008)
Policy Variables		
Policy awareness	−0.022**	(0.010)
Policy use	0.081***	(0.009)
Policy preference: Market	−0.003	(0.021)
Policy preference: Behavioral	−0.023	(0.023)
Policy preference: non-Market	0.025	(0.027)
Constant	1.079***	(0.367)
Observations	1,452	
R^2	0.082	

Note:
*p<0.1; **p<0.05; ***p<0.01
The dependent variable is binary and takes the value of one if respondents did an energy efficiency retrofit during the past five years or plan to do so within the next five years.

5.2 Heterogeneity: Barriers

We now investigate the barriers to energy efficiency retrofits. Table 8 presents three linear probability models, one for each major barrier elicited with the narratives. In these regression models, the dependent variable is a zero-one dummy variable that takes a value of one if the respondent mentioned a particular barrier in the open-ended question. The regressors are the different categories of variables for demographics, building characteristics, behaviors, as well as policy-related variables, which we further explain below.

In order to obtain more detailed information to help target policies, we included five policy variables as covariates in each regression: the index for policy awareness, the index for policy usage, and three dummy variables from the open question that capture policy preferences. The first two indices are based on questions that assess how many existing policies respondents are aware of and how many of these policies respondents have made use of or intend to use. For the policy-preference variables, we use the same classification for policy preferences as described earlier: market-based instruments (subsidies and tax deductions), behavioral instruments (such as information provision and standards), and non-market-based policy instruments (reducing bureaucracy). Then, we created a dummy variable for each of those categories taking a value of one if a respondent mentioned a policy option in this category, and zero otherwise.

Column 1 presents the heterogeneity for the barrier "the building is already energy-efficient". Most of the coefficients for traditional observable covariates are not statistically significant, except for the variable that measures energy literacy, which has a strong positive correlation with that barrier. However, respondents with this barrier tend to have a high degree of both policy knowledge and policy usage. In Column 2, for the financial barrier, income has a strong negative correlation. This result is intuitive: it simply implies that higher-income households are less likely to express financially-related issues as a barrier. In addition, these respondents have made less use of existing policies for retrofits. Column 3

investigates the old age barrier, where tenant age is significant and positive, which is again very consistent with the nature of the barrier expressed in the narratives. Here, none of the policy variables shows any statistical significance. Although, there are a few variables that are strongly correlated with each of the main barriers, observables explain little of the overall variance. This shows the difficulty of policy targeting and tagging for policy-makers. Results indicate that there is heterogeneity in both policy awareness and policy usage for the three main barriers. There is however no particular policy preference associated with any barrier, which suggests that the policy preferences are uniformly distributed over the different barriers.

Table 8: Linear Probability Model on Major Barriers of Retrofits

	Barrier		
	Already Efficient (1)	Expensive (2)	Old Age (3)
Building Age	−0.003**	0.002	0.0003
	(0.001)	(0.001)	(0.001)
Floor Size	0.001	−0.00004	−0.0004
	(0.001)	(0.001)	(0.0003)
Log Rental Value	0.045	0.160**	−0.035
	(0.068)	(0.063)	(0.036)
Log Income	0.034	−0.244**	−0.017
	(0.106)	(0.098)	(0.056)
Age	−0.001	−0.003	0.003
	(0.004)	(0.004)	(0.002)
Male	−0.052	0.121*	−0.066
	(0.077)	(0.072)	(0.041)
Children	−0.064	0.013	0.019
	(0.078)	(0.073)	(0.041)
University Degree	0.137**	−0.129**	0.042
	(0.069)	(0.063)	(0.036)
Pensioner	−0.021	−0.053	0.034
	(0.102)	(0.094)	(0.054)
Energy Literacy	0.130***	−0.0003	0.024
	(0.049)	(0.046)	(0.026)
Took Econ	−0.034	0.054	0.013
	(0.061)	(0.056)	(0.032)
Math Proficient	−0.092	0.053	−0.033
	(0.060)	(0.056)	(0.032)
Energy Saving Score	−0.022	0.019	−0.026
	(0.040)	(0.037)	(0.021)
Donated	0.052	−0.045	0.048
	(0.057)	(0.052)	(0.030)
Happiness Score	0.007	0.005	−0.005
	(0.022)	(0.020)	(0.012)
Policy Variables			
Policy Awareness	0.057**	−0.022	−0.017
	(0.024)	(0.022)	(0.013)
Policy Usage	0.086***	−0.067***	−0.007
	(0.027)	(0.025)	(0.014)
Policy Market	0.036	0.005	−0.034
	(0.056)	(0.052)	(0.029)
Policy Behavioral	0.031	−0.034	−0.015
	(0.063)	(0.058)	(0.033)
Policy Non-Market	0.062	0.023	0.022
	(0.074)	(0.069)	(0.039)
Constant	−0.714	1.394	0.421
	(0.998)	(0.924)	(0.526)
Observations	298	298	298
R^2	0.168	0.115	0.092

Note: *p<0.1; **p<0.05; ***p<0.01
Each column presents a separate linear probability model where the outcome is the respective barrier. The barrier is a binary variable that takes the value of one if the respondent mentioned it in the open text answer and zero otherwise.

5.3 Heterogeneity: Determinants

Turning to determinants, Table 9 presents four linear probability models, one for each major determinant of retrofits. Column 1 shows the result for the determinant "replacement of existing parts", i.e., obsolescence. Tenant age is significant and negatively correlated with that determinant, which can be explained by the shorter period older tenants expect to live in their home. Building age is positively correlated with that determinant, which is expected because older buildings need more frequent repairs compared to new buildings. Moreover, these respondents tend to have a higher policy awareness but no difference in the actual use of various subsidies. The "replacement" determinant is also associated with a policy preference for both market- and non-market oriented policies, however there is no difference for behavioral policies. In Column 2, we present the heterogeneity for the financial determinant "to save money". Interestingly, both a university degree and previous donations to environmental organizations are negatively associated with that determinant. In addition, respondents with the financial determinant do not have a higher policy awareness but rather a higher usage of policies. This group does not have any specific policy preferences. Column 3 investigates the comfort-related determinant. There are only few weakly significant variables associated with that determinant, namely a negative correlation for both pensioners and floor size and a positive correlation for math proficiency. Moreover, this determinant does not seem to show any statistically significant heterogeneity with respect to their awareness, usage of policies as well as policy preferences. Finally, in Column 4, the results indicate that several dimensions of heterogeneity are correlated with environmental concerns. Both income and previous donations to environmental organizations have a strong and statistically significant association with that determinant. Similar to respondents who renovate to save money, environmental concerns are associated with a higher degree of policy usage but not with more awareness. Furthermore, environmental motivations to renovate are strongly associated with behavioral policy preferences.

Table 9: Linear Probability Model on Major Determinants of Retrofits

	Determinant			
	Replacement	Save Money	Comfort	Environmental
	(1)	(2)	(3)	(4)
Building Age	0.002***	0.0004	0.001*	−0.0001
	(0.001)	(0.001)	(0.001)	(0.001)
Floor Size	0.0004	0.00001	−0.001**	−0.0002
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Log Rental Value	−0.003	−0.004	−0.016	−0.040
	(0.034)	(0.029)	(0.032)	(0.032)
Log Income	−0.056	0.075	0.033	0.113**
	(0.055)	(0.047)	(0.051)	(0.051)
Age	−0.007***	−0.002	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)
Male	−0.043	0.034	0.027	−0.013
	(0.038)	(0.033)	(0.035)	(0.036)
Children	−0.045	0.015	−0.024	0.005
	(0.039)	(0.033)	(0.036)	(0.037)
University Degree	0.021	−0.093***	−0.032	0.016
	(0.035)	(0.030)	(0.032)	(0.033)
Pensioner	0.015	0.024	−0.072	0.037
	(0.051)	(0.044)	(0.048)	(0.048)
Energy Literacy	0.064**	−0.007	0.008	−0.022
	(0.029)	(0.025)	(0.027)	(0.027)
Took Econ	−0.003	0.026	0.037	−0.004
	(0.032)	(0.027)	(0.029)	(0.030)
Math Proficient	−0.013	0.019	0.044	0.044
	(0.031)	(0.027)	(0.029)	(0.029)
Energy Saving Score	−0.019	−0.032*	0.028	0.029
	(0.022)	(0.018)	(0.020)	(0.020)
Donated	−0.034	−0.062**	−0.005	0.072**
	(0.030)	(0.025)	(0.028)	(0.028)
Happiness Score	−0.014	0.010	0.007	0.003
	(0.011)	(0.009)	(0.010)	(0.010)
Policy Variables				
Policy Awareness	−0.041***	0.007	0.004	0.017
	(0.015)	(0.013)	(0.014)	(0.015)
Policy Usage	−0.006	0.033***	0.014	0.035***
	(0.013)	(0.011)	(0.012)	(0.013)
Policy Market	0.086***	0.013	0.006	0.012
	(0.029)	(0.025)	(0.027)	(0.028)
Policy Behavioral	0.003	0.004	−0.006	0.079**
	(0.033)	(0.028)	(0.031)	(0.031)
Policy Non-market	0.102***	−0.042	−0.020	−0.040
	(0.037)	(0.032)	(0.034)	(0.035)
Constant	1.145**	−0.348	−0.063	−0.661
	(0.517)	(0.442)	(0.481)	(0.486)
Observations	1,154	1,154	1,154	1,154
R^2	0.060	0.038	0.020	0.041

Note: *p<0.1; **p<0.05; ***p<0.01
Each column presents a separate linear probability model where the outcome is the respective determinant. The determinant is a binary variable that takes the value of one if the respondent mentioned it in the open text answer and zero otherwise.

5.4 Targeting

Based on the insights from the heterogeneity analysis, we can draw conclusions for targeting on a broad scale between takers and non-takers of retrofits, and on a more granular scale within those groups, specific to each barrier and determinant.
Between takers and non-takers, the main difference is a lower age for takers. Apart from age, we find no other observables that policy makers could robustly use to target policies: policy awareness is actually lower for takers and there are no significant differences in policy preferences. For this reason, we propose to focus on the specific barriers and determinants that respondents mentioned in their open answers. By analyzing how to target specific barriers and determinants, policy makers can address the core reasons why homeowners decide for or against a retrofit. Within the two groups, we can use the results from this section to cluster respondents into groups and see if different household types emerge.

For the barriers, two types of homeowners emerge: those who do not renovate because they perceive their house already as energy-efficient and those who face financial constraints. Respondents who do not renovate because their building is already efficient do not differ in income or age from other non-takers. The main difference can be found in a higher educational level. Those respondents also have a high awareness and experience of using policies, they do not have any particular policy preferences. For this reason, there are few options to target these homeowners with specific policies. In contrast, homeowners who did not renovate due to financial constraints have a lower income, but also do not differ in terms of policy awareness or policy preferences, they have less experience in using policies though. It seems that policy preferences do not play a role in explaining why certain consumers perform retrofits and others do not. Possibly, existing policies are not of a sufficient magnitude to help overcoming these barriers to retrofits.

The determinants have three major groups that are of interest for policy targeting: households who invest to replace parts of the building, households who see retrofits as profitable investments and households who renovate out of ecological concerns. The replacements-category is primarily motivated by a greater building age, a lower respondent age and a high energy literacy. Moreover, this group has a lower awareness of existing policies and would prefer policies that reduce bureaucracy and higher subsidies. Most of the characteristics for this group are very rational given that the primary motivation is to replace broken parts of the home (which is often a necessity). Even though the policy awareness is low for this group, respondents do not favor more information on policies but would prefer less bureaucracy and higher subsidies. It seems that respondents with this motivation renovate out of necessity and would like to facilitate this process. The second group that renovates to save money is characterized by less education and fewer donations to environmental organizations. This group also has a weakly higher income compared to others. For their retrofits, these respondents made more use of existing policies compared to other groups. They do not have any particular policy preferences though. As for the previous group that renovates out of necessity, it seems that respondents who renovate to save money act out of financial opportunity. For this reason, these homeowners make use of existing policies but do not have any particular preference about what policy makers could improve. For targeting, this group does not show any particular angle policy makers can address. The last group consists of respondents who renovated out of environmental concerns. Those respondents have a higher income and previously donated to environmental organizations. They made more use of policies compared to other groups and would strongly favor behavioral policies such as information campaigns and ecological standards. For policy targeting this group is particularly interesting because those respondents do not renovate out of necessity or financial opportunity. Policy makers could target this group by providing easier access to information regarding policies.

6 Conclusions

In this paper, we propose a novel approach to elicit the barriers and determinants of energy efficiency investments. Narratives offer a powerful way to elicit and rank important barriers and determinants of households' retrofit decisions. Our results first suggest that energy efficiency investments are highly opportunistic. Non-takers believe, rightfully or not, that little opportunities for energy efficiency exist in their home. Takers primarily invest in energy efficiency out of necessity to replace old parts of the building or out of financial opportunity when they perceive the investment as profitable. The monetary aspect is also important as a major barrier, because many respondents stated that they face financial constraints with respect to renovation plans. However, several co-benefits of energy efficiency also emerge as important determinants, mainly the increased comfort. Finally, environmental concerns showed to be a major determinant for energy efficiency retrofits.

A more granular analysis of barriers and determinants gave further insights how to target energy efficiency policies. The majority of characteristics that influence a respondents' decision to perform a retrofit are difficult to target for policy makers. For instance, the main difference between takers and non-takers is a higher tenant age for non-takers. We asked respondents about their awareness and usage about existing policies and about preferences for additional policy measures. These policy measures were particularly useful in explaining heterogeneity within the group of takers, between different determinants. Based on these results, the most promising group for future policy targeting are home owners who renovate out of environmental concerns. These homeowners tend to have a higher income and would favor behavioral policies that consist of information campaigns and building standards.

From a methodological standpoint, eliciting barriers and determinants with closed- vs. open-ended questions shows the difficulties of precise policy targeting. Both methods yield broadly consistent results but there are some important differences. If we were to target

energy efficiency policies based on closed- vs. open-ended questions, we might achieve a very different allocation.

References

A. Alberini, S. Banfi, and C. Ramseier. Energy efficiency investments in the home: Swiss homeowners and expectations about future energy prices. *The Energy Journal*, 34(1), 2013.

H. Allcott and M. Greenstone. Is there an energy efficiency gap? *Journal of Economic Perspectives*, 26(1):3–28, 2012.

H. Allcott and M. Greenstone. Measuring the welfare effects of residential energy efficiency programs. *NBER Working Paper*, (w23386), 2017.

H. Allcott, C. Knittel, and D. Taubinsky. Tagging and targeting of energy efficiency subsidies. *American Economic Review*, 105(5):187–91, 2015.

D. Angelov. Top2vec: Distributed representations of topics. *arXiv preprint arXiv:2008.09470*, 2020.

S. Banfi, M. Farsi, M. Filippini, and M. Jakob. Willingness to pay for energy-saving measures in residential buildings. *Energy economics*, 30(2):503–516, 2008.

J. Blasch, M. Filippini, and N. Kumar. Boundedly rational consumers, energy and investment literacy, and the display of information on household appliances. *Resource and Energy Economics*, 56:39–58, 2019.

J. Blasch, N. Boogen, C. Daminato, and M. Filippini. Empower the consumer! energy-related financial literacy and its implications for economic decision making. *Economics of Energy & Environmental Policy*, 10(2), 2021.

E. Cagno, E. Worrell, A. Trianni, and G. Pugliese. A novel approach for barriers to industrial energy efficiency. *Renewable and Sustainable Energy Reviews*, 19:290–308, 2013.

C. Cattaneo. Internal and external barriers to energy efficiency: which role for policy interventions? *Energy efficiency*, 12(5):1293–1311, 2019.

N. Egami, C. J. Fong, J. Grimmer, M. E. Roberts, and B. M. Stewart. How to make causal inferences using texts. *arXiv preprint arXiv:1802.02163*, 2018.

European Commission. Report of the work of the task force on mobilising efforts to reach the eu energy efficiency targets for 2020. January 2019.

U. Fischbacher, S. Schudy, and S. Teyssier. Heterogeneous preferences and investments in energy saving measures. *Resource and Energy Economics*, 63:101202, 2021.

T. D. Gerarden, R. G. Newell, and R. N. Stavins. Assessing the energy-efficiency gap. *Journal of Economic Literature*, 55(4):1486–1525, 2017.

K. Gillingham, R. G. Newell, and K. Palmer. Energy efficiency economics and policy. *Annu. Rev. Resour. Econ.*, 1(1):597–620, 2009.

K. Gillingham, A. Keyes, and K. Palmer. Advances in evaluating energy efficiency policies and programs. *Annual Review of Resource Economics*, 10:511–532, 2018.

E. Grave, P. Bojanowski, P. Gupta, A. Joulin, and T. Mikolov. Learning word vectors for 157 languages. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)*, 2018.

M. Honnibal, I. Montani, S. Van Landeghem, and A. Boyd. spaCy: Industrial-strength Natural Language Processing in Python, 2020.

N. Hrovatin and J. Zorić. Determinants of energy-efficient home retrofits in slovenia: The role of information sources. *Energy and Buildings*, 180:42–50, 2018.

A. B. Jaffe and R. N. Stavins. The energy-efficiency gap what does it mean? *Energy policy*, 22(10):804–810, 1994.

M. Jakob. Marginal costs and co-benefits of energy efficiency investments: The case of the swiss residential sector. *Energy policy*, 34(2):172–187, 2006.

M. Jakob et al. The drivers of and barriers to energy efficiency in renovation decisions of single-family home-owners, 2007.

J. A. Krosnick. Survey research. *Annual review of psychology*, 50(1):537–567, 1999.

K.-H. Lee. Drivers and barriers to energy efficiency management for sustainable development. *Sustainable Development*, 23(1):16–25, 2015.

A. Lusardi and O. S. Mitchell. Planning and financial literacy: How do women fare? *American economic review*, 98(2):413–17, 2008.

S. Lyubomirsky and H. S. Lepper. A measure of subjective happiness: Preliminary reliability and construct validation. *Social indicators research*, 46(2):137–155, 1999.

A. K. McCallum. Mallet: A machine learning for language toolkit. http://mallet.cs.umass.edu, 2002.

C. E. Moody. Mixing dirichlet topic models and word embeddings to make lda2vec. *arXiv preprint arXiv:1605.02019*, 2016.

M. E. Roberts, B. M. Stewart, D. Tingley, C. Lucas, J. Leder-Luis, S. K. Gadarian, B. Albertson, and D. G. Rand. Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4):1064–1082, 2014.

M. E. Roberts, B. M. Stewart, and D. Tingley. stm: An R package for structural topic models. *Journal of Statistical Software*, 91(2):1–40, 2019. doi: 10.18637/jss.v091.i02.

J. Schleich, X. Gassmann, C. Faure, and T. Meissner. Making the implicit explicit: A look inside the implicit discount rate. *Energy Policy*, 97:321–331, 2016.

J. Schleich, X. Gassmann, T. Meissner, and C. Faure. A large-scale test of the effects of time discounting, risk aversion, loss aversion, and present bias on household adoption of energy-efficient technologies. *Energy Economics*, 80:377–393, 2019.

R. J. Shiller. *Narrative economics*. Princeton University Press, 2020.

S. Sorrell, E. O'Malley, J. Schleich, and S. Scott. *The Economics of Energy Efficiency: Barriers to Cost-Effective Investment*. Edward Elgar Publisher, 2004.

S. Stantcheva. Understanding economic policies: What do people know and how can they learn. *Harvard University Working Paper*, 2020.

R. J. Sutherland. Market barriers to energy-efficiency investments. *The Energy Journal*, 12 (3), 1991.

G. Trotta. Factors affecting energy-saving behaviours and energy efficiency investments in british households. *Energy Policy*, 114:529–539, 2018.

D. Ürge-Vorsatz, S. T. Herrero, N. K. Dubash, and F. Lecocq. Measuring the co-benefits of climate change mitigation. *Annual Review of Environment and Resources*, 39:549–582, 2014.

7 Appendix

7.1 Variable Description

Building Characteristics

Building Age

The respondent's building age in years.

Floor size

The floor-size in square meters.

Number of rooms

Number of rooms, excluding kitchen, bathroom and WC.

Garden

A dummy variable that takes the value of one if the respondent's house has a garden and zero otherwise.

Rental Value

The self-estimated monthly rental value that a respondent would obtain for renting their house on the market. Respondents usually have a proxy for that rental value because it is important in Switzerland for tax purposes.

Heating

Respondents were asked what primary source of heating they use for their house. They could choose between four options: oil, gas, heat pump and "other". Oil and gas were taken together as one variable.

Solar PV

A dummy variable that takes the value of one if the respondent's house has solar panels.

Demographics

Income

The respondents gross household income. Respondents could choose between the following brackets: "below 8 000 CHF", "8 000 - 12 000 CHF", "12 000 - 16 000 CHF", "16 000 - 20 000 CHF", "above 20 000 CHF" and "no answer". Respondents with "no answer" were omitted from the data-set. We converted "below 8 000 CHF" to 8 000 CHF and "above 20 000 CHF" to 22 000 CHF, for all other brackets we chose the average number between the two bounds (10 000, 14 000 and 18 000 CHF respectively).

Age

The respondents age in years.

Male

A dummy variable that takes the value of one if the respondent's gender is male and zero otherwise.

Children

A dummy variable that takes the value of one if the respondent's household includes children.

University Degree

A dummy variable that takes the value of one if the respondent holds a university degree. We also inquired if a respondent's spouse holds a university degree, when applicable.

Employment

Three categories for respondent's current employment situation: full-time employment, pensioner and other employment (including part-time employment). We also inquired the same information for a respondent's spouse if applicable.

Allergies

A dummy variable that takes the value of one if a respondent or a member of their household suffers from any of the following allergies: dust, mites pollen, animal hair or feathers.

Psychographics

Energy Literacy

In order to obtain a proxy on financial literacy in the context of energy efficient investments, we used a reduced version of a score that is based on Blasch et al. (2021). Specifically, we first use the three classical financial literacy questions by Lusardi and Mitchell (2008): the first question inquires the knowledge about interest rates, the second about the effect of inflation on investment and the third question addresses the importance of portfolio diversification. Each question can be answered correctly or incorrectly which gives each respondent a total score from zero to three. Following Blasch et al. (2021) we added two questions to this score: the first questions asks for an estimate for the electricity price in the Canton of Zurich. The actual price is around 0.20 CHF/kWh but we considered all responses in the range between 0.06 CHF/kWh and 0.30 CHF/kWh as being in the correct order of magnitude. The second question gave a hypothetical investment decision in two heating systems with different initial costs and different energy savings per year. The respondent had to calculate which of the two heating systems is less expensive after 20 years (without considering inflation or alternative investments). In total, respondents answered 5 questions and could obtain a score between 0 and 5.

Took Econ

A dummy variable that takes the value of one if the respondent took economics classes during their education.

Math Proficient

We asked respondents how they self-assess their math-proficiency while they were in school. The possible answers were "I do not remember anymore", "below average", "average" and "above average". Out of these answers we constructed a dummy variable that takes the value of one if the respondent answered "above average" and zero otherwise.

Energy Saving Score

We presented respondents with three everyday activities that consume energy but also allow to save energy: "use washing machine and dishwasher only if it is fully loaded", "turn off the light when leaving the room, even for a short amount of time" and " fully turn-off electrical appliances such as TV or computer (no standby)". For each situation, respondents could choose between "never", "rarely", "sometimes", "often" and "always". We constructed for each of the three situations a dummy variable that takes the value of one if the respondent chose "often" or "always" and zero otherwise. To obtain a score between zero and three, we added the three dummy variables.

Donated

A dummy variable that takes the value of one if the respondent donated to an environmental organization during the past 12 month and zero otherwise.

Happiness Score

We asked two questions by Lyubomirsky and Lepper (1999) where respondents could rate their own happiness as well as their perceived happiness relative to their peers on a scale of 1 to 7. Similarly, we asked respondents to rate their happiness with their home on a scale of 1 to 7 and their happiness with their home relative to their peers. For each question we created a dummy variable that takes the value of one if the score if the individual score is above the mean score for the entire sample and zero otherwise. We then took the sum of the four dummy variables to obtain a score between zero and four for each respondent.

Working Papers of the Center of Economic Research at ETH Zurich

(PDF-files of the Working Papers can be downloaded at www.cer.ethz.ch/research/working-papers.html).

21/359 S. Houde, T. Wekhof
The Narrative of the Energy Efficiency Gap

21/358 F. Böser, H. Gersbach
Leverage Constraints and Bank Monitoring: Bank Regulation versus Monetary Policy

21/357 F. Böser
Monetary Policy under Subjective Beliefs of Banks: Optimal Central Bank Collateral Requirements

21/356 D. Cerruti, M. Filippini
Speed limits and vehicle accidents in built-up areas: The impact of 30 km/h zones

21/355 A. Miftakhova, C. Renoir
Economic Growth and Equity in Anticipation of Climate Policy

21/354 F. Böser, C. Colesanti Senni
CAROs: Climate Risk-Adjusted Refinancing Operations

21/353 M. Filippini, N. Kumar, S. Srinivasan
Behavioral Anomalies and Fuel Efficiency: Evidence from Motorcycles in Nepal

21/352 V. Angst, C. Colesanti Senni, M. Maibach, M. Peter, N. Reidt, R. van Nieuwkoop
Economic impacts of decarbonizing the Swiss passenger transport sector

21/351 N. Reidt
Climate Policies and Labor Markets in Developing Countries

21/350 V. Britz, H. Gersbach
Pendular Voting

21/349 E. Grieg
Public opinion and special interests in American environmental politics

21/348 N. Ritter, J. A. Bingler
Do homo sapiens know their prices? Insights on dysfunctional price mechanisms from a large field experiment

20/347 C. Daminato, M. Filippini, F. Haufler
Personalized Digital Information and Tax-favoured Retirement Savings: Quasi-experimental Evidence from Administrative Data

20/346 V. Britz, H. Gersbach
Open Rule Legislative Bargaining

20/345 A. Brausmann, E. Grieg
Resource Discoveries and the Political Survival of Dictators

20/344 A. Jo
The Elasticity of Substitution between Clean and Dirty Energy with Technological Bias

20/343 I. van den Bijgaart, D. Cerruti
The effect of information on market activity; evidence from vehicle recalls

20/342 H. Gersbach, R. Wattenhofer
A Minting Mold for the eFranc: A Policy Paper

20/341 L. Bretschger
Getting the Costs of Environmental Protection Right

20/340 J. A. Bingler, C. Colesanti Senni
Taming the Green Swan: How to improve climate-related financial risk assessments

20/339 M. Arvaniti, T. Sjögren
Temptation in Consumption and Optimal Redistributive Taxation

20/338 M. Filippini, S. Srinivasan
Voluntary adoption of environmental standards and limited attention: Evidence from the food and beverage industry in Vietnam

20/337 F. Böser, C. Colesanti Senni
Emission-based Interest Rates and the Transition to a Low-carbon Economy

20/336 L. Bretschger, E. Grieg, P. J.J. Welfens, T. Xiong
Corona Fatality Development and the Environment: Empirical Evidence for OECD Countries

20/335 M. Arvaniti, W. Habla
The Political Economy of Negotiating International Carbon Markets

20/334 N. Boogen, C. Daminato, M. Filippini, A. Obrist
Can Information about Energy Costs Affect Consumers Choices? Evidence from a Field Experiment

20/333 M. Filippini, N. Kumar, S. Srinivasan
Nudging the Adoption of Fuel-Efficient Vehicles: Evidence from a Stated Choice Experiment in Nepal

20/332 L. Bretschger, E. Grieg
Exiting the fossil world: The effects of fuel taxation in the UK

20/331 H. Gersbach, E. Komarov
Research Bubbles

20/330 E. V. Dioikitopoulos, C. Karydas
Sustainability traps: patience and innovation

19/329 M. Arvaniti, C. K. Krishnamurthy, A. Crepin
Time-consistent resource management with regime shifts

19/328 L. Bretschger, K. Pittel
Twenty Key Questions in Environmental and Resource Economics

19/327 C. Karydas, A. Xepapadeas
Climate change financial risks: pricing and portfolio allocation

19/326 M. Filippini, S. Srinivasan
Investments in Worker Health and Labor Productivity: Evidence from Vietnam

19/325 H. Gersbach
Democratizing Tech Giants! A Roadmap

19/324 A. Brausmann, M. Flubacher and F. Lechthaler
Valuing meteorological services in resource-constrained settings: Application to smallholder farmers in the Peruvian Altiplano

19/323 C. Devaux and J. Nicolai
Designing an EU Ship Recycling Licence: A Roadmap

19/322 H. Gersbach
Flexible Majority Rules for Cryptocurrency Issuance

19/321 K. Gillingham, S. Houde and A. van Benthem
Consumer Myopia in Vehicle Purchases: Evidence from a Natural Experiment

19/320 L. Bretschger
Malthus in the Light of Climate Change

19/319 J. Ing and J. Nicolai
Dirty versus Clean Firms' Relocation under International Trade and Imperfect Competition