Using narratives to infer preferences in understanding the energy efficiency gap

Journal Article

Author(s):
Wekhof, Tobias (iD); Houde, Sébastien

Using narratives to infer preferences in understanding the energy efficiency gap

Tobias Wekhof [1,2] ✉ **& Sébastien Houde[3]**

Investing in energy efficiency is crucial for a low-carbon economy, particularly in the building sector. Despite various subsidy programmes, meeting energy targets is challenging because households do not invest sufficiently. Here we study the low numbers of energy efficiency retrofits carried out by homeowners. We use narratives, an emerging method based on open-ended survey responses, to identify the barriers and determinants behind renovation decisions. Using natural language processing, we transform narratives into quantifiable metrics. Whereas financial considerations are a major barrier for homeowners, their main reasons for renovating are not related to energy savings. Most homeowners delay energy-saving investments until their buildings require renovations. Co-benefits such as environmental concerns and comfort gains are equally or more important than financial motivations. Many homeowners are unaware of existing policies and would favour reducing the bureaucracy of retrofits. Subsidies, although popular, are likely to be mistargeted. Effective policies should also consider institutional factors such as the bureaucratic burden and the accessibility of information.

Investments in energy efficiency are essential for the transition to a low-carbon economy. Most, if not all, developed economies have set ambitious targets for energy retrofits in the building sector. However, energy efficiency targets are difficult to achieve[1]. This observation is well known—for more than 40 years, researchers have pointed out that society is lagging behind in energy efficiency investments, a phenomenon known as the 'energy efficiency gap'[2]. The key question regarding this gap is whether or not systematic barriers and household preferences are impeding the adoption of seemingly cost-effective, energy-efficient technologies.

On the basis of personal narratives combined with natural language processing (NLP), we propose a method to study the energy efficiency gap. To elicit narratives related to energy efficiency investments, we conducted an extensive survey among single-family homeowners in the Canton of Zurich, Switzerland. In particular, we asked respondents to express their thoughts on specific topics in open-ended responses: What are the main barriers and determinants to energy efficiency retrofits? What policy measures would households prefer, and how are these measures related to the barriers and determinants of energy efficiency investments?

Our Article makes two main contributions. First, we add new evidence to the empirical literature on the barriers and determinants of households regarding their energy efficiency investments[3–10]. As discussed in a review of the literature[11], there are important gaps in these data—most studies focus on a predetermined set of explanatory variables due to the limitation of closed-ended questionnaires. This makes it difficult to generalize findings, and this approach is also prone to researchers' biases. It does not enable the decision-makers' thought processes to be uncovered about the most relevant barriers and determinants.

Second, we contribute to the emerging literature in economics that uses narratives to explain the drivers of decisions[12]. Our approach enables us to uncover more nuanced barriers and determinants of technology adoption than closed-ended survey questions (that is, multiple-choice questions). We provide a proof of concept that researchers and policy-makers can easily implement, scale to large samples and replicate across contexts. Our method has the potential to examine the energy efficiency gap in a broader context and, more generally, to robustly identify behavioural barriers and determinants related to energy efficiency.

[1]ETH Zurich, Center of Economic Research, Zurich, Switzerland. [2]University of Zurich, Department of Banking and Finance, Zurich, Switzerland. [3]Université de Lausanne, Faculty of Business and Economics (HEC), Lausanne, Switzerland. ✉e-mail: twekhof@ethz.ch

We find that narratives can help understand the personal factors of households, particularly the barriers and determinants to undertaking energy efficiency retrofits. Our main finding states that energy efficiency investments are highly opportunistic. Non-takers (that is, homeowners who did not invest in energy efficiency) often believe that their homes are already energy efficient enough and therefore do not seek such investments. For takers, that is, homeowners who have invested in energy efficiency, a large share did so because a building technology no longer functioned and needed replacing. Financial considerations are a major barrier to renovation, but for those households that have invested, they are not the primary reason. Co-benefits, namely comfort gains and reducing the environmental footprint, play a similar or even more important role.

Survey data

We collaborated with the Statistical Office of the Swiss Canton of Zurich to recruit participants for our research project. We sent personalized invitation letters to a random sample of single-family homeowners, stratified along the following criteria: those owning a home that was constructed before 1990, and where 50% were to those with renovation permits during the 5 years leading up to the survey. We stratified the tenant age and the number of tenants by considering buckets for both variables; in addition, we targeted homeowners with new buildings who adopted the primary certification for energy-efficient buildings in Switzerland, that is, 'Minergie' certification.

Of the 16,700 letters sent, the response rate was high (20.8%), with 3,471 respondents starting the survey. A detailed description of the variables is given in Supplementary Tables 5–7. For a detailed overview of the survey procedure, see 'Survey procedure' in Methods.

We identified two groups of households based on their past and intended energy efficiency behaviours. The first group, who adopted the Swiss energy efficiency certification for new buildings (Minergie), was not included in this analysis. The second group comprised households with houses built before 1990 that had either performed energy efficiency retrofits in the past 5 years or planned to do so within the next 5 years. From these, we categorized households as either non-takers (21% of the sample) or takers (79% of the sample) based on whether they had invested in energy efficiency or planned to do so.

Table 1 shows the different household types and how they differ for key building characteristics, demographics and psychographics (variables that measure the degree of energy-related literacy, support for environmental causes and self-reported happiness). Overall, takers and non-takers are similar in the building characteristics and demographic features, except for a slightly higher age among non-takers. For this reason, we focus on the specific barriers and determinants to the energy-efficient retrofits mentioned by takers and non-takers.

Barriers and determinants of energy efficiency investments

This section presents the barriers and determinants to energy-efficient retrofits from our survey. Takers, that is, respondents who had carried out a retrofit in the 5 years before the survey or who intended to do so in the next 5 years, provided their motivations for the renovation. Non-takers stated their reasons against retrofitting. For more information on the method used to analyse the text responses, see 'Eliciting energy efficiency narratives' and 'Semi-manual classification' in Methods.

We grouped all barriers and determinants into four categories that correspond to the different types of market barriers: classic 'market'-type barriers are normal components of markets that impact decisions, such as heterogeneity in building stock; 'non-market'-type barriers refer to co-benefits and hidden costs, for example, hassle costs or increased comfort; 'financial'-type barriers are related to prices and costs; and 'behavioural'-type barriers describe the psychological, cognitive and educational factors of decision-making for households.

Table 1 | Summary statistics by household type

Variable	Non-takers	Takers
Building characteristics		
Building age (years)	57.31	56.13
Floor size (m²)	167.24	168.78
Rental value (CHF per month)	4,045.96	3,772.20
Oil/gas heating (%)	52.40	48.50
Heat pump (%)	26.10	34.40
Other heating (%)	21.50	17.10
Solar PV panels (%)	13.40	20.00
Demographics		
Income (CHF per month)	12,455.70	12,517.97
Age (years)	61.36	58.13
Female (%)	20.50	22.50
Children (%)	42.50	49.50
University degree (%)	61.80	59.70
Psychographics		
Energy literacy (scored out of 5)	3.73	3.76
Studied economics (%)	43.40	47.60
Proficient in mathematics (%)	47.10	45.20
Energy-saving score (scored out of 3)	2.25	2.31
Donated environment (%)	53.50	55.40
Happiness score (scored out of 4)	2.17	2.10

This table presents the summary statistics for the entire survey sample, with the exception of respondents who live in a Minergie-certified building. The final sample contains a total of 2,187 observations, out of which 461 (21%) are non-takers and 1,726 (79%) are takers. Takers are defined as respondents who carried out an energy efficiency retrofit either in the past 5 years or who plan to do so within the next 5 years. In the subsequent analysis, to avoid limiting the sample size, we did not use all the variables. Specifically, the variables not used were the heating modes and solar photovoltaic (PV) devices.
CHF, Swiss francs.

For a detailed taxonomy of barriers and determinants, see 'Taxonomy of barriers and determinants' in Methods.

Table 2 shows the barriers for the non-takers that were elicited from the open- and closed-ended questions. Primarily, non-takers believe that their home is already energy efficient. In terms of consistency between the open-ended and closed-ended questions, this barrier is found to be the most prevalent for both approaches. When using narratives, ~49% of respondents wrote that their home was already energy-efficient, but only ~38% chose this option as a potential barrier in the closed-ended question.

However, the two building characteristics that generally have the greatest impact on a dwelling's energy efficiency potential—the year of construction and type of heating system—did not differ drastically between non-takers and takers (see Table 1). Therefore, the belief of the non-takers about the energy efficiency of their homes could either be a personal preference (a normal part of markets) or a misperception. In the latter case, information campaigns and subsidized audits may be warranted to address this barrier.

The second most important barrier is the cost of retrofitting. Although important compared with the other barriers, it was mentioned by ~26% in the narratives and by ~22% in the closed-ended answers. Financial barriers are at the heart of energy efficiency programmes, which have led to generous subsidy programmes in Switzerland and elsewhere. Our results suggest that policymakers may shift their focus away from this financial barrier.

The third most important barrier shows a discrepancy between the open-ended and closed-ended questions. In the narratives, the

Table 2 | Barriers to energy efficiency retrofits for non-takers

Barrier	Type	Open (%)	Closed (%)
The building is already energy efficient	Market	49.5	38.4
Too expensive	Financial	26.2	21.9
Old age (of the respondent)	Market	8.9	0.0
Too complicated	Non-market	7.2	10.0
Aesthetics	Market	3.0	8.0
Difficulties due to historic building	Market	2.8	6.1
Expert recommended against	Behavioural	2.6	1.7
Other priorities	Non-market	1.1	0.0
Difficulties in applying for permits	Non-market	0.9	3.3
Difficulties in obtaining financing	Financial	0.7	4.1
Craftsman recommended against	Behavioural	0.7	2.4
Hassle	Non-market	0.4	0.0
Architect recommended against	Behavioural	0.2	1.7
Homeowner did not think of it	Behavioural	0.0	4.6
Planning to move	Market	0.0	6.5
Lack of information	Behavioural	0.0	8.7
The investment is too risky	Behavioural	0.0	2.2
Leaving the house during the renovation	Non-market	0.0	5.9
It is difficult to find experts or materials	Non-market	0.0	4.6
Bad experiences with previous renovations	Non-market	0.0	1.7

Non-takers are households that did not undertake a retrofit in the past 5 years and did not plan to do so in the next 5 years, who were asked to choose among several options for the reasons they decided against a retrofit. The column 'Closed' corresponds to the share of non-takers that select each particular barrier (the number of choices a respondent could make was not limited). Later in the survey, we asked the same respondents an open-ended question for the reasons they did not carry out a retrofit. We then classified the text answers into the same categories as the closed-ended answers and added several new topics, such as old age. The column 'Open' corresponds to the share of non-takers that mentioned a particular barrier.

Table 3 | Determinants of energy efficiency retrofits for takers

Determinants	Type	Open (%)	Closed (%)
Replace broken elements	Market	45.5	57.6
To reduce my ecological footprint	Non-market	30.0	69.0
To save money	Financial	28.8	36.7
To increase comfort	Non-market	24.6	68.5
Increase resale value	Financial	4.8	25.2
Regulatory	Non-market	4.3	0.0
Increase size of home	Market	3.2	0.0
Recommended by another expert	Behavioural	2.1	6.0
Aesthetics	Non-market	2.0	0.0
Safety	Non-market	1.0	0.0
Recommended by an architect	Behavioural	0.9	4.9
Recommended by a craftsman	Behavioural	0.5	4.1

Takers are households that either performed a retrofit in the past 5 years prior to the survey or who plan to do so in the next 5 years, and they were asked to choose among several options for why they decided to perform a retrofit (the number of choices a respondent could make was not limited). Later in the survey, we asked the same respondents an open-ended question as to why they decided to carry out a retrofit. We then classified the text answers into the same categories as the closed answers and added several new topics, such as the 'regulatory' determinant.

age of the respondent, for advanced ages in particular, is the third most mentioned barrier. Old age is not a topic we listed a priori in the closed-ended question. With hindsight, we recognize that this can be an important barrier. Retrofits are long-term investments that older homeowners may not fully realize during their expected lifetimes. In our taxonomy of barriers, heterogeneous life expectancy is a normal component of markets that does not require policy intervention.

In the closed-ended question, the hassle costs, that is, the fact that homeowners perceive such an investment to be too complicated, are the third most important barrier, with 10% mentioning it in the closed-ended answers. The narratives show a similar but slightly lower percentage of respondents mentioning this barrier.

The respondents selected several other barriers with low frequency in the closed-ended questions but did not mention them in the open-ended responses. When respondents choose among a predefined list of elements, it is almost costless for them to select an additional option. Thus, the closed-ended question leads to a greater variety of barriers, but cheap talk could also be involved. By contrast, writing about an additional barrier in response to an open-ended question requires more effort. Therefore, open-ended questions could lead to more truthfulness in identifying the most important barrier(s) faced

by each non-taker. For example, aesthetics and the difficulties associated with renovating landmarked buildings are two barriers that were highlighted as important in the closed-ended questions; however, respondents rarely mentioned them in their narratives.

Table 3 presents the determinants of energy efficiency renovations for takers. The elicitation procedure plays an even more important role compared with the barriers. In particular, it leads to greater variation in the share of each topic and ranking between open- and closed-ended questions.

From the narratives, the most important determinant is the need to replace broken building parts. The importance of technological obsolescence is consistent with the main barrier for non-takers, who perceive that they have few options for improving energy efficiency. This observation suggests that energy efficiency investments are highly opportunistic and not the result of well-planned replacements. Homeowners often do not start thinking about energy efficiency until a building technology malfunctions.

In the closed-ended question, the top-ranking determinants were reducing the environmental footprint and comfort—more than two-thirds of the respondents chose these options. In the open-ended question, these two factors ranked second and fourth, respectively; however, they were mentioned by less than one-third of the respondents. Nonetheless, both elicitation procedures show that the non-market benefits associated with energy efficiency investments are important determinants.

The relative decrease in the importance of the environmental footprint between the closed- and open-ended answers may be related to the 'intention–behaviour' gap in sustainable consumption, where respondents indicate that they value ecological product aspects but do not reflect this in their purchases[13].

The third most important determinant relates to financial motivations. Again, there is a large discrepancy between the share of respondents who indicated this determinant in the open- and closed-ended answers. Specifically, ~29% of respondents mentioned this determinant in the narratives, but ~37% selected this topic in the closed-ended question. In addition, the impact of energy efficiency on the resale value received

Table 4 | Policy preferences from open-ended answers

Policy preference	Non-takers (%)	Takers (%)
More subsidy (market)	40.6	44.4
More information (behavioural)	20.8	19.8
Less bureaucracy (non-market)	16.3	19.5
Focus on PV panels	12.6	17.9
Focus on heating	11.3	13.8
Standards (behavioural)	10.0	10.7
Tax deduction (market)	9.1	8.9
Pollution tax (market)	8.5	8.6
Focus on insulation	2.2	4.6
Focus on new buildings	4.6	4.2
Technology	2.4	2.8
Property tax (market)	0.4	1.6
Subsidy threshold	0.9	1.4
Credit (market)	0.2	0.5

This table presents policy preferences obtained by classifying the responses to the open-ended question on policy preferences. The keywords from the topic dictionary are unique to each topic, and responses can contain multiple topics.

little attention in the open-ended question, at less than 5%. However, around 25% of respondents still selected this determinant in the closed-ended question.

Overall, the results clearly show the main reasons for investing in energy efficiency: obsolescence, non-market benefits and financial considerations. However, the respective importance of each determinant depends on the elicitation procedure. This, in turn, has important implications for targeting energy efficiency measures—depending on whether closed- or open-ended questions were used, policymakers may choose different measures.

Policy preferences

For the design of energy efficiency policies, it is crucial to identify the level of awareness and preferences for specific policy instruments, as well as how these affect key barriers and determinants. We first show the level of general awareness of the policy landscape and the respondents' experiences with these policies. In the second step, we analyse the responses to the open-ended question on policy preferences.

We asked respondents about their awareness of the four main energy efficiency policies in Switzerland: rebates on mortgage interest, tax exemptions or deductions, various subsidies from cantons and municipalities, and the Swiss Federal Building Program. For each policy, the respondents were asked to choose one of four options: not aware; aware; used; or intend to use. The results show that the takers were slightly more aware of the policies than the non-takers, and that they used them more frequently. This trend is also reflected in two indices that we created, that is, policy awareness and policy use. Each index would score a maximum of four points if the respondent was aware of all four policies or had used them in the past. For more information on policy awareness and use, see Supplementary Methods 3.

In the second step, we relied on narratives to determine the policy preferences. Unlike for barriers and determinants, we used only an open-ended question. Table 4 shows the respondents' policy preferences for different instruments. A wide range of topics emerged from the narratives. When asked how policies could encourage energy efficiency investments for all households, the top suggestion was more generous subsidies, followed by policymakers providing more information.

The third most common suggestion was to reduce bureaucracy (especially related to subsidies and building permits). Other

suggestions included promoting standards and a greater focus on PV panels and on heating systems. Other topics, which accounted for a smaller share, also appeared in the narratives. Tax-related measures were discussed but were not a popular topic, especially compared with subsidies. Although subsidies were the most frequently mentioned topic, more than half of the respondents favoured other measures.

We grouped all of the policy options suggested by respondents into three broad categories, depending on what type of barriers an instrument may best address. The first category consists of market-based instruments: policy options related to subsidies and taxes. The second category consists of behavioural instruments, including instruments motivated by behavioural biases, most notably information provision and standards. The third category consists of non-market-based policy instruments that involve other interventions, such as reducing bureaucracy.

Heterogeneity and targeting

Targeting policies to populations where they will be effective is critical to improving energy efficiency. Numerous studies that have conducted ex-post evaluations of energy efficiency subsidies have found them to be an expensive way to reduce carbon emissions—far beyond the social cost of carbon[14–17]. The main reason is that these subsidies are not targeted, and many recipients are so-called free riders, that is, these households would have made the same energy efficiency investments even without the often generous subsidies. One way to address the free-rider problem and increase the cost-effectiveness of energy efficiency subsidies is to target these measures to specific types of household[18].

Our results suggest that it will be difficult to encourage non-takers to invest in energy efficiency through targeted interventions that are based on observable characteristics. In Supplementary Methods 4, we present a linear probability model to analyse the differences in observable variables between takers and non-takers. Except for age, we found few observable variables that predict which households do and do not take up energy efficiency investments—hence, policymakers have few opportunities to use observable information.

Instead of targeting the entire group of non-takers, policymakers could target at a more granular level within takers and non-takers. Next, we analyse the correlation between key barriers/determinants and various variables that policymakers may observe or elicit, such as demographics, building characteristics, policy preferences and psychographics (that is, energy-related literacy, support for environmental causes and self-reported happiness).

We first examine the drivers of heterogeneity in barriers within the non-takers category. Table 5 presents three linear probability models, one for each main barrier elicited with the narratives. In these regression models, the dependent variable is a zero-one dummy variable that takes the value of 1 if the respondent mentions a particular barrier in the open-ended question.

Model column 1 shows the heterogeneity for the barrier that the building was considered to be already energy efficient. None of the coefficients for the traditional observable covariates is statistically significant.

For the financial barrier in model column 2, the respondents are less likely to be female; they have less education but report a higher rental value for their homes. In addition, these respondents were less likely to have used existing policies for retrofits.

Model column 3 examines the barrier of old age, with the tenant age being significant and positive, again consistent with the nature of the barrier in the narratives. Here, none of the policy variables showed statistical significance. Although several variables strongly correlate with each of the main barriers, the observable variables explain only a small portion of the total variance. The lack of statistically significant characteristics shows the difficulty of policy targeting for policymakers. Results suggest some heterogeneity in policy use; however, no particular policy preference is associated with any barrier,

Table 5 | Linear probability model on the major barriers

Barrier	Already efficient	Expensive	Old age
Log income	0.013 (0.107)	–0.129 (0.097)	–0.004 (0.058)
Age	0.000 (0.004)	–0.003 (0.003)	0.004** (0.002)
Female	0.014 (0.083)	–0.135* (0.075)	0.082* (0.045)
Children	–0.056 (0.084)	0.022 (0.076)	0.010 (0.046)
University degree	0.076 (0.073)	–0.118* (0.066)	0.045 (0.040)
Building age	–0.001 (0.002)	0.000 (0.001)	0.000 (0.001)
Floor size	0.001 (0.001)	–0.001 (0.001)	–0.001 (0.000)
Log rental value	0.018 (0.073)	0.140** (0.066)	–0.070* (0.040)
Energy literacy	0.075 (0.053)	0.036 (0.048)	0.028 (0.029)
Studied economics	0.021 (0.066)	0.036 (0.059)	0.004 (0.036)
Proficient in mathematics	–0.060 (0.065)	0.068 (0.059)	–0.018 (0.035)
Energy-saving score	–0.070 (0.043)	0.013 (0.039)	–0.014 (0.023)
Donated	0.064 (0.061)	–0.002 (0.055)	0.060* (0.033)
Happiness score	–0.012 (0.024)	0.005 (0.021)	0.006 (0.013)
Policy awareness	0.019 (0.025)	–0.016 (0.023)	–0.019 (0.014)
Policy usage	0.027 (0.031)	–0.086*** (0.028)	–0.003 (0.017)
Policy market	0.032 (0.060)	–0.003 (0.055)	–0.016 (0.033)
Policy behavioural	0.022 (0.064)	–0.026 (0.058)	0.013 (0.035)
Policy non-market	0.112 (0.078)	0.024 (0.070)	–0.040 (0.042)
Intercept	0.001 (1.049)	0.613 (0.950)	0.446 (0.571)
Number of observations	297	297	297
R^2	0.056	0.092	0.095

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 1 if the respondent mentioned it in the open-ended answer and zero otherwise. Standard errors are in parentheses. The asterisks indicate P values from a two-sided t-test with the null hypothesis being that the coefficient is equal to zero: *$P < 0.1$; **$P < 0.05$; ***$P < 0.01$.
R^2, coefficient of determination.

suggesting that policy preferences follow a uniform distribution over the different barriers.

Next, we analyse the heterogeneity within the takers. Table 6 shows four linear probability models, one for each major determinant of retrofitting.

Model column 1 shows the result for the determinant of replacement of existing parts (that is, obsolescence). Tenant age is significantly and negatively correlated with this determinant, possibly due to the shorter time that older tenants expect to live in their homes. As expected, building age positively correlates with this determinant, as older buildings are more likely to need repairs than newer ones. Respondents who renovated to replace broken parts were less likely to know about available subsidies. In addition, the replacement determinant is associated with a policy preference for market-oriented and non-market-oriented policies; however, these respondents have no preference for behavioural policies.

Model column 2 presents the heterogeneity for the financial determinant of saving money. Interestingly, a university degree and previous donations to environmental organizations are negatively associated with this determinant. Not surprisingly, respondents who renovate to save money show higher policy usage (but not a higher level of awareness) and a preference for market-based policies (that is, subsidies).

Model column 3 displays the comfort determinant. Only a positive correlation for building age is strongly associated with this

determinant. Moreover, this determinant does not seem to show statistically significant heterogeneity in terms of their awareness or policy preferences but a higher level of policy use.

Model column 4 shows heterogeneity among respondents who renovate out of environmental concerns. Both income and previous donations to environmental organizations have a strong and statistically significant association with this determinant. Similar to respondents who renovate to save money, environmental concerns are associated with higher levels of policy use but not with more awareness. Moreover, environmental motivations to renovate are strongly associated with behavioural policy preferences and are weakly negatively associated with non-market policies.

Policymakers could target their policies if barriers and determinants correlate with some well-defined household types. Therefore, we group respondents based on the previous heterogeneity analysis.

Among the non-takers, two types of homeowner emerge. Those who do not renovate because they already consider their home to be energy efficient, and those who face financial constraints. The first type of homeowner is not significantly different from other non-takers; the second type of homeowner tends to be male and has a lower level of education. Overall, policy preferences play a minor role in explaining the barriers of the non-takers, which limits the scope for policy targeting.

For the takers, the main aim of policymakers is to identify homeowners who are renovating to replace broken parts (replacers) and encourage them to plan their retrofits earlier. In addition, policymakers can target respondents who do not see energy-efficient retrofits as a financial opportunity (money savers): households that do not belong to the money-savers group may not be aware of the long-term cost savings from improved energy efficiency.

Compared with other takers, replacers are relatively young and live in older buildings. In addition, replacers are less informed about existing policies and prefer policies that involve less bureaucracy and higher subsidies. They are consistent with a pragmatic decision-maker whose primary motivation is to replace broken parts of the house (which is often a necessity). Even though policy awareness is low for this group, respondents prefer less bureaucracy and higher subsidies rather than more information on policies. The replacers are thus looking for ways to facilitate the retrofitting process, which policymakers can achieve by providing more information and reducing bureaucracy.

Money savers are less educated and less likely to donate to environmental organizations. They are more inclined than other groups to use existing policies for retrofits and favour market-based policies (such as subsidies). It is, therefore, difficult for policymakers to target specific policies.

Aside from age and income, there are few observable variables that enable policymakers to target non-takers. Most homeowners who renovate do so to fix broken parts; this can be seen as free riding because these homeowners would have renovated regardless of receiving subsidies. By contrast, replacers are less aware of policies and favour reducing bureaucratic hurdles. Policymakers could therefore limit free riding by targeting the group of replacers, especially by improving the institutional framework for energy-efficient retrofits—through reducing bureaucracy and increasing information campaigns.

Discussion and conclusions

Using an emerging approach based on open-ended survey questions, we elicited the narratives of homeowners about their renovation behaviour. Our key finding is that energy efficiency investments are highly opportunistic. Non-takers believe, rightfully or not, that there are few opportunities for energy efficiency in their homes. Whereas financial considerations are a major barrier to renovation, we found that homeowners who do make energy-efficient retrofits do not see them as financial opportunities. Strikingly, most homeowners delay renovations until they need to replace broken building components. In such a setting, the renovations would probably have been carried out

Table 6 | Linear probability model on the major determinants

Determinant	Replacement	Save money	Comfort	Environmental
Log income	−0.077 (0.054)	0.008 (0.050)	0.059 (0.048)	0.139*** (0.050)
Age	−0.005*** (0.002)	−0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Female	0.041 (0.039)	−0.012 (0.036)	−0.045 (0.035)	0.025 (0.036)
Children	−0.018 (0.039)	0.040 (0.037)	−0.027 (0.035)	−0.027 (0.036)
University degree	0.023 (0.035)	−0.062* (0.033)	−0.037 (0.031)	0.011 (0.032)
Building age	0.002** (0.001)	0.000 (0.001)	0.002*** (0.001)	0.000 (0.001)
Floor size	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Log rental value	0.010 (0.035)	−0.007 (0.033)	−0.021 (0.031)	−0.052 (0.032)
Energy literacy	0.046 (0.029)	−0.002 (0.027)	0.016 (0.026)	−0.009 (0.027)
Studied economics	0.021 (0.032)	0.039 (0.030)	0.050* (0.029)	0.000 (0.030)
Proficient in mathematics	−0.008 (0.031)	0.023 (0.029)	0.038 (0.028)	0.038 (0.029)
Energy-saving score	0.000 (0.022)	−0.015 (0.020)	0.022 (0.020)	0.024 (0.020)
Donated	−0.035 (0.030)	−0.051* (0.028)	0.003 (0.027)	0.082*** (0.028)
Happiness score	−0.007 (0.011)	0.008 (0.010)	−0.001 (0.010)	0.003 (0.010)
Policy awareness	−0.037** (0.015)	0.006 (0.014)	0.005 (0.014)	0.022 (0.014)
Policy usage	−0.001 (0.013)	0.022* (0.012)	0.025** (0.012)	0.020* (0.012)
Policy market	0.083*** (0.030)	0.075*** (0.028)	−0.006 (0.027)	0.021 (0.028)
Policy behavioural	0.019 (0.032)	−0.021 (0.030)	−0.003 (0.029)	0.111*** (0.030)
Policy non-market	0.100*** (0.036)	−0.028 (0.034)	0.035 (0.032)	−0.057* (0.033)
Intercept	1.129** (0.518)	0.302 (0.486)	−0.380 (0.465)	−0.802* (0.482)
Number of observations	1,158	1,158	1,158	1,158
R^2	0.047	0.024	0.026	0.047

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 1 if the respondent mentioned it in the open-text answer and zero otherwise. Standard errors are in parentheses. The asterisks indicate P values from a two-sided t-test with the null hypothesis that the coefficient is equal to zero: *$P < 0.1$; **$P < 0.05$; ***$P < 0.01$.

even without subsidies, suggesting that free riding plays an important role in Switzerland's generous subsidy programmes.

Our approach consists of eliciting narratives in surveys that are based on open-ended questions. Narratives, combined with NLP methods, offer a powerful means of identifying and assessing the key barriers and determinants of households' retrofitting decisions. Notably, narrative-based rankings differ from rankings using closed-ended questions. In the closed-ended questions, respondents indicated that the primary reason for performing a renovation was to save money and address environmental concerns. By contrast, the narratives showed a different picture, where the main reason was to replace old parts of the building. In addition, the importance of several co-benefits of energy efficiency decreased when using closed-ended questions, especially regarding comfort gain and environmental concerns.

Furthermore, we found a generally low level of policy awareness and, among takers, a low policy usage. In particular, homeowners who have renovated to repair broken parts had a lower policy awareness, which provides an opportunity for information campaigns. Unsurprisingly, most homeowners supported higher subsidies; however, subsidies in Switzerland are already generous. The second and third most favoured policies were related to improving the institutions related to retrofits, that is, reducing bureaucracy and providing more information for homeowners.

Most of the characteristics that influence the decision of respondents' to carry out retrofits are challenging to target for policymakers. For example, the main difference between takers and non-takers was the older age of the non-takers. It is difficult to imagine a means-tested subsidy programme using age as a criterion. However, a programme with a 'senior discount' and a more practical approach for older households may be politically feasible. We found no other demographic variables that policymakers could use to target policies besides age. Moreover, there are no significant differences in policy preferences.

Since the primary determinant of energy-efficient retrofits was to repair defective parts, more generous subsidies may have a small impact on the retrofitting decision and encourage free riding. Most homeowners renovate out of necessity rather than planning their retrofits for the long term. However, takers are often unaware of existing measures and would welcome a reduction in the bureaucratic burden of retrofitting. Instead of purely monetary incentives, an effective policy should also consider institutional factors, such as bureaucratic burden and accessibility of information.

In the future, rapid improvements in artificial intelligence and chatbots will make text-based interactions more common, providing fertile ground for researchers and policymakers. Respondents' narratives about energy efficiency offer precise explanations for their preferences and often differ from traditional multiple-choice answers. Future research can analyse these differences and uncover new insights for policymakers, which can help to close the energy efficiency gap.

Methods
Related literature
Our study relates primarily to the literature on the energy efficiency gap that has developed a taxonomy for barriers and determinants of energy efficiency investments. It also relates to the broader literature in energy economics studying technology adoption and policy preferences using stated and revealed preferences. We briefly review these two strands of literature to put our study into context.

Taxonomy of barriers and determinants. In the literature, there are several frameworks that have identified and categorized the barriers and determinants of energy efficiency investments. One particularly influential taxonomy distinguishes between three different perspectives: economic, behavioural and organizational[3].

The economic perspective considers rational utility-maximizing agents as the benchmark to understand the choices of agents regarding the adoption of energy-efficient technologies. The behavioural perspective departs from this purely neo-classical framework: it considers different manifestations of bounded rationality, which have also been referred to in the literature as behavioural failures[19] or internalities[20]. Finally, the organizational perspective considers the role of institutions with which the agents interact. These could be institutions that governments have little ability to transform, such as values and culture, or others over which they have considerable influence, such as fiscal, competition and regulatory policies.

Although this taxonomy is helpful in navigating the different explanations of the source of the energy efficiency gap, effective policy design needs a more precise categorization. Based on the economic, behavioural and organizational perspectives from the literature[3], our taxonomy thus distinguishes between behavioural, financial, non-market and market barriers.

Behavioural barriers. The empirical research has focused mainly on determining whether or not households correctly perceive the energy-savings component of the net investment costs. Some of these inefficiencies may be behavioural. Hence, we consider them to be internalities[20], such as inattention and biased beliefs about energy prices, to name but a few. These behavioural barriers can, however, be confounded with neo-classic market barriers such as different access to credit or time-discounting preferences.

Financial barriers. The second set of economic barriers focuses on the role of external financially related factors. Energy prices may be too low, investment costs too high, subsidies may not be generous enough, and various financial distortions could exist. Formally, in a household-investment framework, the role of financially related barriers operates through the price variables and the preference parameters that dictate the price sensitivity.

Non-market barriers. The third subcategory of economic barriers consists of the non-market components of the investment, which includes the various co-benefits and hidden costs of such investments. The literature has pointed out that specific co-benefits can be important, which researchers investigated mainly using contingent valuation methods[21,22]. External organizational constraint factors can affect the hassle costs, which we also consider to be a form of a non-market barrier.

Market barriers. Finally, classic market barriers are typical components of well-functioning markets. These barriers can arise due to heterogeneity in the building stock, technologies and preferences. Understanding this heterogeneity can enable policymakers to target energy efficiency policies to increase their cost-effectiveness[18]. Researchers have recognized this point for a long time, and several studies have investigated different dimensions of heterogeneity in the decision to adopt energy-efficient technologies, for example, in ref. 23.

Apart from these different types of barrier, there are also market failures at the source of the energy efficiency gap. Externalities associated with energy systems that are not systematically taken into account, asymmetries of information and imperfect competition are three market failures that interact with the different barriers.

Our focus is on uncovering the barriers and determinants that arise at the level of the decision-makers. The market failures listed above are features of the entire market. For instance, pollution from energy production affects all market participants and may not be reflected in market prices. However, this externality does not arise from the individual decision-making process of households, which is the focus of our elicitation approach (for example, some households may invest in energy efficiency to profit from co-benefits such as comfort).

Methods to elicit barriers and determinants. Empirical researchers in the energy sector have used two types of approach to elicit the barriers and determinants of technology adoption: stated and revealed preferences.

With revealed-preference methods, the underlying determinants of economic decision-making are inferred from observed, real-life choices (for example, in ref. 24) or from experimental data (for example, refs. 25–27).

With stated-preference methods, analysts will typically construct a survey with well-defined options of barriers and determinants that they think households might face (for example, refs. 28–30); analysts can also construct hypothetical-choice situations to have a perfectly controlled environment from which they can infer underlying preferences (for example, refs. 31–35). These data, together with a model that provides the micro-foundations for mapping between preferences and observed choices, are used to infer preferences and, incidentally, particular barriers and determinants.

Eliciting narratives using open-ended survey questions is a third and complementary approach. Researchers have used this question type since the 1940s, although not systematically[36]. With small samples, however, open-ended questions are used regularly, for example, in refs. 37–40. On a larger scale, this approach should yield very noisy, hard-to-interpret qualitative data. However, researchers suggested revisiting the concept of open-ended survey questions[41]. Recent advances in NLP enabled us to turn narratives into quantifiable metrics to elicit proxies for household preferences and market barriers. Using this approach, open-ended questions have become increasingly popular to elicit opinions on critical societal issues such as immigration[42], climate change[43] and macroeconomic shocks[44], as well as to elicit policy preferences[12,45].

Narratives offer two advantages over typical closed-ended survey questions in stated-preference studies. First, at the respondent level, narratives elicit a narrower but more specific set of barriers and determinants. In particular, respondents tend to focus on a few but more important topics to explain their decision-making. Second, at the population-wide level, compared with closed-ended questions, narratives uncover a broader set of barriers and determinants.

Most research uses NLP with already existing text data, for example, public comments[46], newspaper articles[47,48], congressional speech[49] or Twitter texts[50,51]. By contrast, open-ended questions enable the collection of new text data that are specific to the research question. However, owing to the smaller sample size, shorter texts and specialized vocabulary, open-ended responses require a different methodological approach, as we will outline in Supplementary Methods 1.

Important findings from other studies. Two recent reviews[11,52] aim to consolidate the empirical findings related to the barriers and determinants of energy efficiency investments. The first of these reviews is the most relevant study for us as it explicitly focuses exclusively on empirical investigations and discusses the methods used in detail[11]: it covers 26 empirical studies, where most of them used standardized survey procedures with predetermined answers (that is, closed-ended questions) and only one used semi-standardized qualitative interviews.

Overall, the review identified 167 different explanatory variables across the 26 studies, although their main conclusions are humbling[11]. Despite extensive empirical work, only a few robust findings emerge. Our findings are consistent with some patterns they identified. For instance, the positive relationship between income and energy efficiency investments is present in several studies and ours. Age also tends to predict a lower take-up rate. Higher financial and energy literacy are positively associated with such investments. Comfort, when measured, is also positively correlated with investments. Finally, policies, such as audits and energy-provision programmes, tend to be associated with higher take-up rates.

However, the authors include several caveats regarding their findings[11]. They conclude that the field is not mature enough to obtain a general consensus on the main barriers and determinants of energy efficiency. As they pointed out, one main problem is the need for comparable elicitation procedures that truly uncover the thought processes of decision-makers. Until now, each study has used its own structured questions and focused on easy-to-measure variables. Not only does this approach make it hard to compare studies and draw general conclusions, but it is also prone to biases of the authors based on their own credences. Moreover, most of these studies rely on cross-sectional regression analyses that look for associations between behaviours and the households' characteristics, beliefs and other constructed variables. This approach only indirectly infers households' thought processes and the barriers to and determinants of their investment.

Our proposed elicitation approach based on open-ended questions combined with NLP aims to address these shortcomings by being scalable and comparable across domains and contexts, while directly uncovering the thought processes we ultimately want to learn about.

Survey procedure

The first survey module collected information on past and future energy efficiency-related behaviours: whether households had performed or intended to perform retrofits, as well as the types of retrofit. We used these different behaviours to distinguish between takers and non-takers of energy efficiency investments.

The goal of the remaining modules was to determine the components of the households' decisions that influenced these behaviours. One of the most important modules focused on different barriers and determinants. To elicit these components, we used open-ended questions, which provided narratives about specific aspects of the decision-making process. We also used structured closed-ended questions that closely mirrored the open-ended questions. Our goal was to provide a benchmark for open-ended questions.

Another module focused on preferences for different types of energy efficiency policies. Finally, the remaining modules elicited household and building characteristics, including some related to the decision-making process, such as financial and energy-related literacy. We used these variables to investigate heterogeneity along several dimensions.

To recruit participants, we collaborated with the Statistical Office of the Canton of Zurich. We sent personalized invitation letters via postal mail to a random sample of homeowners. The letter contained a short description of our research project and a link to an online survey. Household respondents had to type the link into a web browser to complete the survey using the software SurveyMonkey. To incentivize participation, respondents could win one of 100 gift certificates, each worth about US$200, in a lottery. We obtained informed consent from all participants.

We stratified the sample according to the following rules: single-family homes, the year of construction before 1990, and 50% with renovation permits during the past 5 years; furthermore, we stratified the tenant age and the number of tenants by considering buckets for both variables. We also stratified the sample to target homeowners who adopted the primary certification for energy-efficient buildings in Switzerland, that is, Minergie certification.

In the Canton of Zurich, there are 127,950 single-family homes for the period of our survey; 10,737 had applied for renovation permits between 2014 and 2019. The Statistical Office of the Canton of Zurich sampled this population and sent out 16,700 letters on our behalf on 3 February 2020. A household member could complete the online survey until 13 March 2020. Of the 16,700 letters sent, the response rate was high: 3,471 respondents started the survey, which is a response rate of 20.8%. Furthermore, there was a completion rate of 82%, with an average survey completion time of 30 min.

Although our sampling strategy targeted a population of homeowners of single-family houses, a small number of respondents did

not fall into this category. There were 161 (renting) tenants and a small number of respondents living in an apartment ($n = 23$). We excluded those observations from our analysis.

Sample composition: classifying household types. We used the past and intended energy-efficiency-related behaviours to classify households into two broad segments. First, we distinguished homeowners who adopted the Swiss energy efficiency certification for buildings (Minergie). Our stratified sampling strategy ensured that we observed a large number of those households ($n = 524$). We used these households for a separate study and therefore omitted these observations for the present analysis. Second, we distinguished households depending on whether they performed an energy efficiency retrofit in the past 5 years or planned to do so within the next 5 years. Furthermore, we used only answers from respondents who answered the open-ended question on the barriers/determinants to retrofitting. Based on this criterion, 2,187 households fell into two mutually exclusive categories:

- Non-takers: households that had not performed energy efficiency retrofits in the past and were not planning to do so in the future (461 observations, 21% of the sample).
- Takers: households that either performed energy efficiency retrofits in the past 5 years or who planned to perform at least one in the next 5 years (1,726 observations, 79% of the sample).

Our sample is representative of homeowners in Zurich with respect to the respondent age and household size. The average age for homeowners in the Canton of Zurich is 57.2 years. The percentage of single-person households is 11.6% for the entire population (in our sample, this share is 12.5% for non-takers and 8% for takers). For the other variables, there exist no official statistics, and the homeowner population differs from the rest of the population. However, our response rate was high, which supports that our sample is representative.

Eliciting energy efficiency narratives

This section describes how we elicit energy efficiency narratives and turn them into quantitative variables. Our survey first contained a closed-ended question on renovation behaviour, followed by an open-ended question on the same topic. Both questions asked why respondents decided for or against an energy-efficient retrofit. This setting enables a comparison between the two question types. Separately, we elicited barriers from the non-takers and determinants from the takers. The structure of the questions was similar for the question about barriers and the question about determinants. Finally, at the end of the survey, we asked a second open-ended question on the policies that the respondents would favour to increase the number of retrofits. This last open-ended question did not have a closed-ended counterpart.

In the closed-ended questions, respondents had to choose among several options in a multiple-choice question. For barriers, we listed 17 potential barriers discussed in the literature on the energy efficiency gap. Those options were presented to non-takers, whom we asked to select all the barriers that were relevant to them. We use a similar format to elicit the determinants. We established a list of eight potential determinants, and the takers selected those that were important to their retrofit decision.

Providing predefined options in the closed-ended question could lead to an elicitation bias. The goal of the open-ended follow-up question was to provide the respondents with time to think carefully and to elicit a more nuanced response. Writing the text answer about the retrofit decision forces respondents to reflect on their decision, to consider the points from the preceding multiple-choice options again and to decide which factors influenced their decision. When designing the survey, our initial hypothesis was that the open-ended questions would induce the survey participants to discuss fewer, but also the most important barriers or determinants, to their renovation decision. Compared with the closed-ended questions, this format enabled us to discriminate

better between the different options, as well as to discover barriers and determinants we may not have thought of while designing the survey.

Our approach relates closely to deliberations in experiments. In these settings, participants make an initial choice, and subsequently reflect on and discuss their decision with other participants to revise their decision and select their true optimal choice[53]. In our case, the aim is that, in asking the respondents to answer the open-ended question, this implements deliberation in the survey and helps to identify true barriers and determinants.

In our design, the respondents could reflect on their choice but could not discuss it with other respondents. Instead, we leverage the different cognitive processes for closed- and open-ended responses, that is, recognition and recall. From a cognitive point of view, the process induced by an open-ended question is called recall. By contrast, responses from closed-ended questions are based on recognition, where a respondent identifies the correct answer among multiple options. The underlying processes needed to answer recognition-type questions are different. They may be less complex than the more individual task in answering recall questions[54,55].

Next, we describe in detail how we design the open-ended survey questions. We then discuss our NLP-based method to classify the text responses into topics. In the final subsection we also compare our method with other text-classification approaches.

Description of open-ended questions. In designing open-ended questions, it is first essential to provide context for the participants and indicate why we ask these questions. We thus structured our survey by initially presenting the following short introduction explaining the rationale for asking open-ended questions (as the survey was conducted in German, we present our own translation of the answers here):

"The reasons for energy efficiency retrofits are complex and different for each household. We would like to learn more about why you decided (or did not) to renovate. What was important to you? Were there alternatives? Your response will help us better understand how we can support energy efficiency retrofits."

After providing some context, we then asked the following question to elicit the determinants (or barriers) of energy efficiency retrofits:

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

At the end of the survey, in addition to barriers and determinants, we also extracted narratives about policy preferences. For that purpose, we presented the following 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 buildings and renovation."

As for the determinants/barriers, after providing some context, the open-ended question was introduced with the following introduction:

"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 open-ended questions worked very well. By inspecting a large number of responses, we found that the respondents provided meaningful answers. The length of the answers to the three open-ended questions varied, on average, between 19 and 24 words. For all questions, the standard deviation is slightly below the size of these averages, and some respondents wrote very long and detailed answers. The median number of words was between 12 and 21, and depending on the questions, the 90% percentile is between 44 and 47. The mean and median sentence length is between 1 and 2, with a standard deviation of 1.3–1.6, and the 90th percentile between 3 and 4 sentences. Hence, most respondents wrote less than the four sentences requested in the open-ended question. This result indicates that the requested length of between three and four sentences did not limit the respondents in the length of their answers.

The questions on barriers and determinants were mandatory for all non-Minergie participants. As previously mentioned, we asked only the non-takers about the barriers to renovate, and only the takers about their determinants to renovate. An overview of the summary statistics for the open-ended questions is shown in Supplementary Table 8. For these two open-ended questions, we observed an attrition rate of only 0.6% for the non-takers and 0.9% for the takers—that is, upon having to answer one of these particular questions, only 0.6% and 0.9% decided to stop the survey altogether. At the beginning of the survey, we asked three open-ended questions, which we did not analyse in this study. These questions focused on the sentimental value of the home. They asked the respondents to describe the elements of their home that are associated with positive emotions and with negative emotions, and to describe what they would buy if they won a lottery. These initial three open-ended questions had an attrition rate of 12%, 0.03% and 1.1%, respectively. The question on policy recommendations was not mandatory and was placed at the end of the survey. Furthermore, this question was presented to all respondents, including the Minergie subsample. We observed a higher but still low attrition rate of 8.5%. The self-selection of respondents is thus not a major concern.

In the following subsection, we describe our approach to classify the answers into topics, which we based on a method called 'keyword dictionaries'.

Semi-manual classification

The first step in the text classification consists of extracting the entire text corpus for each question. Supplementary Table 1 shows the ten most used words for each of the three questions. The table also shows the original German keywords, an English translation and their total frequency in the answers. For all three questions, the most frequent words are not very informative for explaining the text answers. Instead, the most common words refer to the question itself. For instance, the answers for the barriers contain most frequently the words 'renovation', 'energetic' and 'house'. The high prevalence of these words is not surprising, given that the question was the about energy efficiency renovations of houses.

For this reason, the most prominent words are of little help in identifying major topics in the text answers. We thus need to focus on words with a lower frequency. For example, 'expensive' and 'costs' are very informative in identifying barriers to retrofitting. In this particular case, these words indicate that financial reasons may be a barrier. One challenge is that the frequency of these words is relatively low, which makes it hard to systematically identify a specific topic. To identify topics in a robust manner, we propose a keyword approach: this consists of identifying a large number of keywords, with relatively low frequency, that map a response to a predefined topic.

The main idea of our method consists of classifying the text answers using a dictionary-based approach. Using this method, a set of keywords defines each topic: if an answer contains any of these predefined words, the algorithm will classify it into the respective topic. We included as many words as possible in the dictionaries to obtain a precise classification of the topics. For this reason, we considered all words from the open-ended questions as potential keywords for the classification.

The following approach enables us to create dictionaries with a large number of words and with high precision. The classification method proceeds via three steps: pre-processing, clustering and topic extraction. First, the pre-processing step reduces the dimensionality of the text and adds to each keyword its word embedding, which describes the semantic distance from other words. In the clustering step, we cluster the keywords into groups based on their semantic similarity. In the subsequent topic extraction, we manually build a dictionary based on the clusters from the second step.

Pre-processing. The goal of the pre-processing step is to reduce the number of words of the text to facilitate the subsequent clustering, followed by the topic extraction. We first extracted all the words from the respondents' answers and transformed them into tokens. In this study, we use unigrams, that is, one word per token, which is considered sufficiently precise for NLP[56]. We then lemmatized all words using the spaCy open-source software library for NLP[57] (for example, 'better' will be transformed into 'good').

Next, we sorted all words according to their part of speech (POS). This step is necessary because many words in our text corpus have a relatively small semantic distance when they belong to the same POS. The underlying reason for this is that word embeddings rely on the 'distributional hypothesis', which means that words occurring in similar contexts are assigned a small semantic distance[58]. Therefore, word embeddings for words of the same POS can be relatively similar if these words occur in the same context. For instance, two adjectives may be considered similar, not because of their descriptions but because both are adjectives. We used the spaCy algorithm for POS tagging, marking all lemmatized words as nouns, adjectives, verbs or adverbs.

To complete the pre-processing, we used word embeddings to map each remaining word to a semantic-distance metric whenever possible. Word embeddings are matrices with a column of values for each word that indicate the relative semantic distance between words (for example, the distance between the words 'heating' and 'oil' is smaller than the distance between 'heating' and 'pencil'). We can calculate their semantic distance using the cosine similarity for any two words in our corpus. To construct such a matrix for our corpus of unique words, we mapped all the words present in the answers to the predefined German fastText word-embedding vectors[59]. Not all words could be mapped to the word embeddings and were omitted; most omitted words occurred only once or twice in the entire text corpus and have little value for the subsequent clustering and classification steps.

Clustering. Next, we clustered all unique nouns, adjectives, verbs and adverbs based on their semantic distance from the word embeddings. This step aims to create groups of similar words to facilitate the subsequent topic allocation of the keywords. As a result, for the final step of our approach, the researcher does not have to scan through an unsorted list of keywords and decide which topic each keyword belongs to. Instead, it becomes possible to analyse groups of 10–40 similar words and decide which of these words belongs to a specific topic.

To implement the clustering, we consider that words of the same POS tend to have a small semantic distance due to the word-embeddings construction. To avoid the influence of the POS in clustering words, we cluster words separately according to their POS. We use k-means clustering, which enables us to use the cosine-similarity metric with the word-embedding data; this k-means clustering is an unsupervised clustering method that groups data into a fixed number of clusters. Finally, it should be noted that the clustering step is not sensitive to either the number of clusters or to a specific clustering algorithm, as long as the resulting groups are sufficiently manageable to be read by a researcher.

Topic extraction. In the final step, we extracted the topics. We assigned each word, when possible, to one of the existing topics from the corresponding closed-ended question. Each word could belong only to a single topic. This step was not automated but was performed manually. During the topic-extraction process, we also discovered additional topics, which we then added to the list of predefined topics for the closed-ended question. Of all the unique words, we could assign between 15% and 20% of the words to a topic.

After creating the dictionary with the lemmatized words, we searched the non-lemmatized text for words for which the lemma was contained in our dictionary, and added these words. For example, the initial dictionary contained the word 'cost', which is the lemma of 'costs', which we added in this last step. Adding the non-lemmatized words enabled the dictionary on the non-lemmatized text to be used.

Finally, after assigning words to topics, we labelled the text answers by automatically searching each answer for the presence of the words that define a topic. We applied the same approach to the question with barriers but identified different keywords and topics. An overview of the most important words that define the topics can be found in Supplementary Table 2 for the two major barriers, in Supplementary Table 3 for the four major determinants and in Supplementary Table 4 for the policy preferences. All tables show the original German words, their English translations and their total frequency.

Using the above approach, we ranked the barriers and determinants to energy efficiency investments by tabulating the topic frequency. We then contrasted the rankings with those obtained from the closed-ended questions. Compared with the closed-ended questions, the answers from the open-ended questions were only sometimes consistent: a respondent could check the box for a certain topic but not mention it in the text answer and vice versa. Possibly, the text classification was inaccurate. As a robustness check, we checked all inconsistent answers manually for the major topics. In most cases, the initial classification was correct, meaning that the topic shares changed only marginally. We present a more detailed analysis of the consistency in Supplementary Methods 2. Note that the survey did not aim to specifically elicit the underlying reasons for the inconsistencies between open- and closed-ended questions. A study regarding opinions on US trade policies observes a similar inconsistency between closed- and open-ended survey responses[60].

Supplementary Table 9 provides an overview of how each pre-processing step reduced the number of words. Initially, all unique words in the corpus were considered to be potential keywords. In the first pre-processing step, selections were made based only on the POS and only nouns, verbs, adjectives and adverbs were retained. Next, we lemmatized the words and dropped words with three characters or less (except for 'CO_2' and 'old' with the determinants and 'PV' in place of photovoltaic with the policy recommendations).

For the barriers and determinants, we did not have a limit on the frequency of the words, but for the policy recommendations, we only selected words that occurred at least twice. This slight change in procedure was because we did not have a predefined list of topics from a closed-ended question. Therefore, we did not aim for maximum precision and could improve the topic extraction considerably. Selecting words with a frequency equal to or higher than two substantially 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 was not necessary. Working with a corpus with a lower dimensionality also facilitated the initial discovery and definition of topics.

The second pre-processing step mapped the remaining keywords to the word embedding whenever possible, further reducing the number of words. Most words that could not be mapped occurred just once or twice in the text corpus, which suggests that these words would not influence the classification importantly. Overall, our pre-processing reduced the number of unique words by 50–75%.

On the basis of the lists of words for each open-ended question, we constructed the initial dictionaries. The dictionaries contain between 15% and 20% of the remaining words after pre-processing. In a final step, we re-introduced the non-lemmatized words, increasing the size of the dictionaries by 7–21%.

Validation and comparison with other approaches. We compared our method with both manual classification and machine-learning-based approaches using two types of model validation: the first based on human coding and the second based on the semantic-distance measure from the word embedding.

Our first validation consisted of comparing our model against two human coders, similar to approaches in the literature[56]. The human coders are consistent with our model for a subsample of our data. Moreover, the consistency between the human coders was only marginally higher than their consistency with our dictionary method.

The second validation is based on the semantic distance obtained from the word embeddings, and is similar to the metrics used for unsupervised topic models to find the optimal topics[61]. We calculate a 'quality' measure that describes the semantic similarity of words within a topic compared with words of other topics. The underlying idea is that the keywords defining a topic should be closely related compared with keywords from other topics. The quality measure is the fraction of the intra-topic coherence and the inter-topic similarity. For all topics, we found that the quality measure was greater than unity, which means that the intra-topic coherence is larger than the inter-topic similarity. A detailed description of both validation approaches can be found in Supplementary Methods 1.

Our validation with human coding is similar to qualitative methods that researchers have relied on to classify open-ended survey responses before advances in NLP. Our semi-automated approach has, however, several advantages over manual approaches. It is scalable, which means that the additional time to construct a dictionary decreases with more answers. Future surveys that contain the same open-ended question can use the same dictionary to analyse the responses. By contrast, a manual classification must start from scratch for a new survey with the same open-ended question. The more topics a researcher must consider for manual classification, the more challenging the classification becomes. Our keyword approach is less affected by the number of topics because the researcher selects keywords from a pre-compiled list of similar words.

To classify text into topics, researchers differentiate between text-classification methods with known and unknown categories[56]. The most commonly used method for text analysis consists of topic models that detect categories in the text unknown to the researcher. Among these, latent Dirichlet allocation (LDA) is the most commonly used method[62].

LDA is a topic model that assumes the following text generation process: first, each document is generated by sampling topics from a topic distribution. Conditional on the sampled topics, words are sampled from each topic-word distribution. This unsupervised method is very efficient with large texts and without any additional covariates about the text.

However, some of the characteristics of LDA make it unsuitable for deriving the exact topics from open-ended survey responses; with short text samples and a small number of documents, LDA tends to perform poorly[63]. Moreover, if the researcher aims to recover many topics, LDA risks giving results that human readers cannot interpret[64]. Open-ended survey responses fall into that category because they consist of small samples and short texts. For this reason, LDA and similar topic models may not give the exact distributions of topics from open-ended survey responses[65]. Other topic models that are based on LDA share these characteristics, for example, the Structural Topic Model[66], LDA2Vec[67] or Top2Vec[68].

Finally, because LDA is an unsupervised model, it does not consider the information from the open-ended questions that the respondents provide. For instance, when asked why they performed a retrofit, some respondents answered that they exchanged their gas heating because it was no longer functioning. Others answered that they installed new windows because the old ones were broken. In both cases, the relevant topic that answers the question would be that these respondents renovated because they had to exchange broken parts of the building. An unsupervised model, however, cannot consider this information and will instead focus on the parts of the response that describe which building parts were exchanged, that is, the gas heating or the windows.

Classification methods with unknown categories, such as LDA, can be differentiated from methods where the categories are known to the researcher[56]. Our method to classify the open-ended responses is based on a dictionary approach, where the categories are known to the researcher. Dictionary approaches rely on pre-chosen keywords that can describe sentiments or topics. Consequently, a text containing these keywords will be associated with a specific sentiment or a topic. By contrast, machine-learning methods use the linguistic features of a text[69]. Because we use machine-learning methods to construct the dictionary, we consider our method to be a hybrid approach between the two[69].

In contrast to machine-learning-based methods, our dictionary approach provides an exact distribution of topics. Moreover, multiple topics can be associated with text answers with the same high precision. Furthermore, the categories are transparent because they depend on an accessible list of keywords. Finally, once a dictionary has been compiled, it can be used for a future survey that relies on the same open-ended question. Overall, our approach thus provides a way to move forward and address the inconsistencies in explaining renovation decisions[11]. We elicit the barriers and determinants of energy efficiency investments using a scalable and easy-to-replicate questionnaire across a large set of domains and contexts.

Ethics statement

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability
The data that support the findings of this study are not openly available to ensure the anonymity of the survey participants. Upon reasonable request and with a data protection/security agreement, the data can be made available from M. Filippini at the Research Collection at ETH Zurich (mfilippini@ethz.ch). A response to the request will be given within 4 weeks.

Code availability
The code is available from the corresponding author upon request. A response to the request will be given within 4 weeks.

References

1. *2022 Report on the Achievement of the 2020 Energy Efficiency Targets* (European Commission, 2022).
2. Jaffe, A. B. & Stavins, R. N. The energy-efficiency gap. What does it mean? *Energy Policy* **22**, 804–810 (1994).
3. Sorrell, S., O'Malley, E., Schleich, J. & Scott, S. *The Economics of Energy Efficiency: Barriers to Cost-Effective Investment* (Edward Elgar Publishing, 2004).
4. Cagno, E., Worrell, E., Trianni, A. & Pugliese, G. A novel approach for barriers to industrial energy efficiency. *Renew. Sustain. Energy Rev.* **19**, 290–308 (2013).
5. Achtnicht, M. & Madlener, R. Factors influencing German house owners' preferences on energy retrofits. *Energy Policy* **68**, 254–263 (2014).
6. Schleich, J., Gassmann, X., Faure, C. & Meissner, T. Making the implicit explicit: a look inside the implicit discount rate. *Energy Policy* **97**, 321–331 (2016).

7.   Gerarden, T. D., Newell, R. G. & Stavins, R. N. Assessing the energy-efficiency gap. *J. Econ. Lit.* **55**, 1486–1525 (2017).

8.   Gillingham, K., Keyes, A. & Palmer, K. Advances in evaluating energy efficiency policies and programs. *Annu. Rev. Resour. Econ.* **10**, 511–532 (2018).

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

10. Christensen, P., Francisco, P., Myers, E. & Souza, M. Decomposing the wedge between projected and realized returns in energy efficiency programs. *Rev. Econ. Stat.* https://doi.org/10.1162/rest_a_01087 (2023).

11. Kastner, I. & Stern, P. C. Examining the decision-making processes behind household energy investments: a review. *Energy Res. Soc. Sci.* **10**, 72–89 (2015).

12. Ferrario, B. & Stantcheva, S. Eliciting people's first-order concerns: text analysis of open-ended survey questions. *AEA Pap. Proc.* **112**, 163–69 (2022).

13. White, K., Hardisty, D. J. & Habib, R. The elusive green consumer. *Harv. Bus. Rev.* **11**, 124–133 (2019).

14. Grösche, P. & Vance, C. Willingness to pay for energy conservation and free-ridership on subsidization: evidence from Germany. *Energy J.* https://doi.org/10.5547/ISSN0195-6574-EJ-Vol30-No2-7 (2009).

15. Boomhower, J. & Davis, L. W. A credible approach for measuring inframarginal participation in energy efficiency programs. *J. Public Econ.* **113**, 67–79 (2014).

16. Houde, S. & Aldy, J. E. Consumers' response to state energy efficient appliance rebate programs. *Am. Econ. J. Econ. Policy* **9**, 227–55 (2017).

17. Buettner, T. & Madzharova, B. *Promoting Sales of Energy Efficient Household Appliances: Outcomes and Cost Effectiveness of Rebate Programs* Working Paper No. 9048 (CESifo, 2021).

18. Allcott, H., Knittel, C. & Taubinsky, D. Tagging and targeting of energy efficiency subsidies. *Am. Econ. Rev.* **105**, 187–91 (2015).

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

20. Allcott, H. & Sunstein, C. R. Regulating internalities. *J. Policy Anal. Manage.* **34**, 698–705 (2015).

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

22. Ürge Vorsatz, D., Herrero, S. T., Dubash, N. K. & Lecocq, F. Measuring the co-benefits of climate change mitigation. *Annu. Rev. Environ. Resour.* **39**, 549–582 (2014).

23. Jakob, M. *The Drivers of and Barriers to Energy Efficiency in Renovation Decisions of Single-Family Home-Owners* Working Paper 56 (Centre for Energy Policy and Economics, Swiss Federal Institute of Technology, 2007).

24. Kahn, M. E., Kok, N. & Quigley, J. M. Carbon emissions from the commercial building sector: the role of climate, quality, and incentives. *J. Public Econ.* **113**, 1–12 (2014).

25. Allcott, H. & Greenstone, M. *Measuring the Welfare Effects of Residential Energy Efficiency Programs* Working Paper 23386 (National Bureau of Economic Research, 2017).

26. Gilbert, B. & Zivin, J. G. Dynamic salience with intermittent billing: evidence from smart electricity meters. *J. Econ. Behav. Organ.* **107**, 176–190 (2014).

27. Asensio, O. I. & Delmas, M. A. The dynamics of behavior change: evidence from energy conservation. *J. Econ. Behav. Organ.* **126**, 196–212 (2016).

28. Lee, K.-H. Drivers and barriers to energy efficiency management for sustainable development. *Sustain. Dev.* **23**, 16–25 (2015).

29. Trotta, G. Factors affecting energy-saving behaviours and energy efficiency investments in British households. *Energy Policy* **114**, 529–539 (2018).

30. Hrovatin, N. & Zorić, J. Determinants of energy-efficient home retrofits in Slovenia: the role of information sources. *Energy Build.* **180**, 42–50 (2018).

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

32. Alberini, A., Banfi, S. & Ramseier, C. Energy efficiency investments in the home: Swiss homeowners and expectations about future energy prices. *Energy J.* https://doi.org/10.5547/01956574.34.1.3 (2013).

33. Blasch, J., Filippini, M. & Kumar, N. Boundedly rational consumers, energy and investment literacy, and the display of information on household appliances. *Resour. Energy Econ.* **56**, 39–58 (2019).

34. Fischbacher, U., Schudy, S. & Teyssier, S. Heterogeneous preferences and investments in energy saving measures. *Resour. Energy Econ.* **63**, 101202 (2021).

35. Schleich, J., Gassmann, X., Meissner, T. & Faure, C. 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 Econ.* **80**, 377–393 (2019).

36. Krosnick, J. A. Survey research. *Annu. Rev. Psychol.* **50**, 537–567 (1999).

37. Kolko, J. & Neumark, D. Do some enterprise zones create jobs? *J. Policy Anal. Manage.* **29**, 5–38 (2010).

38. Langbein, L. I. Responsive bureaus, equity, and regulatory negotiation: an empirical view. *J. Policy Anal. Manage.* **21**, 449–465 (2002).

39. Kathlene, L. & Martin, J. A. Enhancing citizen participation: panel designs, perspectives, and policy formation. *J. Policy Anal. Manage.* **10**, 46–63 (1991).

40. Schneider, M., Elacqua, G. & Buckley, J. School choice in Chile: is it class or the classroom? *J. Policy Anal. Manage.* **25**, 577–601 (2006).

41. Roberts, M. E. et al. Structural topic models for open-ended survey responses. *Am. J. Pol. Sci.* **58**, 1064–1082 (2014).

42. Bursztyn, L., Haaland, I. K., Rao, A. & Roth, C. P. *Disguising Prejudice: Popular Rationales as Excuses for Intolerant Expression* Working Paper 27288 (National Bureau of Economic Research, 2020).

43. Tvinnereim, E. & Fløttum, K. Explaining topic prevalence in answers to open-ended survey questions about climate change. *Nat. Clim. Chang.* **5**, 744–747 (2015).

44. Andre, P., Haaland, I., Roth, C. & Wohlfart, J. *Narratives about the Macroeconomy* Working Paper 18/21 (Center for Economic Behavior and Inequality, 2021).

45. Stantcheva, S. *Understanding Economic Policies: What Do People Know and Learn?* (Harvard University, 2020).

46. Dokshin, F. A. Variation of public discourse about the impacts of fracking with geographic scale and proximity to proposed development. *Nat. Energy* **6**, 961–969 (2021).

47. Belmonte, A. & Rochlitz, M. The political economy of collective memories: evidence from Russian politics. *J. Econ. Behav. Organ.* **168**, 229–250 (2019).

48. Benesch, C., Loretz, S., Stadelmann, D. & Thomas, T. Media coverage and immigration worries: econometric evidence. *J. Econ. Behav. Organ.* **160**, 52–67 (2019).

49. Gentzkow, M., Shapiro, J. M. & Taddy, M. Measuring group differences in high-dimensional choices: method and application to congressional speech. *Econometrica* **87**, 1307–1340 (2019).

50. Baylis, P. Temperature and temperament: evidence from Twitter. *J. Public Econ.* **184**, 104161 (2020).

51. Morales, J. S. Legislating during war: conflict and politics in Colombia. *J. Public Econ.* **193**, 104325 (2021).

52. Liu, G., Ye, K., Tan, Y., Huang, Z. & Li, X. Factors influencing homeowners' housing renovation decision-making: towards a holistic understanding. *Energy Build.* **254**, 111568 (2022).

53. Lienhoop, N. & Völker, M. Preference refinement in deliberative choice experiments for ecosystem service valuation. *Land Econ.* **92**, 555–577 (2016).

54. Lane, S. Validity of high-stakes assessment: are students engaged in complex thinking? *Educ. Meas. Issues Pract.* **23**, 6–14 (2004).

55. Lindner, M. A., Strobel, B. & Köller, O. Multiple-Choice-Prüfungen an Hochschulen? *Z. Padagog. Psychol.* **29**, 133–149 (2015).

56. Grimmer, J. & Stewart, B. M. Text as data: the promise and pitfalls of automatic content analysis methods for political texts. *Polit. Anal.* **21**, 267–297 (2013).

57. Honnibal, M., Montani, I., Van Landeghem, S. & Boyd, A. spaCy: industrial-strength natural language processing in Python. *Zenodo* https://doi.org/10.5281/zenodo.1212303 (2020).

58. Goldberg, Y. & Levy, O. word2vec explained: deriving Mikolov et al.'s negative-sampling word-embedding method. Preprint at https://arxiv.org/abs/1402.3722 (2014).

59. Grave, E., Bojanowski, P., Gupta, P., Joulin, A. & Mikolov, T. Learning word vectors for 157 languages. In *Proc. Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (Eds Calzolari, N. et al.) (European Language Resources Association, 2018); https://aclanthology.org/L18-1550

60. Chang, A., Ferguson, T., Rothschild, J. & Page, B. I. *Ambivalence About International Trade in Open- and Closed-ended Survey Response* Working Paper No. 162 (Institute for New Economic Thinking, 2021).

61. Dieng, A. B., Ruiz, F. J. & Blei, D. M. Topic modeling in embedding spaces. *Trans. Assoc. Comput. Linguist.* **8**, 439–453 (2020).

62. Blei, D. M., Ng, A. Y. & Jordan, M. I. Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3**, 993–1022 (2003).

63. Tang, J., Meng, Z., Nguyen, X., Mei, Q. & Zhang, M. Understanding the limiting factors of topic modeling via posterior contraction analysis. In *Proc. 31st International Conference on Machine Learning* Vol. 32 (Eds Xing, E. P. & Jebara, T.) 190–198 (PMLR, 2014).

64. Nikolenko, S. I., Koltcov, S. & Koltsova, O. Topic modelling for qualitative studies. *J. Inf. Sci.* **43**, 88–102 (2017).

65. Pietsch, A.-S. & Lessmann, S. Topic modeling for analyzing open-ended survey responses. *J. Bus. Anal.* **1**, 93–116 (2018).

66. Roberts, M. E., Stewart, B. M. & Tingley, D. stm: an R package for structural topic models. *J. Stat. Softw.* **91**, 1–40 (2019).

67. Moody, C. E. Mixing Dirichlet topic models and word embeddings to make lda2vec. Preprint at https://arxiv.org/abs/1605.02019 (2016).

68. Angelov, D. Top2vec: distributed representations of topics. Preprint at https://arxiv.org/abs/2008.09470 (2020).

69. Maynard, D. & Funk, A. Automatic detection of political opinions in tweets. In *The Semantic Web: ESWC 2011 Workshops* (Eds García-Castro, R. et al.) 88–99 (Springer, 2012).

Author contributions

T.W. contributed to the project planning, survey design, data analysis, writing of the paper and methodological development. S.H. contributed to the project planning, survey design, data analysis and writing of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41560-023-01303-x.

Correspondence and requests for materials should be addressed to Tobias Wekhof.

Peer review information *Nature Energy* thanks Yueming (Lucy) Qiu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

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

nature portfolio

Corresponding author(s): Tobias Wekhof

Last updated by author(s): May 31, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The data collection consisted of an online survey, administered using SurveyMonkey.
Data analysis	The data was analyzed using R (version 4.0.5) and Python (version 3.8.5) with the spaCy package (version 2.3.4), gensim (version 3.8.3) and nltk (version 3.5).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data that support the findings of this study are not openly available to ensure the anonymity of the survey participants. Upon reasonable request and with a data protection/security agreement, the data can be made available from the Research Collection at ETH Zurich. For access, please contact Massimo Filippini (mfilippini@ethz.ch). A response to the request will be given within four weeks.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Survey participants self-declared their gender/sex (with the options female, male, and other). Note that survey was conducted in German, and the German word "Geschlecht" does not differentiate between gender and sex.
Population characteristics	See below.
Recruitment	Invitation to participate in an online survey with invitation letters sent by postal mail. Respondents with a high interst in energy efficiency and retrofits may had a higher response rate, which could lead to self selection. However, we included a large selection of co-variates in the statistical analysis to control for self-selection.
Ethics oversight	Approved by the Ethics Commission of ETH Zurich (EK 2019-N-130). Survey respondents gave informed consent to their participation.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Statistical analysis of quantitative survey data.
Research sample	Single-family homeowners in the Canton of Zurich. The sample is representative for the general homeowner population in the Canton of Zurich. We chose homeowners of single family houses because, unlike tenants of rented appartments, these homeowners can decide about renovations. The mean age was 61 years (non-Takers) and 58 years (Takers); 20.5% among non-Takers were female and 22.5% among Takers were female.
Sampling strategy	In the Canton of Zurich, there are 127,950 single-family homes; 10,737 of those have applied for renovation permits between 2014 and 2019. The Statistical Office of the Canton Zurich sampled this population and sent out 16,700 letters February 3, 2020, on our behalf. We stratified the sample according to the following rules: only single-family homes, year of construction prior to 1990, and 50% with renovation permits during the last five years; there was a wide range for tenant-age and the number of tenants. We also stratified to target homeowners who adopted the main certification for energy-efficient buildings in Switzerland, the "Minergie" certification. This sample size is sufficient because it allows to perform multivariate regression analysis. Based on our budget and expected response rate, we chose the obtained sample size.
Data collection	The Statistical Office of the Canton Zurich sampled this population and sent out 16,700 letters February 3, 2020, on our behalf. These letters contained a link to an online survey. Apart the respondents, no other person was involved in completing the survey. A household member could complete the online survey until March 13, 2020.
Timing	The data collection started on February 3, 2020 and ended on March 13, 2020.
Data exclusions	There were 161 (renting) tenants and a small number of respondents living in an apartment (n = 23). Those observations are excluded from our analysis.
Non-participation	Of the 16,700 letters sent, the response rate was high: 3,471 respondents started the survey, which is a response rate of 20.8%.
Randomization	Conditional on the stratification criteria, the Statistical Office of the Canton Zurich randomly selected 16,700 homeowners to receive an invitation letter to the survey.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	Antibodies
☒	Eukaryotic cell lines
☒	Palaeontology and archaeology
☒	Animals and other organisms
☒	Clinical data
☒	Dual use research of concern

Methods

n/a	Involved in the study
☒	ChIP-seq
☒	Flow cytometry
☒	MRI-based neuroimaging