Tackling energy poverty through behavioral change: A pilot study on social comparison interventions in social housing districts

Nicolas Caballero$^{1,2}$ and Nives Della Valle$^{1,3}$

$^1$Eurac Research, Institute for Renewable Energy
$^2$Department of Economics, Ruhr Bochum University
$^3$Joint Research Centre, European Commission

July 24, 2020

Abstract

Behavioral Economics has in recent years played a key role in informing the design of non-price interventions aimed at promoting energy conservation behaviors in residential areas. Some of the most influential contributions of the discipline in an applied setting have centered around the development of norm-based interventions. The success that these interventions have had in specific contexts presents an opportunity to utilize them as tools for tackling a prevalent type of poverty at the EU level: energy poverty. Recent contributions to the literature highlight the role of inefficient energy behavior as a significant driver of this particular type of poverty, which is characterised by an inability to afford the basic energy services necessary to guarantee a decent standard of living. Therefore, the effectiveness of norm-based interventions in vulnerable populations merits further investigation to determine whether this approach can suitably address the behavioral components of energy poverty by promoting efficient energy consumption and conservation efforts. This is particularly imperative when combined with retrofitting innovations, as it can help avoid negative behavioral responses often associated with the implementation of efficiency upgrades, such as rebound effects.

This study reports on a pilot conducted in an exemplary social housing context (located in Bolzano, Italy) with the aim to assess the effectiveness of social comparison interventions in energy vulnerable groups. Using a design that combines appeals to injunctive and descriptive norms embedded within In-Home Devices (IHD) in recently retrofitted homes, our objective is to set a basis for the assessment of effectiveness of these types of interventions in social housing populations. Our study seeks to provide useful methodological insights to policy makers on how to effectively design behaviourally informed interventions aimed at tackling energy poverty.

JEL Classification

C93; D03; D04; D12; D19; D91; Q40
1 Introduction

The field of Behavioral Economics has in recent years contributed greatly to informing the design of non-price interventions aimed at promoting energy conservation behaviors in residential areas (Andor and Fels, 2018). Increasingly more applied research has focused on uncovering the effectiveness of the provision of feedback on personal energy consumption in promoting conservation efforts. The main methodological contribution of the discipline in this regard has been in expanding the use of Randomised Controlled Trials (RCTs) (Banerjee and Duflo, 2009) to identify causal links between feedback intervention and reduced energy consumption. In addition in recent years, behavioral economics has expanded the implementation of feedback interventions by integrating appeals to social norms, that is by designing norm-based interventions that rely on social comparisons to encourage energy conservation.

Norm-based interventions in the energy domain refers to the provision of information to households about how their individual levels of energy consumption compare to that of a reference group of comparable households (Andor and Fels, 2018) (who ideally should be chosen to have lower levels of consumption, signalling to the target household a pre-existing societal norm towards lower levels of consumption). This approach is intrinsically linked to the provision of feedback on one’s consumption, and, at the same time, also introduces appeals to social norms through the provision of social information. The specific strand of norm-based interventions based on leveraging social comparisons are known in the literature as social comparison interventions.

The main objective of social comparison interventions in the energy domain is to promote behaviors that consume less energy, reducing therefore our dependence on CO2 emissions. Given our urgent need as world citizens to take action on climate change, as highlighted by the recent IPCC special report (Intergovernmental Panel on Climate Change (2018) Global Warming of 1.5C), all efforts promoting energy conservation and decarbonisation are of paramount importance. Given ample evidence that behavior is the main driver of variations in energy household energy use (Chen et al. (2015)), adapting our energy decisions must be a key component in our efforts as a society to mitigate global warming. Furthermore, behavioral interventions of this kind can help reduce the energy performance gap (Galvin, 2014), described as the difference between expected energy savings and actual energy savings. Particularly in what concerns retrofitted apartments, these interventions can help align behavior with the technological innovation in order to achieve greater levels of energy savings.

Despite the existence of a large literature testing norm-based interventions in the field, little attention has been given to understanding the effects of these interventions on specific demographics. Evidence by Khosrowpour et al. (2016) suggests the need for tailored interventions for households with different energy consumption patterns, highlighting the fact that different populations may react in different way to the provision of social information. The issue of intervention success on target demographics is of crucial importance particularly when testing interventions that could be especially beneficial to the energy vulnerable, a group characterised by sub-optimal energy behaviors and high energy needs (Kearns et al., 2019). This study aims at addressing this gap in the literature, by setting the basis for the evaluation of the effectiveness of norm-based interventions in a social housing context.

We present a methodology to study the effects of providing social information through In-Home Devices (IHDs) installed in recently retrofitted social houses, on energy consumption behaviors. More generally, we set the methodological basis for a more extensive evaluation of these interventions in a specific context of vulnerability. We begin to uncover how effective these interventions may be in these contexts by using an experimental approach that allows us to make causal in-
ferences, and make the case for a wider implementation of this methodology in social housing demographics to advise policy-making.

Previous research on these tenants (DellaValle et al., 2018) uncovered context-specific factors that affect the energy consumption patterns of this case study demographic. These include a high proportion of retirees and older individuals who are amongst the most energy vulnerable in our society. Furthermore, conditions of resource scarcity have proven to exacerbate behavioral biases such as myopia (Shah et al., 2012), and other context-specific factors, like stigma, play a crucial role in the development of inefficient energy behaviors within the most vulnerable populations (Hall et al., 2014; DellaValle, 2019). For these reasons, studying the effectiveness of a well-known behavioral intervention on the energy behavior of this specific population is of particular interest.

Our study also adds to the literature on interventions aimed at tackling energy poverty. Recent research (Kearns et al., 2019; DellaValle, 2019) has begun to pay attention to energy behaviors as an additional driver of energy poverty, recognizing factors such as use of household spaces and failure to adopt “adaptive thermal comfort” as significant determinants. Our study is unique in its emphasis on using behavior-change strategies to tackle energy poverty by encouraging the adoption of more efficient behaviors and, as a result, improving access to basic capabilities that derive from an efficient use of energy, such as good physical and mental health, education, and social integration for the most vulnerable (Day et al., 2016).

Our pilot is embedded within a wider EU project called Sinfonia, ran in social housing districts in Bolzano, South Tyrol (Italy). Consenting tenants were provided with IHDs to monitor and better control their energy use, and it is within this technology that we embed our intervention. In the context of recent retrofits, the wider goal of the IHD installation is to reduce the so-called energy performance gap (Galvin, 2014) by aligning the technological and behavioral components of efficiency improvements, and ensuring that the increased energy efficiency of the dwelling is matched with the adoption of more efficient behaviors. It is important to note that in this paper, we focus only on the effects of providing social information, not the effects of providing the IHD technology in general.

At the moment, we only report from one of the investigated districts and only in an initial subset of 13 apartments (housing 28 occupants) which were the first to receive the intervention. We plan to investigate the effects of our intervention in more districts and over a wider period of time in a follow-up study, as well as enlarging our scope to capture the effect on thermal energy demand as well. In this study, our aim is to provide useful methodological insights to policy makers on how to design complementary interventions aimed at tackling energy poverty. In particular, we aim to provide insights useful not only to integrate behaviorally informed interventions in social housing retrofitting projects, but also to account for contextual features in vulnerable demographics.

The results from our pilot seem to tentatively suggest that a one-size-fits-all application of a social comparison intervention in a vulnerable context is ineffective. Households receiving the normative appeals do not consume significantly less electricity than those households not receiving it during the first 3 months of implementation. In fact, when we consider the average electrical consumption across all household over the relevant time period, households in the former group consume slightly more, although the difference is statistically insignificant. This may be due to a variety of reasons, including the psychological impacts of scarcity (DellaValle and Zubaryeva, 2019), an initially very low level of energy consumption pre-intervention (Andor et al., 2018), or the limited salience of the descriptive information together with a negative perception of the injunctive information. All these potential mechanisms are discussed in the Discussion section of this paper, where we point that future research could further explore the precise mechanisms that limit the effectiveness of the intervention in a social housing context.

The remainder of the paper is structured as follows: Section 2 details the theoretical framework
we position our pilot study in, with particular emphasis on why social housing tenants make for an interesting and important target group to investigate in the context of energy efficiency behavior-change interventions. Section 3 introduces a methodology for the application of social comparison interventions in social housing districts and explains the context and application of our pilot intervention. Section 4 presents the results from our pilot intervention, with a particular focus on the adoption of a suitable analytical accounts for important household characteristics and preferences. Section 5 discusses the results, limitations of the study as well as directions for future research. Finally, Section 6 provides preliminary conclusions.

2 Theoretical background

2.1 Norm-based interventions

Norm-based interventions have proven amongst the most successful non-price interventions to achieve behavior change in applied settings. In a variety of pro-environmental domains, the provision of social information has proved an effective tool in shifting preferences towards more sustainable behaviors. In the domain of recycling (Schultz (1999)), towel reuse (Schultz et al. (2008)), household water use (Ferraro et al. (2011)), and crucially energy use (Allcott (2011)), norm-based interventions have been proven to affect both intention and actual behavior in the field.

The psychological processes by which the provision of social information affects individual behavior is still a matter of debate in the literature, but prevailing research seems to emphasize the effect that normative appeals have on our empirical expectations. When these expectations on other people’s behavior condition our own behavior, the resulting behavioral pattern is described as a ‘descriptive norm’ (Bicchieri (2005)). As noted by Bicchieri and Dimant (2019), while the term ‘descriptive norm’ is widely used in the psychological literature to mean a perception of what is commonly done, it is important to clarify that descriptive norms relate to interdependent behaviors, or those behaviors where motivation to undertake is dependent on a person’s beliefs of what is commonly done. Our expectations based on unconditional (shared) behavior, therefore, are distinct from descriptive norms (such as our expectation that people will wear an umbrella when it is raining). According to these expectations, people may wish to stick to descriptive norms for fear of social disapproval, or seeking social esteem (Farrow et al. (2017a)). Eventually, they condition their own behavior based on the empirical observations of others’ behavior (Bicchieri (2005)).

It is also important to discern between descriptive norms and injunctive norms. While descriptive norms relate to behavior motivated by empirical expectations on how people behave, injunctive norms are behavioral patterns that are conditional on our perceptions of what is perceived to be desirable or approved from our peers (therefore, like in the case of descriptive norms, also being interdependent behaviors). The key difference are the relevant underlying expectations, whether they are related to what other people are doing or what other people believe ”ought” to be done (Bicchieri and Dimant (2019)).

In the context of norm-based interventions therefore, at least two things need to be clearly outlined before designing an intervention. First, we need to diagnose the targeted behavior, whether it is conditional on our expectations of others or not (interdependent or independent) (Bicchieri and Dimant (2019)). Assuming the targeted behavior is interdependent, we then need to define what expectations to target in order to achieve the desired behavioral change, whether expectations on

\[i.e.: \text{interventions relying on social influence}\]

\[\text{For a recent review of the literature : Farrow et al. (2017a); Andor and Fels (2018)}\]

\[i.e.: \text{how we expect other people to behave}\]
what people do or expectations on what people think is right, therefore appealing to descriptive norms or injunctive norms, respectively (Bicchieri and Dimant, 2019). Here, the answer is likely to be highly dependent on the context, but at least in the energy domain there is ample research that supports appealing to both of these norms simultaneously when designing an intervention, as explained in the following section.

2.2 Norm-based intervention in the Energy domain

Norm-based interventions in the energy domain for the most part rely on allowing energy users to compare their consumption levels with other users, therefore they can be classified as social comparison interventions. Some ambiguity in terminology exists in the literature regarding the use of the term social comparison interventions compared to norm-based interventions. This is precisely because in most norm-based interventions, social comparison is the mechanism by which researchers and practitioners try to leverage social influence and obtain a desired behavior. For the purposes of our paper, social comparison interventions are taken to be a subset of applications within a wider set of norm-based interventions.

Social comparison interventions in the energy domain refer to the provision of information to households about how their individual levels of energy consumption compare to that of a reference group of comparable households (Andor and Fels, 2018). This approach is intrinsically linked to the provision of feedback on one’s consumption, and, at the same time, also introduces appeals to norms through the provision of social information. As outlined in the theoretical framework, the use of a norm-based intervention to target energy behaviors assumes that at least a part of people's level of energy consumption is interdependent with how others behave, i.e.: that some of our energy behaviors are socially constructed. This is intuitive (while a certain level of energy consumption is required to meet our needs, a large portion of our daily energy behaviors depend on what we believe to be socially acceptable, as well as the behavior of our peers (Wolske et al., 2020)), and is also backed by previously researched successful interventions that leverage social norms (Andor and Fels, 2018). Furthermore, using the framework of Bicchieri and Dimant (2019), we recognise that a large part of the literature is primarily concerned on appealing to descriptive norms by altering empirical expectations on social behaviors.

The most heavily researched social comparison interventions have been ran by the US utility company OPower, where consumers are sent Home Energy Reports (HERs) through the mail with varying levels of frequency (Allcott, 2011). More recently, digital devices such as ”smart meters” and other In-Home Displays (IHDDS) have allowed for more flexibility and a higher frequency in the delivery of social information, as well as the combination of several types of interventions to study their aggregate and interactive effects (Schultz et al., 2015). Our methodology uses IHDDS as feedback mechanisms that integrate appeals to social norms in order to obtain a desired behavioral change (i.e.: reduction in energy consumption). Despite the great potential offered by IHDDS for the implementation of behavioral interventions, their effectiveness as delivery modes in social comparison interventions is still under-researched. Therefore, our study contributes to this emerging body of literature, and analyses the efficacy of such an intervention in a social housing context.

Despite differences in feedback frequency and delivery mode (Farrow et al., 2017a), implementations of norm-based interventions in the energy domain share several commonalities. One common attribute of norm-based interventions in the energy domain is the combination of appeals to injunctive norms as well as descriptive norms. Evidence suggests that descriptive norms are more effective in encouraging behavior change than injunctive norms, however appeals to descriptive norms in isolation can lead to what is known in the literature as a boomerang effect (Clee and Wicklund, 1980).
The boomerang effect in this context refers to an increase in energy consumption from households initially consuming less than the norm once they have access to the social information. This risks backfiring the intervention’s expected efficacy, and can have consequences on the net results of the intervention. However, when descriptive norms are used in conjunction with injunctive norms, the boomerang effect has been shown to disappear (Schultz et al., 2007).

Another common aspect is the target demographics that these interventions are aimed at. For the most part, these interventions have been limited to residential energy use, and primarily in the private sphere. The effectiveness of these interventions on energy use in the public sphere has been largely ignored. In this paper, we start to contribute to this line of research by studying the effectiveness of social comparison interventions in social housing. Further research could look at their effectiveness in other aspects of the public sphere, such as public schools, hospitals, etc.

There is no general consensus in the literature as to the success of social comparison interventions in the energy domain, but estimates from applications in private households seem to suggest the interventions lead to reduced energy consumption of 1.2% to 30% (Andor and Fels, 2018). However, very few of these studies use IHD devices in their delivery. In comparison, Schultz et al. (2015) finds a reduction of approximately 7% in energy consumption from households receiving norm messages integrated in IHD devices. However, this can vary widely on a case-to-case basis, with some backfiring effects observed in some contexts (Farrow et al., 2017b), particularly in low energy users (Schultz et al., 2007). Furthermore, some evidence from Germany (Andor et al., 2018) seems to suggest these interventions are less effective with European populations, who typically consume less energy on average.

The ambiguity of the existing evidence suggests that these norm-based interventions should be designed carefully, with a clear understanding of what discreetly defined behavior we aim to achieve, what are the underlying expectations we want to affect in order to do this and, most importantly, who we are targeting and how do they make energy decisions. For example, some households (particularly those with lower incomes or more restrictive budgets) have been shown to exhibit a “prebound effect” (Sunikka-Blank and Galvin, 2012) wherein they consume less energy pre-retrofitting than expected from techno-centric estimates, at cost to basic quality of life given that they usually live in energy-inefficient buildings. Behavioral patterns such as the prebound effect constitute a challenge for behavior-change interventions, but also illustrate why it is so important that technical innovations aiming to make the housing stock more efficient are accompanied by a good understanding of pre-intervention behavior and leverage it to increase potential energy savings.

2.3 Conceptualisation of Energy Poverty

By focusing on the specific context of retrofitted social housing, our study adds to the literature on energy poverty, particularly in relationship to behavioral-change interventions that tackle the issue (DellaValle, 2019). While currently there is no academic or policy consensus regarding the definition of energy poverty, a leading conceptualisation that we will adopt for the remainder of this study is the capabilities approach, first applied to the energy domain by Day et al. (2016).

In particular, energy poverty is conceptualized as a “situation of inability to realize the essential capabilities as a result of insufficient access to affordable, reliable and safe energy services, and taking into account the alternative means of realizing those capabilities in a reasonable manner”. Our motivation for adopting the capabilities definition is two-fold.

Firstly, the theoretical basis of this approach is grounded in the link between energy and well-being by explicitly acknowledging the relationship between energy services and the realisation of basic capabilities (good mental and physical health, social acceptance, access to education,
etc.), more so than other measures discussed in the literature. This is particularly important when considering the social housing context of our study, a demographic typically characterised by vulnerable conditions in socio-economic terms (low-income households, ageing populations, large families) and a high level of energy vulnerability (high amount of hours at home, troubles in paying energy bills, etc). In these contexts, basic capabilities are not always realised, making it of paramount importance to acknowledge their connection with energy services, particularly in regards to guiding policy-actions and interventions.

Secondly, the relative flexibility of this definition suggests a holistic approach to tackle energy poverty. In their conceptualisation, Day et al. (2016) explicitly describe the multi-step relationship between energy sources, such as biomass and solar, energy services such as lighting, space heating and cooling, and outcomes, defined as secondary and basic capabilities. This framework is not only crucial in understanding the relationship between energy and well-being, but can be used to locate policies and interventions aiming to tackle energy poverty in relation to which aspect of this relationship they address. Interventions improving dwelling efficiency intervene on the relationship between energy source and domestic energy supply, while behavioral interventions, which act by altering and optimising the usage of energy services, intervene on the relationship between energy services and capabilities. As an illustrative example, we take heating. Retrofitting an apartment will improve technical efficiency, acting on the relationship between energy supply and the specific energy service (heating). Our behavioral intervention, which aims to optimise heating and other energy behaviors, will act on the relationship between heating and its outcomes in terms of capabilities, such as being capable to keep the home appropriately warm at night (secondary capability). This will lead to a more effective use of the energy service, which in turn increases freedom degrees to act (DellaValle and Sareen 2020), i.e. increases the capacity to keep the home warm at night, avoiding health problems and being able to access education in the evenings. By promoting a better use of energy services, our intervention can also be effective at enhancing agency perceptions of vulnerable individuals, who usually experience a lack of agency (Sen, 1999).

Our study takes the view of recent studies recognising behavior as a driver of energy poverty. A number of papers have suggested that a key factor determining energy poverty is the interaction between low household incomes and thermally inefficient homes (Bouzarovski, 2014). In this sense, it is clear that energy inefficiency is a contributing factor to energy poverty which goes beyond limitations in disposable income from low-income households. However, recent literature (Kearns et al., 2019) has begun to pay attention to energy behaviors as an additional driver of energy poverty, recognizing factors such as use of household spaces and failure to adopt ‘adaptive thermal comfort’ as significant in leading households to energy poverty. These factors lead to inefficient behaviors that cause increasing difficulty in managing and paying energy bills, causing detrimental effects on physical and mental health, which can further contribute to the worsening of energy poverty conditions (poor mental health can lead to the adoption of poor heating regimes, and increasing challenge in a households ability to keep warm/cool). Inefficient occupant behaviour affected by poor mental health is particularly prevalent in vulnerable groups, causing the the adoption of inefficient energy regimes which then aggravate health problems, leading to the self-fulfilment of more sub-optimal behaviors (Mould and Baker, 2017). This self-fulfilling cycle is a clear example of why energy poverty is of particular concern in vulnerable groups, and of the importance of recognizing behavior as an additional policy target to tackle energy poverty.

The broader conceptualisation we depicted in this section allows us to locate our intervention in a wider set of policy-actions by explicitly describing the relationship it tries to affect and its objectives, differentiating it from other interventions (such as technical efficiency improvements) and suggesting a holistic approach to policy-making.
### 2.4 Energy behaviors in a social housing context

A notable feature of our study is the choice of the specific target group for our intervention. Social housing tenants are a demographic that is often overlooked in energy behavior change research (Hafner et al., 2020), yet they present a particularly interesting and important group to study for a number of reasons.

Firstly, due to the very aim of social housing being to provide affordable housing for all, there is usually a high representativeness of vulnerable demographics in social housing populations. This includes low-income households, unemployed individuals, retirees, disabled individuals, and large families who are typically home for large portions of a day. These groups are exposed to a number of energy vulnerabilities, for example low-income groups spend a larger share of their income on energy costs than high-income households (Schaffrin and Reibling, 2015). In some cases tenants may need to make energy-consuming adjustments to the dwelling, or add consumptive appliances for health-related reasons (keeping house warm, medical equipment, etc.). Additionally, social housing is typically energy-inefficient, and even in recently retrofitted housing (as is the case in our pilot study), empirical evidence highlights critical behavioral responses that limit the effectiveness of efficiency upgrades (Sorrell et al., 2007). This all suggests that the failure to adopt efficient energy behaviors can have substantial negative distributional or health-related consequences for social housing tenants. Subsequently, these tenants have the most to gain from an intervention that leverages their behavior to achieve energy savings, while also possessing unique energy needs that have to be considered by policymakers and practitioners. Research carried out using this particular group of tenants in Bolzano confirms that vulnerable situations are also apparent in the investigated Sinfonia districts (DellaValle et al., 2018), where the majority of individuals are identified as low educated or retired.

Secondly, there exists a large literature on the psychology of scarcity that points at the potential impacts that living in precarious conditions may have on energy decision-making. For example, scarcity has been shown to focus attention on the most immediate concerns (for the vulnerable this may be paying rent and bills, improving health, caring for children or the elderly), while significantly depleting attention for decisions that are not considered of immediate importance (Shah et al., 2012). Paradoxically, research also shows that this attention depletion leads to sub-optimal decision-making in some domains that would have helped individual combat their existing conditions of scarcity (Tomm and Zhao, 2016). This large body of evidence could also be applied to energy decision-making in vulnerable demographics. In particular, the psychology of scarcity could lead to the adoption of sub-optimal energy behaviors and the lack of interaction with behavior-change interventions, that actually contribute to helping reduce scarcity in the form of lower energy costs.

We should also expect that resource scarcity will worsen the individual tendency towards myopia in the energy domain (DellaValle, 2019). This refers to the over-weighting of present costs and benefits, and the under-weighing of future ones in a time-inconsistent fashion (Loewenstein and Prelec, 1992), leading to sub-optimal choices in the long-run. In the energy domain, such myopic behavior results in the undervaluing of future benefits associated with adopting energy efficient behaviours (Hershfield, 2011). Overall, the literature gives us ample reason to believe that the specific vulnerable conditions that social housing tenants are exposed to will cognitively impact them, leading to the adoption of sub-optimal energy behaviors. Targeting the behaviors or tenants may partially overcome the implications of this cognitive taxation.

Finally, the cognitive impact of stigmatization, deeply linked with social housing residency, poses barriers to the achievement of several benefits accrued by the adoption of energy efficient behaviors. Stigmatization has been shown to be linked to under-performance (Mani et al., 2013),

---

Electronic copy available at: https://ssrn.com/abstract=3659866
due to the depletion of executive resources deriving from efforts to suppress negative thoughts and emotions in the service of self-regulation (Hall et al., 2014). Furthermore, stigmatization has also been shown to result in social distancing, whereby individuals distance themselves from a prescribed social identity (Horan and Austin, 2014). These factors deriving from scarcity, pose important barriers to the adoption of energy efficient behaviors. Cognitive under-performance can have severe implications on economic decision-making and can lead stigmatized tenants to fail to recognize the potential financial benefits that accrue from efficient energy behaviors and energy investments. Similarly, social distance can lead to a deterioration in the willingness to contribute to a public good (Horan and Austin, 2014); being the decision to consume and invest in energy efficiency a type of decision to contribute to a public good (Brekke and Johansson-Stenman, 2008a), social distance is likely to pose additional barriers to the adoption of efficient energy behaviors.

2.5 Sources of heterogeneity in energy conservation

In order to disentangle the effect of our intervention from other motivations to conserve energy, we need to understand the decision-making process of energy conservation and, subject to data availability, control for ulterior factors that may affect the underlying choice of conserving energy.

Energy conservation can be interpreted as a pro-environmental behavior (Brekke and Johansson-Stenman, 2008b). Accordingly, we need to account for heterogeneity of factors underlying the decision-making process to act pro-environmentally. While random assignment into our treatment group can partially account for heterogeneity in these factors (which can be assumed to be equally distributed across the population), we control for them in our analysis in order to identify more closely the effects of our intervention. For example, a tenant’s decision to act more pro-environmentally by consuming less energy could be motivated by (i) a desire to act in accordance to empirical and normative expectations (targeted by our intervention) (Bicchieri, 2005), (ii) possessing a high-degree of intrinsic pro-environmental self-identity (Whitmarsh and O’Neill, 2010) or (iii) possessing an intrinsic motivation to contribute to a public good (Bénabou and Tirole, 2006); being the environment the most prominent public good (Brekke and Johansson-Stenman, 2008b), amongst others. We thus measure and control for pro-environmental self-identity and a number of primary predictors of contribution to a public good: trust, altruism, and reciprocity (Kollock, 1998), in order to allow us to more closely understand the effects of our intervention.

The decision to conserve energy can be also understood as an inter-temporal utility trade-off between present consumption and future financial benefits (in the form of a lower energy bill). As highlighted in the previous section, there is reason to believe myopic decision-making in the energy domain is prevalent due to the psychological burden of scarcity. However, a proportion of our population may be intrinsically very patient and willing to sacrifice some consumption now to benefit from lower energy-related expenses in the future. Therefore, the decision to conserve less energy, similarly to the decision to invest in energy-efficient appliances (Newell and Siikamäki, 2015), can be motivated by an intrinsic preference for delayed returns. In our analysis we therefore elicit and control for time preferences.

Finally, the decision to conserve more energy may simply be due to a better understanding on how energy behaviors relate to environmental and financial outcomes. For example, even if an occupant self-identifies as environmentally-friendly, she may not adjust to more conservatory behavior if she fails to recognize the link between her energy behaviors and environmental outcomes. Therefore, following (Blasch et al., 2017), we control for a general level of energy literacy in our analysis.
2.6 Pilot application aims

Our pilot experiment has been designed to address two main research questions:

1. What are the effects of social comparison interventions, integrated within IHDs, in a social housing context?

2. Can a social comparison intervention applied to a target demographic comprised primarily of vulnerable individuals, help alleviate energy poverty?

To tackle these questions with our accessible set of data, we make a simplifying assumption on the drivers of energy vulnerability in our target demographic, which allows us study differences in the evolution of energy consumption between groups. In particular, we assume pre-existing energy behaviors are sub-optimal and exacerbating a household’s position of energy vulnerability (Kearns et al., 2019). Therefore, if we observe a larger reduction in energy consumption during the investigated period in our treatment group than in our control group, we can take this result as signalling that our intervention was successful in optimizing energy behaviors, and in turn in reducing the energy vulnerability of households. Of course, these assumptions are limiting and a closer study on energy poverty conditions (whether through indoor temperature monitoring or self-reported measures) would have allowed us to tackle the second question more carefully. Obtaining this data however is either outside of the scope of this paper (temperature will be studied in a forthcoming paper on thermal energy behaviors) or impossible in practice at this stage of the project (dynamics in self-reported energy capabilities have not yet been collected). For the scope of this paper, this assumption is instrumental to study the impact of the intervention on energy poverty by looking at electricity consumption only.

Overall, our study aims to discern how social comparison interventions fare in an environment characterized by resource scarcity and inefficient energy behaviors. The success of our intervention faces several barriers deriving from the specific context of social housing, such as a high level of energy vulnerability, and the cognitive impacts of scarcity and stigmatization. On the other hand, it is also because of these contextual reasons that understanding the effects of this widespread behavioral intervention is of paramount importance. It can highlight pathways to the successful implementation of energy efficiency investments that account for and target social aspects by promoting the adoption of more virtuous energy behaviors, thus contributing to drawing social housing tenants out of a situation of energy poverty. If our intervention is unambiguously successful, it can further promote the roll-out of these uniform normative appeals in the context of social housing retrofits. Alternatively, if we find substantial resistance in the intervention success, or encounter unique difficulties that limit the intervention’s effectiveness, our results can further support the importance of targeted feedback programs (Khosrowpour et al., 2016) that address the particular needs and characteristics of the most vulnerable.

Our results unveil practical recommendations for policy makers who wish to maximize the impact of retrofit interventions in social housing settings, mindful of the contextual influence: this latter has to be carefully examined before designing behavior-change interventions in vulnerable demographics. Specifically, behavioral policies in vulnerable demographics should be financially assessed vis-a-vis price interventions in order to choose the most efficient policy instrument to achieve the desired social and environmental objectives.
3 Materials and methods

3.1 Context and pilot design

Our pilot study is embedded within a wider EU-funded project called Sinfonia, ran in social housing districts in Bolzano, South Tyrol. The Sinfonia Smart city project, born from the cooperation between Bolzano and Innsbruck, aims at finding integrated solutions to achieve significant levels of energy savings in social housing districts\(^5\). As part of this project’s activities, a number of apartment buildings in different districts throughout Bolzano were retrofitted to make them more energy efficient. After the works finished, a number of consenting apartments were installed sensors and ”smart meter” (IHD) technology providing timely feedback about several household characteristics relating to energy efficiency and comfort. These characteristics include humidity, temperature, air quality and, notably, electrical and thermal energy consumption (in terms of kWh and Wh/m\(^2\)).

The home page of the smart meter display can be seen in Figure\([1]\). These displays were shown on a tablet installed near the tenants front door, which is being transmitted the information recorded by the sensors.

---

\(^5\)Sinfonia website, http://www.sinfonia-smartcities.eu/

![Figure 1: Control Group Home Display](image-url)
history” tab, users are brought to a separate page as seen in Figure 2. Users can then navigate this page to find information on their past energy consumption aggregated at different levels (daily, weekly, monthly), and they can visualise the evolution of their consumption levels through these visualisations.

![Control History Display](image)

Figure 2: Control History Display

It is important to note that due to the necessity of receiving consent to install the IHD technology, our experiment is susceptible to self-selection bias. It could be for example that subjects consenting to the installation of the monitors already share pro-environmental attitudes, or an awareness about the role of behavior on environmental outcomes. This limitation is discussed further in Section 5.

3.2 Pilot Study Design

Households with installed IHDs are randomized into two different groups, the control group and the norm group. The control group receives feedback on their own energy consumption through the home and history pages as described above. The layout of the pages in their display is identical to that shown in Figures 1 and 2.

The norm group receives identical information on their own consumption as the control group, but their level of energy consumption is also compared in their display to that of a reference group of neighbors. This comparison is represented both in terms of last-day averages (as shown in Figure 3) and in different formats through the "history" tab (shown in Figure 4). In short, tenants in the control group receive only information on their own electric and thermal consumption, whereas tenants in the norm group receive own-information as well as social information, all displayed in the same modules within their display. This allows the user to easily compare their own consumption.
Close attention was paid to the selection of the reference groups. Evidence suggests that the choice of reference group is crucial for the effectiveness of norm-based interventions (Abrahamse et al. (2005), Bicchieri and Dimant (2019)). Towards this end, we use a restrictive similarity criterion to group households into different reference groups on the basis of observable characteristics which reflect actual energy use. These are: number of household occupants and average number of hours spent at home by household members. Owing to the necessity to create clusters within district level (due to the deployment of different technologies in the different districts) and the existing limitations in the amount of data collected at time of analysis, we were forced to trade-off the inclusion of some potentially relevant comparable characteristics around which to group households (number of children, number of retired household members, etc.) in order to avoid minimizing the size of the reference groups. If all households were to be compared with only one or two more neighbors for example, the relevance of the social information would have been significantly diluted.

Furthermore, in line with previous literature (Alberts et al. (2016), Anderson et al. (2017)), we choose to compare individual household behavior not to an average of other households in the reference group, but rather the behavior of top performers of the reference group to provide a truly virtuous example to follow. We were able to create comparable groups comprising of 3-4 households in our investigated apartments, and picked the average energy consumption of the top two highest performers in that group as the shared social information. Given the limited data available at time of the initial analysis however, the use of a restrictive similarity criterion may have had unintended consequences on the significance of generated social information, something that will be discussed in Section 5.
The high feedback frequency of consumption information is a notable contribution of our study to the overarching literature on norm-based interventions using IHDs. As mentioned, users can visualize current consumption information, which is updated every 5-minutes. There has been mixed evidence on how the frequency of feedback affects energy conservation efforts. Fischer (2008) argued that frequent feedback on energy consumption was more effective than infrequent feedback due to the closer link it creates between actions and consequences, but later empirical evidence has refuted this claim (Ehrhardt-Martinez et al., 2010), finding real time feedback to result in lower conservation efforts than weekly/daily feedback. In an experimental environment, Casal and Valle (2016) also find that the frequency of feedback does not impact individual performance.

For our purposes, studying whether the intervention is effective in promoting conservation in a social housing context, the frequency of feedback is only relevant insofar as all tenants have access to information on their consumption (and others’ consumption for the norm group) at the same frequency. This is the case in our pilot study. Future research could focus on the effects of increased feedback frequency on energy conservation with a particular focus on social housing.

In our norm-based intervention we appeal to both descriptive and injunctive norms. We employ "smiley face" emoticons similar to those used in the HERs in Allcott (2011). These emoticons are meant to appeal to injunctive norms by suggesting the social desirability of a behavior. If a specific household were currently consuming less than their reference group of neighbors, they would be presented with a smiling face together with the social information, while if the household was currently consuming more than the reference group of neighbors, they would receive a red frowning face. Following the literature, this was done to reduce the likelihood of observing a boomerang effect, whereby low consuming households start increasing their consumption once they realise they are consuming less than the norm.
Our experiment also draws from previous evidence highlighting that social comparison interventions are more effective when complemented with actionable tips (Dolan and Metcalfe, 2013). The technology of the IHDs allows us to suggest targeted actions to reduce consumption of energy in households while maintaining a suitable level of comfort (such as opening a window and turning down heating when the outside temperature is greater than the inside temperature). These tips are available to households both in the control and treatment groups, meaning that in our analysis we only isolate for the effect of exposure to social information, and not the inclusion of the tips.

3.3 Data description

To study whether changes in behavior have taken place in the short-term (as this is the period when a potential change in behavior is likely to occur due to the novelty of the display), we studied the effect of the intervention during the first three months of implementation. The project is expected to last longer than the 3-month period considered in the scope of this study, but to enhance project accountability, it is important to highlight immediate results of our intervention. In a later study, we will analyse also the long-term effects of our intervention and for households from the remaining districts. It is also important to note that the project is complex, due to the different technical characteristics of the dwelling and the different interventions being implemented in each investigated district. In this study, we are solely interested in studying the effects of the social comparison intervention, and adopt our analytical approach accordingly: we compare households in the treatment group with their appropriate counterfactual (control group). Due to the randomized assignment of households into the two groups at district level, the only discernible difference between these households at the aggregate level is whether or not they see the social comparison module.

This initial analysis includes only the first apartments to have their displays activated in one of the project districts. The households included in the survey all had their displays activated at the same moment, meaning they were exposed to the intervention for an equal amount of time. This included 13 apartments initially. Subsequently, one of the tenants asked to have the sensor uninstalled and was removed from the sample. This left us with 12 apartments across 2 groups analysed over a period of 3 months, from November 22nd to February 23rd. This included 27 tenants.

Using the sensor technology installed after the retrofitting works, we gathered data on hourly energy consumption for each of the 12 apartments. Taking advantage of both the longitudinal and cross-sectional nature of our data, we created a panel dataset that collected highly granular information on energy consumption across the 12 households. We aggregated our granular data at the hourly level for the sake of tractability.

In this study, we focus only on electrical energy consumption, which was measured in kWh. As anticipated earlier, thermal energy consumption will be investigated in a follow-up study.

It is worth mentioning that we encountered some technical difficulties during the data collection process that resulted in us receiving distorted hourly data on electricity consumption. Due to the nature of the sensing technology and a margin of error, for a small proportion of hours and in a limited number of apartments, consumption was recorded but not reported immediately. Instead, the recording system aggregated the results from multiple hours of consumption into the observation for a single hour. This led to some incorrect observations in our dataset that could bias our regression analysis in particular. We decided to drop these biased observations from our dataset before running our regression analysis: whenever there was a gap in reporting of more than 1 hour for any apartment, the following observation was dropped. While this makes it so that we lose a small number of observations (hence ending up with an unbalanced panel), it allows our analysis to
be unaffected by technical difficulties in the sensor and recording technology, ensuring that every observation in our dataset is indeed collecting consumption during the span of a single hour. It is important to note however that, as the incidence of these errors disappeared when aggregating consumption at the daily level, these observations were not dropped when completing our DID analysis.

A notable limitation of our dataset is the inability to access long-term pre-intervention data on energy consumption. This does not allow us to make a comprehensive Difference-in-Differences (DID) Analysis. Therefore, a proxy DID approach was adopted, following Bager and Mundaca (2017), as will be explained in the following section.

We report the distribution of electricity consumption (hourly measurements) in our data in Figure 5 and summary statistics for this variable can be found in Table 1. As we can see, the distribution of energy consumption in both groups is skewed to the left, with the mean hourly level of energy consumption in the overall sample being 0.248.

![Figure 5: Distribution of data on hourly energy consumption during first 3 months of intervention across groups](image)

| Type  | Observations | Mean  | Std. Dev. | Min    | Max    |
|-------|--------------|-------|-----------|--------|--------|
| Overall | 26,576       | 0.248 | 0.294     | 0.0015 | 2.903  |
| Control | 13,190       | 0.248 | 0.290     | 0.0015 | 2.903  |
| Norm   | 13,386       | 0.249 | 0.298     | 0.005  | 2.666  |

Table 1: Summary Statistics of Hourly Electricity Consumption (kWh/h)
Table 2: Descriptive Characteristics of Sample

| Variables        | N  | Mean     | St. Dev. | Min | Max |
|------------------|----|----------|----------|-----|-----|
| Female Dummy     | 27 | 0.518    | 0.509    | 0   | 1   |
| Retired Dummy    | 27 | 0.259    | 0.446    | 0   | 1   |
| > 65 Dummy       | 27 | 0.259    | 0.446    | 0   | 1   |
| 40-59 Dummy      | 27 | 0.481    | 0.509    | 0   | 1   |
| Children Dummy   | 27 | 0.0740   | 0.267    | 0   | 1   |
| N. Children      | 12 | 0.167    | 0.389    | 0   | 1   |
| N. > 65          | 12 | 0.583    | 0.793    | 0   | 2   |
| N. retired       | 12 | 0.583    | 0.793    | 0   | 2   |
| > 12 h at home Dummy | 12 | 0.833    | 0.389    | 0   | 1   |

Our dataset also contains information on observable household characteristics, both regarding the dwelling and household composition. This includes information on contracted energy level (which was subsequently dropped from analysis as all apartments shared the same contracted level of 3 kWh), number of occupants, age of occupants, occupation of occupants, gender, and average number of hours spent at home by each occupant. From this data, we were able to generate several aggregate-level variables relating to age groups, number of children, number of retired occupants, number of unemployed occupants, and so on. Some of these structural aspects are controlled for in our regression analysis to estimate treatment effects more accurately.

A descriptive analysis on the observable characteristics of the sample population is summarised in the Table 2. The first five rows display several of the respondent’s characteristics while the last four represent those of the household.

These descriptive statistics reveal some key points. Firstly, a significant proportion of the tenants in the sample are retired and over 65 years of age (26% of individuals in the sample for both categories). Subsequently, the majority of the households have at least one member that stays at home more than 12 hours each day. This confirms that the sample prominently features vulnerable individuals, such as the elderly and retired, and that the general energy needs of the our sample might be high.

Our sample does not feature a large number of children, meaning the large prevalence of individuals staying long hours at home is not primarily driven by adults staying home with their children. Rather this is likely driven by retirees, or potentially the unemployed.

In addition to data on energy consumption, and observable household characteristics, we collected data on a number of household preferences using surveys administered at the time of installing the smart meter. The objective of the survey is to collect valuable information on the individual preferences of occupants in order to control for factors that may affect the underlying decision-making process of conserving electricity, in absence of our intervention. By controlling for these potential sources of heterogeneity in our analysis, we can determine how much of the resulting change in behavior can be attributed to the intervention. The survey-elicited measures are therefore integrated into our regression analysis as additional explanatory variables.

The survey-elicited data included measures on energy literacy (Blasch et al., 2017), pro-environmental self-identity, altruism, trust, reciprocity, group identity and inter-temporal preferences. The survey items, based on Luhtanen and Crocker (1992) and the experimentally-validated items developed by Falk et al. (2018), were a combination of Likert scales and multiple-choice questions.

It is important to note that many of the items in the survey elicit information on individual characteristic of the respondent, not necessarily the household as a whole. As the survey was
conducted on only the one household member present at the time of installation, the collected measures are likely affected by an individual bias and can at best be used as proxies of general household characteristics. This is a limitation of the collected data, that can lead some of the relevant variables to have an individual-level bias. This is explained further in the Discussion section of this paper.

4 Results

4.1 Descriptive Analysis

We start by studying differences in overall average energy consumption between the two groups in the short-term. Figure 6 and Table 3 show no significant differences at the average level amongst groups, with the control group consuming 0.248 (St. Dev = 2.90) as opposed to the norms group consuming 0.249 (St. Dev = 0.298).

![Figure 6: Average Energy Consumption across groups](https://ssrn.com/abstract=3659866)

This is confirmed by an a cross-sectional Mann-Whitney U test at the household level that fails to reject the null of equal distributions across groups (p=0.6310).

Next, we are interested in understanding how electricity consumption patterns of the tenants evolve over the initial 3 months of intervention, in order to distinguish any time-dependent differences in consumption amongst groups. Figure 7 shows the daily evolution of hourly averages in
Table 3: Summary Statistics of Hourly Energy Consumption of sample (kWh)

| Group | N      | Mean | Std. Dev. | Variance |
|-------|--------|------|-----------|----------|
| Control | 13,190 | 0.248| 0.290     | 0.084    |
| Norm   | 13,386 | 0.249| 0.298     | 0.089    |

electricity consumption across groups. Here too, we fail to identify discernible differences in the evolution of daily consumption behaviors between the two groups.

4.2 Difference-In-Differences

Following Bager and Mundaca (2017), we employ a DID approach by measuring relative change in daily electricity consumption from the first to the last week of intervention for both groups, and comparing the changes between groups. Our variable of interest is weekly average daily electricity consumption.\(^6\)

\[^6\]This was measured by aggregating hourly energy consumption over a single day. This daily energy consumption measure was then averaged at a weekly-level.
An ideal implementation of this approach would compare pre-intervention to post-intervention data on consumption, but as stated earlier, we do not have access to pre-intervention electricity consumption data. The results for the weekly DID analysis can be found in Table 4, and a graphical representation of the weekly evolution of daily consumption can be found in Figure 8.

Table 4: Daily Average Hourly Electricity use for households by Groups

| Group | First week average (kWh) | Mid-period week average (kWh) | Last week average (kWh) | Change in Consumption ($\delta\%$) |
|-------|--------------------------|-------------------------------|------------------------|----------------------------------|
| Control | 6.179                    | 6.446                         | 5.176                  | -16.232%                         |
| Norm   | 6.145                    | 5.866                         | 5.588                  | -9.064%                          |

As is clear from the table, over the relevant period of our intervention both groups experience a reduction in their daily electricity consumption. However, we see a larger reduction in the energy consumption of the control group, rather than the norms group (a differential effect of 7.168%), suggesting a backfiring role of our intervention. From Figure 8 however, we can see there are no marked differences between daily consumption of the two groups throughout the weeks.

We see also from Table 4 and Figure 8 that both groups reduced their weekly average daily energy consumption overall in the investigated period. More detailed analysis, taking into account...
weather effects, is needed to understand why both groups go through a reduction in electricity consumption. We can speculate, based on previous research on the effects of increased feedback on consumption (Faruqui et al., 2010), that the display installation did have an effect in reducing overall electricity consumption, but that the normative appeals were unable to promote further reductions for the treatment group. It must be noted that the displays were installed in all investigated apartments, making it impossible to establish a causal relationship between display installation and the general reduction in electricity consumption we observe in both groups.

To statistically analyse our panel data set, while acknowledging the limitations of working with a small sample size, we use a paired t-test to compare the mean values between the two groups. Following medical research suggestions for small-sample pilot studies such as ours (Bager and Mundaca, 2017), we consider different significance levels for our test ($\alpha = 0.05 - 0.30$) and different confidence intervals than used in most large-scale field experiments (i.e. 70-90%). However, even when adjusting for our smaller sample size, we fail to identify significant statistical effects of our intervention ($p=0.80$) at the hourly level. For the sake of completeness, we again run a two-sample t-test on aggregated daily consumption, again failing to identify statistically significant differences in daily energy consumption between the two groups ($p= 0.759$).

4.3 Regression Analysis

We now turn to regression analysis to study more closely the effect of our intervention on electricity consumption during the initial 3-month period of implementation. The regression analysis method we present here is intended to be a methodological starting point for the complete analysis that will be conducted once we have access to data from other districts. It is important to note however that due to the currently low number of households included in the analysis, indications of statistical significance should be considered with caution.

We start by defining a bivariate model to analyse the impact of treatment assignment on hourly energy consumption levels of individual households:

$$electricity_{it} = c + \beta \text{social}_i + \epsilon_{it}$$

Our dependent variable is kWh of electricity consumed by household $i$ during hour $t$ in our panel data-set ($t=1-2233$). $Social$ is our treatment dummy variable that takes the value of 1 if the household is in the norm group, and 0 if the household is in the control group.

We further enrich our model (following (Schleich et al., 2017)) by adding hourly and monthly dummies, in order to control for variations in electricity demand across months (i.e. due to weather conditions) and across hours of the day (i.e: due to variations in household occupancy and activities). We are interested in studying how our intervention performs when we take into account this variability.

$$electricity_{it} = c + \beta \text{social}_{it} + \sum_{m=1}^{4} M + \sum_{h=1}^{24} H + \epsilon_{it}$$

Finally, we define two multivariate models (MV(1) and MV(2)) that separately control for household structural characteristics, as well as the survey-elicited preferences. The two multivariate models are defined separately in order to circumvent potential biases in our results deriving from analysing data collected at two different levels (household structural variables represent household level characteristics, while survey-elicited preferences represent individually-collected preferences).
In defining MV(2), we use the individual preferences as household-level proxies for overall preferences, in order to control to some extent with potential sources of heterogeneity in energy consumption deriving from individual preferences. This approach however suggests treatment effects derived from estimation of our MV(2) model are to be considered carefully, acknowledging this aggregation of preferences.

The two multivariate models then become:

\[ electricity_{it} = c + \beta_{social_i} + \alpha_{lattermovein_i} + \mu Z_i + \epsilon_{it} \]

and

\[ electricity_{it} = c + \beta_{social_i} + \sigma \text{TimePref}_i + \zeta \text{ReciprocityPref}_i + \rho \text{AltruismPref}_i + \delta \text{GroupId}_i + \lambda \text{EnvSelfId}_i + \gamma \text{TrustPref}_i + \theta \text{EnerLit}_i + \epsilon_{it} \]

\[ \text{latermovein} \text{ is a dummy variable that accounts for when the tenants moved into the apartment in the timeline of the retrofitting project. } \]
\[ Z_i \text{ is a matrix of variables measuring different household characteristics, including number of children, number of males and females, number of people over 65, number of occupants in the household, and average number of hours spent indoors by occupants. We exclude data on dwelling size for this part of the analysis, as it is not expected to be strictly relevant for electricity consumption, but rather will be included in the follow-up study when investigating the effects of the intervention on thermal energy consumption. For the sake of completeness however, we also run all our regressions with the variable of dwelling size (in } M^2 \text{) included and report the results in the Appendix. We get qualitatively similar results for treatment effects both in terms of coefficient and size.} \]

Turning to the variables obtained by survey answers, \text{TimePref}, \text{TrustPref} and \text{ReciprocityPref} are ordinal variables that capture individual-level inter temporal preferences and preferences on trust and reciprocity, respectively, on a scale of 1 to 7. \text{AltruismPref} is a normalised variable that captures individual level altruistic preferences from 0 (extremely non-altruistic) to 1 (extremely altruistic). \text{GroupId} is a measure that captures degree of cohesion to social groups which captures social distance which, as previously explained, can be a predictor to willingness to contribute to public good. This measure is defined as an average to the answers to four questions relating to the level of identification with specific social groups. Finally, \text{EnvSelfId} is a measure of pro-environmental self-identity.

Finally, \text{EnerLit} is a normalised variable measuring the amount of correct answers out of 4 in a series of questions carefully designed to test the general level of energy literacy of the individual.

All models are estimated via the GLS panel random-effects estimator. In order to understand whether or not there are differences in intervention efficacy during weekends and weekdays, we also estimate all models separately for weekends and weekdays.

The results of our estimated model can be seen in Table 5. Our results confirm the limited, potentially backfiring role of our intervention. The estimated coefficient in our two bivariate models is positive, confirming that subjects in the treatment group consumed on average more than those in the control group. As expected by the small sample, the results do not reach statistical significance. Estimates of the average treatment effect in the bivariate model range from a positive effect of 0.03% to 0.07%, with the inclusion of monthly and hourly dummies diluting the positive effect.

Turning to our multivariate models, controlling for household characteristics, we again estimate a positive coefficient for our treatment dummy, but this time reaching the effect of the intervention is
statistically significant. On the other hand, when controlling for individual characteristics, we find that not only the coefficients associated to some individual characteristics (namely trust, reciprocity, altruism, time discounting, and level of energy literacy) are significant, but also that the coefficient of our treatment dummy turns negative. This suggests that heterogeneity in unobserved household preferences might play a crucial role in affecting the efficacy of our intervention.
Table 5: Results from GLS random effects regression.

| VARIABLES          | (1)      | (2)      | (3)      | (4)      |
|-------------------|----------|----------|----------|----------|
|                   | BV(1)    | BV(2)    | MV(1)    | MV(2)    |
| Social            | 0.000688 | 0.000331 | 0.124*** | -0.0476*** |
|                   | (0.0613) | (0.0740) | (0.00587)| (0.00852)|
| EnvSelfId         | 0.0205   |          |          |          |
|                   | (0.0134) |          |          |          |
| TrustPref         | 0.0263***|          |          |          |
|                   | (0.00347)|          |          |          |
| ReciprocityPref   |          | -0.0790***|         |          |
|                   |          | (0.00455)|          |          |
| AltruismPref      |          | -0.0626***|         |          |
|                   |          | (0.00835)|          |          |
| TimePref          |          | 0.0506***|         |          |
|                   |          | (0.00232)|          |          |
| EnerLit           |          | -0.0520***|         |          |
|                   |          | (0.00339)|          |          |
| GroupId           |          | 0.0162***|         |          |
|                   |          | (0.00293)|          |          |
| latermovein       |          | -0.183***|         |          |
|                   |          | (0.00628)|          |          |
| OccupantNumber    |          | -0.215***|         |          |
|                   |          | (0.0124) |          |          |
| HoursAtHome       |          | 0.0113***|         |          |
|                   |          | (0.00115)|          |          |
| MeanAge           |          | -0.00855***|        |          |
|                   |          | (0.000476)|         |          |
| N children        |          | -0.00759 |         |          |
|                   |          | (0.0109) |          |          |
| N > 65            | 0.155*** |          |         |          |
|                   | (0.00837)|          |          |          |
| N female          | 0.238*** |          |         |          |
|                   | (0.00849)|          |          |          |
| Constant          | 0.248*** | 0.160*** | 0.639*** | 0.178*** |
|                   | (0.0433) | (0.0531) | (0.0308) | (0.0655) |
| Observations      | 26,576   | 26,576   | 26,576   | 22,160   |
| Number of Apartments| 12   | 12    | 12    | 10    |

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Dependent variable is energy consumption of apartment i at hour t (t=1-2233). BV(2), MV(1), and MV(2) control also for hourly and monthly variations by including hour and month dummies in the regression (suppressed in output). MV(2) drops observations from two apartments from which we do not have survey data.
Turning now to estimating our model separately for weekdays (Table 6) and weekends, we see that during weekdays the effects of our intervention are not qualitatively different from the effects we observe using our entire range of data.
Table 6: Note: Results from GLS random effects regression including only weekdays.

| VARIABLES                  | (1)          | (2)          | (3)          | (4)          |
|----------------------------|--------------|--------------|--------------|--------------|
|                            | BV(1)        | BV(2)        | MV(1)        | MV(2)        |
| Social                     | 0.00830      | 0.00813      | 0.120***     | -0.0568***   |
|                            | (0.0621)     | (0.0755)     | (0.00688)    | (0.00999)    |
| latermovein                | -0.187***    |              |              |              |
|                            | (0.00737)    |              |              |              |
| OccupantNumber             | -0.191***    |              |              |              |
|                            | (0.0145)     |              |              |              |
| HoursAtHome                | 0.0103***    |              |              |              |
|                            | (0.00135)    |              |              |              |
| MeanAge                    | -0.00786***  |              |              |              |
|                            | (0.000559)   |              |              |              |
| N children                 |              | -0.0157      |              |              |
|                            |              | (0.0128)     |              |              |
| N > 65                     |              | 0.142***     |              |              |
|                            |              | (0.00983)    |              |              |
| N female                   |              | 0.220***     |              |              |
|                            |              | (0.00996)    |              |              |
| EnvSelfId                  |              |              | 0.0270*      |              |
|                            |              |              | (0.0156)     |              |
| TrustPref                  |              |              | 0.0288***    |              |
|                            |              |              | (0.00406)    |              |
| ReciprocityPref            |              |              | -0.0729***   |              |
|                            |              |              | (0.00533)    |              |
| AltruismPref               |              |              | -0.054***    |              |
|                            |              |              | (0.00981)    |              |
| TimePref                   |              |              | 0.0493***    |              |
|                            |              |              | (0.00272)    |              |
| EnerLit                    |              |              | -0.0563***   |              |
|                            |              |              | (0.00398)    |              |
| GroupId                    |              |              | 0.0216***    |              |
|                            |              |              | (0.00343)    |              |
| Constant                   | 0.239***     | 0.151***     | 0.593***     | 0.0665       |
|                            | (0.0439)     | (0.0544)     | (0.0362)     | (0.0766)     |
| Observations               | 18,624       | 18,624       | 18,624       | 15,534       |
| Number of Apartments       | 12           | 12           | 12           | 10           |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Dependent variable is energy consumption of apartment i at hour t (t=1-2233). BV(2), MV(1), and MV(2) control also for hourly and monthly variations by including hour and month dummies in the regression (suppressed in output). MV(2) drops observations from two apartments from which we do not have survey data.
When we estimate our model using only weekend data (Table 7), we find slightly different results. The estimated coefficient of our treatment dummy for MV(1) is still positive, further suggesting that sociodemographic characteristics are important determinants of the effectiveness of social comparison interventions in this context. However, the estimated coefficient for our treatment dummy is negative in both our bivariate models, despite being statistically insignificant.
Table 7: Results from GLS random effects regression including only weekends.

| VARIABLES          | (1)    | (2)    | (3)    | (4)    |
|--------------------|--------|--------|--------|--------|
|                    | BV(1)  | BV(2)  | MV(1)  | MV(2)  |
| Social             | -0.0172| -0.0181| 0.132***| -0.0275*|
|                    | (0.0607)| (0.0662)| (0.0111)| (0.0161)|
| latemovein         | -0.175***|
|                    | (0.0119)|        |        |        |
| OccupantNumber     | -0.270***|
|                    | (0.0234)|        |        |        |
| HoursAtHome        | 0.0134***|
|                    | (0.00217)|        |        |        |
| MeanAge            | -0.0101***|
|                    | (0.000899)|        |        |        |
| N children         | 0.0114|
|                    | (0.0206)|        |        |        |
| N > 65             | 0.185***|
|                    | (0.0158)|        |        |        |
| N female           | 0.281***|
|                    | (0.0161)|        |        |        |
| EnvSelfId          | 0.00714|
|                    | (0.0255)|        |        |        |
| TrustPref          | 0.0207***|
|                    | (0.00656)|        |        |        |
| ReciprocityPref    | -0.0940***|
|                    | (0.00861)|        |        |        |
| AltruismPref       | -0.0814***|
|                    | (0.0157)|        |        |        |
| TimePref           | 0.0534***|
|                    | (0.00441)|        |        |        |
| EnerLit            | -0.0422***|
|                    | (0.00640)|        |        |        |
| GroupId            | 0.00380|
|                    | (0.00555)|        |        |        |
| Constant           | 0.269***| 0.179***| 0.741***| 0.430***|
|                    | (0.0429)| (0.0497)| (0.0582)| (0.125)|
| Observations       | 7,952  | 7,952  | 7,952  | 6,626  |
| Number of Apartments| 12     | 12     | 12     | 10     |

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: Dependent variable is energy consumption of apartment i at hour t (t=1-2233). BV(2), MV(1), and MV(2) control also for hourly and monthly variations by including hour and month dummies in the regression (suppressed in output). MV(2) drops observations from two apartments from which we do not have survey data.
4.4 Interaction effects

We conducted a study of conditional marginal effects in our model for a subset of the significant variables. This was done in order to better understand the direction of our treatment effect at different levels of significant covariates in our model. We completed this study separately for the household structural variables and for the survey-elicited variables of significance.

The results of this analysis for the structural variables (Table 8; Figure 9) are not surprising given our earlier results and much of the literature on socio-demographic determinants of energy use (ˇSˇ cepanovi´ c et al., 2017). The difference in energy consumption of those households assigned to the control group and those assigned to the treatment group increases as the number of females and mature tenants increases. However, only our results relating to the variable “number of females” are significant. We can interpret these results as suggesting that the more females in a household, the wider the positive difference in electricity consumption between treatment and control groups, therefore the less our intervention has been effective. This highlights that household family-composition characteristics may be crucial determinants in how effective these interventions are in a social-housing context, and suggests the need for targeted interventions that take into account household gender composition.

| Delta Method | N female | N > 65 |
|-------------|---------|-------|
| 1.Social    |         |       |
| 0           | -0.099  | -0.047|
| 1           | 0.011   | 0.037 |
| 2           | 0.121** | 0.122 |
| 3           | 0.231** | 0.206 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Moving on to studying the marginal conditional effects of some of the survey-elicited variables, we are interested in studying the conditional marginal effects of reciprocity, energy literacy, and willingness to delay (Table 9). Carrying out a study on conditional marginal effects of these variables
enables us to improve identification of treatment effects at varying levels of individual preferences.

Our results for energy literacy and time preferences are not statistically significant. Turning to reciprocity (Table 10), there seems to be a statistically significant difference in electricity consumption between groups at different levels of this variable. This difference also seems to be increasing at higher levels of reciprocity, which counters the predictions of our theoretical framework, with higher levels of reciprocity signalling higher willingness to contribute to a public good and hence act in a more environmentally friendly manner. The results from our data, seem instead to show a diminishing negative difference, and eventually a positive difference in electricity consumption between groups as reciprocity increases, signalling a reduced effectiveness of our intervention for higher levels of reciprocity. While this result might be an artifact of the limited sample, it might also be driven by the reduced sense of agency from living in sub-optimal contexts that generally leads to a deterioration of social preferences (Becchetti et al., 2013).

![Conditional Marginal Effects of Reciprocity](image.png)

Figure 10: Conditional marginal effects of reciprocity

### Table 9: Conditional marginal effects of energy-related preferences

| Delta Method | ReciprocityPref | EnerLit | TimePref |
|--------------|-----------------|---------|----------|
| 1            | -0.992**        | 0.007   | 0.0555   |
| 2            | -0.812**        | -0.066  | 0.046    |
| 3            | -0.633**        | -0.139  | 0.036    |
| 4            | -0.454**        | -0.211  | 0.026    |
| 5            | -0.274**        |        | 0.016    |
| 6            | -0.094**        |        | 0.006    |
| 7            | 0.085           |        | -0.004   |

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
5 Discussion

5.1 Discussion of results

Our results, while acknowledging significant limitations in the data sample size (from only one district) and scope (electricity consumption only), suggest that context matters when applying social comparison interventions in social housing districts. We fail to identify significant differences in electrical energy consumption between our two groups at the average level, and find statistically insignificant treatment effects in our bivariate regression models. Furthermore, we observe some backfiring effects of our intervention when we control for household characteristics. These results stand in stark opposition to much of the existing literature which focuses primarily on large scale RCTs in similarly residential areas within the private housing sector. We do, however, estimate a statistically significant negative coefficient for our treatment effect when controlling separately for survey-elicited characteristics. The estimated results from this model will be corroborated by a future study with data from other districts.

A first recommendation emerging from this study is that policymakers and practitioners ought to fully consider the characteristics and particular behaviors of the target group before designing an intervention aimed at tackling energy poverty through behavioral change. It is paramount to take into consideration the specific needs, pre-existing behaviors and motivations, and key reference groups relevant for the demographics being targeted, in order to maximise the effectiveness of the intervention. In this respect, our results seem to support findings by (Khosrowpour et al., 2016) that highlight the need for targeted feedback mechanisms in behavior change interventions and acknowledge that "one size does not fit all".

In addition, these results seem to point that for this particular demographic, a more holistic behaviourally-informed intervention than one based solely on nudging, might be desirable. As an example, an intervention that provides individuals with basic facts on energy poverty might be implemented to boost skills and knowledge required for identifying and sharing their needs and problems related to home energy comfort and energy consumption (DellaValle and Sareen, 2020). This might not only a way to harness the local experience, thus truly engaging target individuals in the process of betterment of their conditions, but also a way to further increase their capabilities and optimise their energy-use to fit their specific needs. Additionally, given that uniform behavior-change interventions of this kind may be ineffective in isolation when targeting vulnerable demographics, there is scope to integrate a more holistic set of actions in the delivery of these interventions, for example based on community-based social marketing approaches. At the stage of analysis furthermore, the integration of more qualitative methods could complement our quantitative approach to better understand experienced conditions of tenants in relation to energy vulnerability and engagement with IHDs, particularly since energy poverty is a multidimensional concept. This idea is consistent with the approach of "mixed methods" proposed by (Sovacool et al., 2018).

The difference in the direction of our treatment effects between weekdays and weekends is also significant. Given our sample composition being comprised in large part of retirees, the unemployed, or other groups that stay a minimum of 12 hours at home, we should not observe substantial differences in energy behaviors for these tenants between weekdays and weekends. The fact that our intervention leads to a negative difference in electricity consumption between the treatment and control group when considering only weekend data, could suggest that the composition of occupants in a home at any one time has a moderating effect with our intervention. For example, young adults who are typically away during weekdays could have a higher likelihood to interact with the technology (and therefore see the normative appeals) than some of the older household members.
who are at home during weekdays. Such an interpretation could support the implementation of remote feedback mechanisms with integrated normative appeals and high granularity information on consumption, such as an app. This more accessible information could lead to the diffusion of more efficient energy behaviors in the home, even when the younger members are away. However, it is important to remark that our findings are not statistically significant, and further research needs to be conducted to support these types of expensive policy propositions.

While our pilot design is not well suited to identify what specific contextual features limited the effectiveness of our intervention, we propose a number of directions for future research emerging from patterns in our data and/or supported by preexisting literature, that can further shed some light on the issues we have begun to explore in this study.

Firstly, it could be the case that the conditions of resource scarcity, cognitive strain, and stigmatization, which are common in social housing demographics, are simply too psychologically taxing to allow for the success of a top-down behavior change strategy aimed at reducing energy consumption. It could be that the level of cognitive strain associated from being in a condition of income and energy vulnerability is too large to lead to the required level of interaction with the technologies and initiatives designed to achieve behavioral change. While these issues are surely likely to have impacted the effectiveness of our intervention, it would seem dismissive to assume that they categorically impede social housing tenants from engaging with a behavior-change strategy. This would also be inconsistent with previous findings that have trialled different forms of behavior change interventions in social housing districts with moderate success (Hafner et al. (2020); Sangalli et al. (2020)). It would however be interesting to uncover to what degree these different psychological barriers stifle the adoption of more optimal energy decisions, and how each, in turn, could be addressed. A controlled laboratory could prove a suitable environment to provide further knowledge in this direction (Lunn and N´ı Choisdealbha, 2018).

Secondly, it is plausible that the lack of effectiveness of our intervention does not derive from the uniform design we have applied, but rather that tenants are simply not interacting with the IHD technology enough to be exposed to the normative appeals. Certainly increasing the level of interaction with the IHD technology would be unquestionably beneficial, not only to promote energy conservation but also to increase the agency of vulnerable individuals in the control of their energy consumption. The data however does seem to support this hypothesis as we find no statistically significant interaction effects when adding Number of clicks as a variable in our Bivariate regression models (Appendix). Further researcher would need to be conducted to investigate causality between the level of display interaction and intervention effectiveness. A larger population would have to be tested, perhaps in an laboratory setting (Lunn and N´ı Choisdealbha, 2018) where interaction with a similar interface could be experimentally varied, or in a field setting with an experimentally varied incentive structure that directly rewards interaction with the IHD. It would also be important to conduct further research that can identify behavioral determinants of interaction with IHDs, and find ways to influence tenants to interact further.

Thirdly, it may be that the integration of our intervention in the context of recently retrofitted homes is affecting its behavior-change potential. Evidence has shown that in social housing districts there exists a particularly high tendency to "take-back" a large proportion on energy savings after efficiency upgrades in the form of increased internal temperatures (Coyne et al., 2018). This behavioral response to the retrofitting is likely to be consistent across different forms of energy consumption, including electricity. Taken together with the evidence of a prebound effect prevailing in low-income populations (Sunikka-Blank and Galvin, 2012), it seems likely that our group of tenants (who are part of particularly vulnerable demographics and would be consuming below optimal levels of energy pre-retrofitting), increase their consumption post-retrofitting in order to appropriately meet their basic capabilities now that they can financially afford to do so. The
impact of this behavioral response to retrofitting on tenants' subsequent willingness to adapt their consumption downwards following our intervention is hard to measure with our available data (no pre-intervention electricity consumption information was available). However, it seems plausible to assume that tenants who have recently adjusted to a higher levels of consumption thanks to the efficiency upgrades, would be reluctant to then adjust their consumption downwards when presented with social comparison modules, especially if they were previously consuming sub-optimal levels of energy.

To our knowledge, there currently is no literature studying the impact of rebound and pre-bound effects on the effectiveness of subsequent social comparison interventions, so it is challenging to discern how important these effects are to our observed results. An interesting direction for further research would be to study social comparison interventions in a social housing context not having recently undergone refurbishment, and see if the results qualitatively differ.

It is important to note that just because the intervention did not lead to statistically significant differences in electricity consumption during the time-period investigated, this does not mean that the households were not consuming energy at an optimal level. Their current consumption levels may well have been conducive to them achieving their basic capabilities. Studying how changes in behavior following the retrofit affected the achievement of basic capabilities would be necessary in order to evaluate the success of our intervention from more of a capabilities approach. This would require a longer period of observation and a more qualitative study of household outcomes as the optimal levels of energy consumption needed to satisfy basic capabilities are likely to be highly subjective.

Finally, it may be that the homogeneity in the descriptive information communicated via the normative appeals, combined with a perception that the injunctive messages restrict individual freedom, caused tenants to disregard the social comparison entirely. In the application of social norm interventions, particular attention needs to be paid to target perceptions on the types of messages transferred, with particular attention to those situations where the information prescribed in the normative message can be perceived to restrain individual freedom (DellaValle, 2019). For example, tenants may perceive that they are supposed to maintain a similar, perhaps sub-optimal level of energy consumption to conform to the social norm, and that the normative information (in particular the injunctive norms) restrict their individual freedom to choose their own consumption. If there is such a negative perception to the injunctive appeals, they will likely be ignored. This can be a problem if tenants are consuming electricity at a level below, or roughly very similar to that of the reference group. If the comparisons that are presented to tenants (eg: the top two performers in their assigned reference group) are not significantly different than their own level of consumption, and they disregard the information conveyed by the normative appeals, tenants might reasonably choose to not alter their consumption level. Particularly in this small sample (where reference groups were composed of 4 apartments at most) it is very likely that the descriptive normative information was not different enough from the individual tenants’ own behavior to cause a significant behavioral change. This is supported by our data, where we see a low dispersion of electricity consumption (averaged at the hourly level) amongst clusters (Standard deviation = 0.264-0.315).

While such a methodological limitation could be overcome in a larger RCT with reference groups composed of a large number of households, where the top two performers have extremely virtuous behaviors, there is evidence suggesting the most vulnerable are particularly likely to form energy inefficient behaviors (DellaValle, 2019; Kearns et al, 2019), meaning any comparison group formed solely of social housing tenants, even if the top energy performers in this category, can potentially backfire. Perhaps the most critical suggestion for further research emerging from our study would be to trial social comparison interventions in social housing districts with larger and more diverse
reference groups from a consumption stand-point. For example, the reference social information that social housing tenants receive could be a general national average of energy consumption across different social housing districts, or a particularly virtuous example of a specific social housing district. A clear disadvantage of this approach however, is that this more general information can reduce how much a tenant identifies with the reference group, and decrease how salient the messages are. Studying these issues more in-depth could provide invaluable practical advice to policy-makers in order to design interventions containing the most effective information to leverage behavior.

5.2 Pilot Study Limitations

Due to the field nature of our pilot study, we encountered a number of limiting factors during implementation and in the scope of our study which should be addressed in future research. These limitations are worth discussing in order to better interpret the study’s outcomes and outline future approaches to more closely determine the effect of social comparison interventions in social housing contexts.

At its current stage, the scope of this study is narrow, as we chose to focus on electricity consumption only, over a 3 month period and in a limited number apartments. This narrow scope was taken for three reasons: (i) existing data limitations as there was a delay of apartment display installations following COVID-19, (ii) in order to focus on short-term effects as these are the most relevant for behavior change and (iii) to emphasize the methodological and analytical aspects of our study, so as to serve as a reference point for a larger, more extensive analysis once more data is available. While it is plausible that expanding the scope of analysis by including more apartments, studied over a larger period of time, and additionally considering thermal behavior would result in the identification of significant differences in energy consumption between groups, we have no reason a priori to believe that this will be the case. These concerns however are certainly valid from an analytical point of view and a more exhaustive analysis of the available data will be carried out in a forthcoming study in order to draw rigorous conclusions that can direct policy-making. Furthermore, while in this study we focused on the effects of our intervention on overall electricity consumption, further research could look at how similar interventions affect energy profiles throughout the day, as it may be that the intervention does not reduce overall energy consumption but rather shifts energy habits and consumption patterns across the day, which could have associated environmental and financial benefits for a society (especially in the case of variable energy tariffs).

A common implication of field studies such as ours is that participation is voluntary, creating the possibility that our sample is non-representative of the wider population we intend to study as there may be some systematic relationship between participation and some unobservable characteristics, leading to a self-selection bias. The potential for a positive self-selection bias is well researched in field experiments (Gautier and Klaauw 2012), as well as in the case of our specific type of intervention (Allcott 2015).

In our study it is possible that the households which self-select to allow the display installation are particularly prone to have a higher pro-environmental attitude (the mean for our elicited measure of pro-environmental self-identity is 6.58, considerably higher than the median of 4.5) or be more likely to be willing to take control of their energy consumption. This is indeed a limitation of our study as it threatens to reduce the external validity of our findings. This showcases one of the reasons why it is methodologically demanding to run a field study of this type in a social housing context. Other reasons include low participation in the program, as well as limited interaction with the employed technologies. All these issues should be tackled in further research, although it may be hard to do so in a field setting. Laboratory experiments targeting specific demographics of interests could be a useful methodological contribution to help overcome the shortcomings that
come from a field implementation. Additionally, to complement the internal validity of lab findings with the high external validity of field implementations, these studies could be complemented with field applications that adopt appropriate and salient incentives to the installation and interaction with the IHD technology, targeting a wider population of households and ensuring interaction with the technology necessary for implementation.

Additionally, the apartment-level randomization that took place within the district has the potential of leading to negative spillover effects. These spillover effects could occur as a result of communication between tenants in different groups. For example, if tenants in the control group become aware that other tenants have displays that show social comparison modules, or even observe each other’s displays, this could potentially affect the way they behave, biasing our results. While a better option would have been to randomize treatment assignment at the building or district-level, the existing timeline of the project as well as other technological considerations made this impossible.

Finally, we encountered some issues at time of running our regression analysis while using random effects GLS methods. As previously explained, a number of the variables included in our full regression model are based on individual-level measures obtained from survey responses. This was done because it proved too intrusive and methodologically complex to try and obtain survey answers from all occupants in the dwellings. We opted instead to get answers from one of the occupants present at time of installation. These variables were treated in our analysis as proxies for household-level preferences. While there is no reason to believe answers to the survey are not representative of household preferences, this approach is certain to produce household-level measures which are subject to the individual bias in preferences of the occupant answering, introducing some individual-level bias to our regression results. Moreover, despite basing the survey items on experimentally validated items following (Falk et al., 2018), there is still scope for hypothetical bias in our survey answers. Finally, there is little variation in the answers to some of the survey items (ReciprocityPref St. Dev = 0.674, TimePref St. Dev = 1.794, EnerLit St. Dev = 1.128), which together with the relatively small sample, limits the significance of our survey-elicited variables and the results obtained from the corresponding model estimation.

Therefore, answers from our estimated MV(2) model are to be considered with these methodological shortcoming in mind. The results for this model show a negative, significant treatment effect for our intervention once these preferences are controlled for. Given the opposition of this finding to our average-level results, our DID approach, and our previously estimated models, we should not take this result as proof that our intervention was successful overall. However, the significance level of this coefficient, as well as the highly significant impact in our model of some of the preference variables included, are indication that personal preferences, which are often disregarded in large RCTs, should absolutely be collected and studied in the context of how they affect the effectiveness of behavior change interventions.

6 Conclusions

In this study, we have presented a methodology designed to integrate a popular behavior-change intervention in the context of social housing retrofits, with the aim of addressing social and behavioral elements of energy efficiency improvements that are often overlooked in a social housing context. We introduced our pilot field study, based on a wealth of previously successful social comparison interventions in the energy domain, and discussed why social housing tenants make for a particularly interesting and important case study, due in part to their high level of energy vulnerability and potential to fall in energy poverty. Our primary aim throughout has been to uncover whether

Electronic copy available at: https://ssrn.com/abstract=3659866
this popular intervention could be applied in a standard way within a social housing context, with its unique difficulties and characteristics.

With the initial analysis of electrical energy consumption suggesting an insignificant effect of our intervention on electricity consumption, our results suggest prudence on the part of policy-makers when applying these behavior-change interventions in vulnerable demographics. Interventions of this kind, especially if delivered using the IHD technology, can be costly to implement. If their effectiveness in social housing are miscalculated and overstated based on the evidence of RCTs on a more general residential area, the costs of the intervention could far outweigh its actual benefits. This echoes findings from Andor et al. (2018) who find that the benefits of social comparison interventions may be overstated in European populations that consume relatively less energy than the general US public (where the highly researched OPower RCTs took place), making their indiscriminate implementation potentially unfruitful when the costs and benefits are fully accounted for. Policy-makers might alternatively wish to initially implement more cost-effective interventions in that are less cognitively taxing for vulnerable demographics to engage with, and instead boost the competencies of vulnerable individuals, so as to empower them to make more optimal energy decisions. Of course there may be benefits related to the use of the display, other than reduced energy consumption caused by social comparison modules, which would make their installation cost-effective. Further research could take a more holistic approach and study the benefits of IHD devices on different dimensions to better evaluate the effectiveness of the display installation as a whole.

Overall, while our results at present are somewhat limited from data availability and a narrow research scope, the methodological basis we introduce with this study enriches the emerging field of applied behavior-change interventions in social-housing districts. This field not only has immense practical importance for policy-makers wishing to leverage virtuous behavior in the context of efficiency upgrades of the social housing dwelling stock, but is also deeply important for discussions on energy justice and the tackling of energy poverty. If research in this area can identify ways that behavior-change interventions could be designed to be mindful of the contextual situations of the most vulnerable, we could ensure that behavior is effectively leveraged together with technical upgrades, in order to improve the capabilities of the most vulnerable and tackle energy poverty.

Acknowledgments

We would like to acknowledge the colleagues of the Urban and Regional Energy System at the Institute for Renewable Energy at EURAC Research who supported this research by providing their feedback. A special thanks goes to Amy Segata, Aaron Estrada, Pietro Zambelli, Marco Castagna, Jessica Balest, Alyona Zubaryeva and Daniele Vettorato without whom the visual design and the implementation of the display would have been impossible. The research leading to these results has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 609019 . The European Union is not liable for any use that may be made of the information contained in this document, which is merely representing the authors’ view.

Author contributions

All authors provided input into the draft and final manuscript, including conceptualization and writing–original draft preparation. C.N. compiled the literature review and was in charge of the data curation and formal analysis. DV.N. was in charge of overall supervision.
References

Abrahamse, W., Steg, L., Vlek, C., and Rothengatter, T. (2005). A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, 25(3):273–291.

Alberts, G., Gurguc, Z., Koutroumpis, P., Martin, R., Muûls, M., and Napp, T. (2016). Competition and norms: A self-defeating combination? *Energy Policy*, 96:504–523.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9-10):1082–1095.

Allcott, H. (2015). Site Selection Bias in Program Evaluation *. *The Quarterly Journal of Economics*, 130(3):1117–1165.

Anderson, K., Song, K., Lee, S., Krupka, E., Lee, H., and Park, M. (2017). Longitudinal analysis of normative energy use feedback on dormitory occupants. *Applied Energy*, 189:623–639.

Andor, M., Gerster, A., Peters, J., and Schmidt, C. M. (2018). Social Norms and Energy Conservation Beyond the US. *SSRN Electronic Journal*.

Andor, M. A. and Fels, K. M. (2018). Behavioral Economics and Energy Conservation – A Systematic Review of Non-price Interventions and Their Causal Effects. *Ecological Economics*, 148:178–210.

Bager, S. and Mundaca, L. (2017). Making ‘Smart Meters’ smarter? Insights from a behavioural economics pilot field experiment in Copenhagen, Denmark. *Energy Research & Social Science*, 28:68–76.

Banerjee, A. V. and Duflo, E. (2009). The Experimental Approach to Development Economics. *Annual Review of Economics*, 1(1):151–178.

Becchetti, L., Corrado, L., and Conzo, P. (2013). *Sociability, Altruism and Subjective Well-Being*.

Bicchieri, C. (2005). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press, Cambridge.

Bicchieri, C. and Dimant, E. (2019). Nudging with care: the risks and benefits of social information. *Public Choice*.

Blasch, J., Boogen, N., Filippini, M., and Kumar, N. (2017). The Role of Energy and Investment Literacy for Residential Electricity Demand and End-use Efficiency. *SSRN Electronic Journal*.

Bouzarovski, S. (2014). Energy poverty in the European Union: landscapes of vulnerability: Energy poverty in the European Union. *Wiley Interdisciplinary Reviews: Energy and Environment*, 3(3):276–289.

Brekke, K. A. and Johansson-Stenman, O. (2008a). The behavioural economics of climate change. *Oxford Review of Economic Policy*, 24(2):280–297.

Brekke, K. A. and Johansson-Stenman, O. (2008b). The behavioural economics of climate change. *Oxford Review of Economic Policy*, 24(2):280–297.

Bénabou, R. and Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, 96(5):1652–1678.
Casal, S. and Valle, N. D. (2016). Feedback and consumption behavior. *CEEL Working Papers*, 1608:21.

Chen, V. L., Delmas, M. A., Kaiser, W. J., and Locke, S. L. (2015). What can we learn from high-frequency appliance-level energy metering? Results from a field experiment. *Energy Policy*, 77:164–175.

Clee, M. A. and Wicklund, R. A. (1980). Consumer Behavior and Psychological Reactance. *Journal of Consumer Research*, 6(4):389–405.

Coyne, B., Lyons, S., and McCoy, D. (2018). The effects of home energy efficiency upgrades on social housing tenants: evidence from Ireland. *Energy Efficiency*, 11(8):2077–2100.

Day, R., Walker, G., and Simcock, N. (2016). Conceptualising energy use and energy poverty using a capabilities framework. *Energy Policy*, 93:255–264.

DellaValle, N. (2019). People’s decisions matter: understanding and addressing energy poverty with behavioral economics. *Energy and Buildings*, 204:109515.

DellaValle, N., Bisello, A., and Balest, J. (2018). In search of behavioural and social levers for effective social housing retrofit programs. *Energy and Buildings*, 172:517–524.

DellaValle, N. and Sareen, S. (2020). Nudging and boosting for equity? Towards a behavioural economics of energy justice. *Energy Research & Social Science*, 68:101589.

DellaValle, N. and Zubaryeva, A. (2019). Can we hope for a collective shift in electric vehicle adoption? Testing salience and norm-based interventions in South Tyrol, Italy. *Energy Research & Social Science*, 55:46–61.

Dolan, P. and Metcalfe, R. (2013). Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation. Technical report, Centre for Economic Performance, LSE.

Ehrhardt-Martinez, K., Donnelly, K., and Laitner, J. (2010). Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities. *American Council for an Energy-Efficient Economy*.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., and Sunde, U. (2018). Global Evidence on Economic Preferences*. *The Quarterly Journal of Economics*, 133(4):1645–1692.

Farrow, K., Grolleau, G., and Ibanez, L. (2017a). Social Norms and Pro-environmental Behavior: A Review of the Evidence. *Ecological Economics*, 140:1–13.

Farrow, K., Grolleau, G., Ibanez, L., and Mzoughi, N. (2017b). SOCIAL NORM INTERVENTIONS AS AN UNDERAPPRECIATED LEVER FOR BEHAVIOR CHANGE IN ENERGY CONSERVATION. *The Journal of Energy and Development*, 43(1/2):235–249.

Faruqui, A., Sergici, S., and Sharif, A. (2010). The impact of informational feedback on energy consumption—A survey of the experimental evidence. *Energy*, 35(4):1598–1608.

Ferraro, P. J., Miranda, J. J., and Price, M. K. (2011). The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment. *American Economic Review*, 101(3):318–22.
Fischer, C. (2008). Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency*, 1(1):79–104.

Galvin, R. (2014). Making the ‘rebound effect’ more useful for performance evaluation of thermal retrofits of existing homes: Defining the ‘energy savings deficit’ and the ‘energy performance gap’. *Energy and Buildings*, 69:515–524.

Gautier, P. A. and Klaauw, B. v. d. (2012). Selection in a field experiment with voluntary participation. *Journal of Applied Econometrics*, 27(1):63–84.

Hafner, R. J., Pahl, S., Jones, R. V., and Fuertes, A. (2020). Energy use in social housing residents in the UK and recommendations for developing energy behaviour change interventions. *Journal of Cleaner Production*, 251:119643.

Hall, C. C., Zhao, J., and Shafir, E. (2014). Self-Affirmation Among the Poor: Cognitive and Behavioral Implications. *Psychological Science*, 25(2):619–625.

Hershfield, H. E. (2011). Future self-continuity: how conceptions of the future self transform intertemporal choice. *Annals of the New York Academy of Sciences*, 1235(1):30–43.

Horan, P. M. and Austin, P. L. (2014). The Social Bases of Welfare Stigma*. *Social Problems*, 21(5):648–657.

Kearns, A., Whitley, E., and Curl, A. (2019). Occupant behaviour as a fourth driver of fuel poverty (aka warmth & energy deprivation). *Energy Policy*, 129:1143–1155.

Khosrowpour, A., Xie, Y., Taylor, J. E., and Hong, Y. (2016). One size does not fit all: Establishing the need for targeted eco-feedback. *Applied Energy*, 184:523–530.

Koellock, P. (1998). Social Dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology*, 24(1):183–214.

Loewenstein, G. and Prelec, D. (1992). Anomalies in Intertemporal Choice: Evidence and an Interpretation*. *The Quarterly Journal of Economics*, 107(2):573–597.

Luhtinen, R. and Crocker, J. (1992). A Collective Self-Esteem Scale: Self-Evaluation of One's Social Identity. *Personality and Social Psychology Bulletin*, 18:302–318.

Lunn, P. D. and Ní Choideabhá, (2018). The case for laboratory experiments in behavioural public policy. *Behavioural Public Policy*, 2(1):22–40.

Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, 341(6149):976–980.

Mould, R. and Baker, K. J. (2017). Documenting fuel poverty from the householders’ perspective. *Energy Research & Social Science*, 31:21–31.

Newell, R. G. and Siikamäki, J. (2015). Individual Time Preferences and Energy Efficiency. *American Economic Review*, 105(5):196–200.

Sangalli, A., Pagliano, L., Causone, F., Salvia, G., Morello, E., and Erba, S. (2020). Behavioural Change Effects on Energy Use in Public Housing: A Case Study. In Littlewood, J., Howlett, R. J., Capozzoli, A., and Jain, L. C., editors, *Sustainability in Energy and Buildings*, pages 759–768, Singapore. Springer Singapore.
Schaffrin, A. and Reibling, N. (2015). Household energy and climate mitigation policies: Investigating energy practices in the housing sector. *Energy Policy*, 77:1–10.

Schleich, J., Faure, C., and Klohasa, M. (2017). Persistence of the effects of providing feedback alongside smart metering devices on household electricity demand. *Energy Policy*, 107:225–233.

Schultz, P. W. (1999). Changing Behavior With Normative Feedback Interventions: A Field Experiment on Curbside Recycling. *Basic and Applied Social Psychology*, 21(1):25–36.

Schultz, P. W., Estrada, M., Schmitt, J., Sokoloski, R., and Silva-Send, N. (2015). Using in-home displays to provide smart meter feedback about household electricity consumption: A randomized control trial comparing kilowatts, cost, and social norms. *Energy*, 90:351–358.

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

Schultz, W. P., Khazian, A. M., and Zaleski, A. C. (2008). Using normative social influence to promote conservation among hotel guests. *Social Influence*, 3(1):4–23.

Sen, A. (1999). *Development as Freedom*. Alfred Knopf, New York.

Shah, A. K., Mullainathan, S., and Shafir, E. (2012). Some Consequences of Having Too Little. *Science*, 338(6107):682–685.

Sorrell, S., University of Sussex, and Sussex Energy Group (2007). *The Rebound effect: an assessment of the evidence for economy-wide energy savings from improved energy efficiency*. UK Energy Research Centre, Sussex. OCLC: 243949582.

Sovacool, B. K., Axsen, J., and Sorrell, S. (2018). Promoting novelty, rigor, and style in energy social science: Towards codes of practice for appropriate methods and research design. *Energy Research & Social Science*, 45:12–42.

Sunikka-Blank, M. and Galvin, R. (2012). Introducing the prebound effect: the gap between performance and actual energy consumption. *Building Research & Information*, 40(3):260–273.

Tomm, B. M. and Zhao, J. (2016). Scarcity captures attention and induces neglect: Eyetracking and behavioral evidence. *Cognitive Science*.

Whitmash, L. and O’Neill, S. (2010). Green identity, green living? The role of pro-environmental self-identity in determining consistency across diverse pro-environmental behaviours. *Journal of Environmental Psychology*, 30(3):305–314.

Wolske, K. S., Gillingham, K. T., and Schultz, P. W. (2020). Peer influence on household energy behaviours. *Nature Energy*, 5(3):202–212.

Šćepanović, S., Warnier, M., and Nurminen, J. K. (2017). The role of context in residential energy interventions: A meta review. *Renewable and Sustainable Energy Reviews*, 77:1146–1168.
## Appendix

Table 10: Interaction Effects of Number of Clicks & treatment

| VARIABLES                  | (1)       | (2)       |
|----------------------------|-----------|-----------|
|                            | ener_cons | ener_cons |
| 1.socialcompar             | 0.00384   | 0.00268   |
|                            | (0.194)   | (0.340)   |
| N_clicks_electrical        | -0.00187  | -0.00191  |
|                            | (0.00413) | (0.00725) |
| socialcompar#N_clicks_electrical | -0.00547 | -0.00544  |
|                            | (0.0314)  | (0.0551)  |
| Constant                   | 0.267***  | 0.176     |
|                            | (0.0867)  | (0.153)   |

Observations: 17,652 17,652  
Number of apt_id: 8 8

Standard errors in parentheses  
*** p<0.01, ** p<0.05, * p<0.1

Note: Dependent variable is energy consumption of apartment i at hour t (t=1-2233). BV(2) also controls for hourly and monthly variations by including hour and month dummies in the regression (suppressed in output). We do not include MV(1) and MV(2) as the output is omitted due to issues of multicollinearity.
Table 11: Model Estimation including dwelling size as an explanatory variable

| VARIABLES                        | (1)          | (2)           | (3)           | (4)           |
|----------------------------------|--------------|---------------|---------------|---------------|
|                                  | BV(1)        | BV(2)         | MV(1)         | MV(2)         |
| socialcompar                     | 0.0319       | 0.0315*       | 0.0174**      | -0.0611***    |
|                                  | (0.0436)     | (0.0184)      | (0.00773)     | (0.00862)     |
| refurbishment_entry_type         | -0.133***    |               |               |               |
|                                  | (0.00667)    |               |               |               |
| occupant_number                  | 0.0711***    |               |               |               |
|                                  | (0.0183)     |               |               |               |
| mean_house                       | 0.00182      |               |               |               |
|                                  | (0.00123)    |               |               |               |
| mean_age                         | 0.00625***   |               |               |               |
|                                  | (0.000849)   |               |               |               |
| N_children                       | -0.133***    |               |               |               |
|                                  | (0.0124)     |               |               |               |
| N_mature                         | -0.114***    |               |               |               |
|                                  | (0.0153)     |               |               |               |
| N_female                         | 0.00362      |               |               |               |
|                                  | (0.0140)     |               |               |               |
| size_m2                          | 0.00415***   | 0.00415***    | 0.00525***    | -0.00881***   |
|                                  | (0.00121)    | (0.000512)    | (0.000250)    | (0.000897)    |
| env_identity                     |               | -0.0526***    |               |               |
|                                  |               | (0.0153)      |               |               |
| trust                            | 0.0319***    |               |               |               |
|                                  | (0.00351)    |               |               |               |
| reciprocity                      | -0.229***    |               |               |               |
|                                  | (0.0159)     |               |               |               |
| altruism                         | -0.113***    |               |               |               |
|                                  | (0.00980)    |               |               |               |
| time_pref                        | 0.129***     |               |               |               |
|                                  | (0.00834)    |               |               |               |
| ener_lit                         | -0.147***    |               |               |               |
|                                  | (0.0103)     |               |               |               |
| group_identity                   | 0.0348***    |               |               |               |
|                                  | (0.00348)    |               |               |               |
| Constant                         | -0.0562      | -0.144***     | -0.526***     | 1.873***      |
|                                  | (0.0941)     | (0.0407)      | (0.0634)      | (0.185)       |
| Observations                     | 26,576       | 26,576        | 26,576        | 22,160        |
| Number of apt_id                 | 12           | 12            | 12            | 10            |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Dependent variable is energy consumption of apartment i at hour t (t=1-2233). BV(2), MV(1), and MV(2) control also for hourly and monthly variations by including hour and month dummies in the regression (suppressed in output). MV(2) drops observations from two apartments for which we do not have survey data.

Electronic copy available at: https://ssrn.com/abstract=3659866