A Field Experiment on Pro-Social Nudges: The Snowballing Problem

Presented to the faculty of Lycoming College in partial fulfillment of the requirements for Departmental Honors in Economics

By
Charlie Hunt
Lycoming College
05/05/2023

Approved by:

(Dr. Michael Kurtz, PhD)

(Dr. Elizabeth Moorhouse, PhD)

(Dr. Tina Norton, PhD)

(Dr. Mel Zimmerman, PhD)

A Field Experiment on Pro-Social Nudges: The Snowballing Problem

Charles Hunt
Lycoming College

May 3, 2023

This work was supported by the Joanne and Arthur Haberberger Fellowship, awarded through Lycoming College, Williamsport, PA

Committee Members:

Dr. Mica Kurtz, chair
Dr. Elizabeth Moorhouse
Dr. Tina Norton
Dr. Mel Zimmerman

Abstract

In contrast to traditional economic methods to address negative externalities, the study of pro-social nudges within the field of behavioral economics is quickly developing (Carlsson et al., 2019, Carlsson et al., 2021, Schubert 2017). The aim of pro-social nudges is to reduce a negative externality to a defined group or community. This study uses a natural field experiment to examine the impact of a pro-social nudge on a local negative externality of hot rooms in college residence halls and the resulting more global externality of excess greenhouse gas emissions. These externalities are the result of the interplay between the heating system and the propensity of students to open their windows causing the system to produce more heat to everyone—labeled the snowballing problem. This study suggests the nudge did not reduce the negative externalities, rather it may have backfired and exacerbated the existing problem (room temperatures increased 0.5-1.6F° after the intervention). As the results illustrate, addressing negative externalities with pro-social nudges may be particularly challenging because they often target behavior that benefits a larger society and requires the individual to experience some short-term disutility.

1.0. Introduction

Traditional methods to address negative externalities range from direct government intervention through regulations or taxes (e.g., Pigouvian tax) to more nuanced market-based solutions, such as emissions capping and trading. Yet, with the recent rise of the field of behavioral economics, and an ever-increasing demand for innovative solutions to collective action problems, the adoption of behavioral interventions coined "nudges" to combat these problems has expanded (Carlsson et al., 2019, Carlsson et al., 2021, Schubert 2017). A nudge is a small intentional change in the choice environment of an individual that usually utilize cognitive heuristics and can influence behavior in a meaningful way without restricting freedom of choice (Thaler and Sunstein 2009).

Conventional nudges are used to shift an individual's behavior for their own benefit, addressing negative *internalities* of an individual. Nudges that implement automatic enrollments in defined-contribution plans to boost retirement savings for under-contributors to their optimal life-cycle savings, is a prominent example. (Thaler and Benartzi 2004). On the other hand, a growing subset of nudge theory involves negative *externalities,* labeled "pro-social nudges" (Hagman et al., 2015). The goal of a pro-social nudge is not to alter the behavior of an individual for their immediate benefit, but rather to change a behavior to reduce a negative externality to a defined group or community. The largest area of research intersecting the field of pro-social nudges, is the related study of green nudges: nudges that reduce a negative environmental externality to society (Carlsson et al., 2019, Carlsson et al., 2021). Many pro-social and green nudges utilize social norms as the underlying cognitive heuristic (Schultz et al., 2007; Allcott, 2011; Allcott and Mullainathan 2010; Allcott and Rogers 2014).

This paper first examines the interaction between the heating systems of two residence halls at a small residential college and student's tendency to open or close their windows. Broadly speaking, in a process termed the "snowballing problem," as certain windows are opened in a hall, the heating system overcompensates and produces more heat to all the rooms in that hall, leading to more windows being opened, perpetuating the cycle. This positive feedback loop compounds into a local negative externality of hot rooms and a more global negative externality of excess greenhouse gas emissions. These externalities allow for an interesting opportunity to test the short-term impact of a pro-social nudge—intended to steer students to keep their windows closed—in a natural field experiment design.

Overall, the results of this study indicate the nudge did not reduce the negative externalities. In fact, there is some evidence the nudge backfired, leading to more open windows, hotter rooms, and increased natural gas usage. This highlights the fact that pro-social nudges using social norms may be particularly ineffective since they nudge behavior that often requires the individual to experience some short-term disutility.

The paper proceeds as follows. Section 2 provides background on the snowballing problem. Section 3 reviews the relevant literature of nudge theory, pro-social nudges, green nudges, and the larger literature of social norms. Section 4, 5, and 6 detail the design and implementation of the nudge, the hypotheses of the experiment, and the empirical data used. Section 7, 8, and 9 explain the methodology, empirical results and discussion, and limitations of the study. Section 10 concludes.

2.0. The Snowballing Problem

The onset of the cold months for many freshmen living in two residence halls means one perhaps counterintuitive reality—hot rooms. Reports of hot rooms from freshmen students and Residential Life are a well-documented issue (Schappert et al. 2018). Broadly, the problem of hot rooms occurs when students in certain rooms in a residence hall open their windows causing the heating system to overcompensate and produce excess heat to all the rooms in the residence halls. Specifically, the two residence halls in this study are spilt into two zones down the middle. Each zone has 4 rooms with a temperature sensor installed. The “Average Zone Temperature” (AZT) is the average temperature of those 4 sensors in each zone. During the onset of the cold months, the heating system for the two residence halls is set to a “Target Zone Temperature” (TZT): a temperature the 4 sensors in each zone should average out to. When a student in a sensor-room opens their window, the AZT decreases below the TZT leading the heating system to turn on for all rooms in that zone until the AZT has increased once again to the TZT. Once the AZT reaches the TZT, the heating system shuts off.

As room temperatures rise in a zone, more students may open their windows, and once another student in a sensor room opens their window, the AZT again temporarily decreases below the set TZT, and the heating system turns on once more to reach the TZT. Critically, when the AZT is below the TZT, heat is generated to *all* the rooms in that zone at the same rate, raising the temperatures of all rooms regardless of the presence of a sensor or if a window is opened or closed. The rooms with open windows, however, experience a comparatively lower rate of temperature increase than those with closed windows (since cold air from outside is coming in).

From a system dynamics perspective, this process of feedback between opening a window and heat increasing in the rooms is a positive feedback loop. That is, opening a window could lead to hotter temperatures, which, in turn, leads to more open windows, and the cycle snowballs, thus, the “snowballing problem.”

Importantly, the students who open their windows are not bearing the full costs of their actions. Instead, students with opened windows are imposing their costs—in the form of hotter rooms—onto students who keep their windows closed. This is the local negative externality. Another, more global, environmental negative externality occurs from the excess gas usage that happens when the heating system produce excess heat. This paper intends to model the snowballing problem and examine the impact of a prosocial nudge on the resulting externalities.

3.0. Literature Review

3.1. Behavioral Nudge

Through Richard Thaler and Case Sunstein’s best-selling book, “Nudge: Improving decisions about health, wealth, and happiness”, the term “nudge” has cemented itself into the lexicon of many policy makers and behavioral economists. A “nudge” is defined as a subtle intentional change in the decision-making environment—termed the choice architecture—that does not restrict options and has a profound impact on behavior (Thaler and Sunstein 2009). Traditional economic theorizing conforms to the neoclassical model of rational choice and would advocate market-incentives to alter behavior. Extensive literature within behavioral economics, however, demonstrates economic actors are reactive not only to price-incentives but to other psychological mechanisms, including, among many others, status quo bias (Bruns et al., 2018), salience (Noggle 2018), sunk costs (Haita-Falah 2017), and social norms (Legros and Beniamino 2020).

Nudges, borrowing heavily from the field of behavioral psychology, usually exploit certain cognitive biases or heuristics in the presence of uncertainty that give rise to bounded rationality, the idea that individuals make decisions that are satisfactory in the moment rather than optimal in the long-term (Kahneman 2011). Importantly, well formulated nudges are not mandates; they don't force their will on the recipient. Furthermore, a nudge does not adjust any monetary incentive or limit available options to the individual. Gruesome pictures on cigarette boxes (Fong et al., 2009), mail-reminders to pay your taxes (Hallsworth et al., 2017), and hotel towel reuse signs (Goldstein et al., 2008) are all instances of nudges. One of the most well-known examples of a nudge is the "Save More Tomorrow" program from Thaler and Benartzi (2004), which advocated automatic enrollments in defined-contribution plans to increase the retirement savings rate for individuals who save under their optimal life-cycle savings rate.

3.2. Pro-Social Nudge and Negative Externalities

Conventionally, nudges are used in areas where individuals have limited experience and lack critical information and help steer individuals away from irrational behavior, leading to poor long-term decisions (bounded rationality), such as in the "Save More Tomorrow" program. In another example, a study from Hanks et al., (2012) made healthier foods more salient by creating a second more "convenient" cafeteria line for healthier foods (as opposed to the line for unhealthy foods) —leading to 18% uptick in the sales of healthy foods. Assuming rational behavior, the long-term benefits of healthy foods normally outweigh the short-term satisfaction of the unhealthy foods. These nudges are used to alter an individual's choice for their own benefit. That is, they are generally dealing with negative *internalities* and are aptly named pro-self nudges (Hagman et al., 2015). However, adoption of nudges to reduce negative *externalities* to society—pro-social nudges—are increasingly being evaluated (Carlsson and Johansson-Stenman 2019). The purpose of a pro-social nudge is not to correct individual mistakes, but rather to reduce a negative externality. In fact, the nudged individual might experience immediate disutility since they are being nudged to reduce an activity that, while self-beneficial, creates a negative externality for society.

3.3. Green Nudge

Nudges that reduce a negative environmental externality—green nudges—have started competing with traditional environmental regulation for a spot in a policymaker's toolbox (Carlsson et al., 2019, Carlsson et al., 2021). Market failures in environmental resource management are abundant (Nyborg 2018). Air and water quality and landscape are public goods; fisheries, forests, irrigation systems are common pool resources; noise and air pollution are externalities. Traditional methods to address externalities (such as a Pigouvian Tax) use prices, property rights, and other market-based incentives to discourage individuals from consuming or producing a good that generates a negative externality. Green nudges, on the other hand, change the choice environment of an individual by capitalizing on a cognitive bias or heuristic without altering monetary incentives. In comparison to traditional interventions, green nudges are particularly enticing since they usually involve little financial investment and avoid the thorny political obstacles of, for example, a pollution-tax (Schubert 2017). Green nudges have been applied in many areas of behavioral environmental conservation such as energy consumption (Allcott 2011), water conservation (Nayar 2017), carbon offsets for air travel (Tyers 2018), and recycling behavior (Czajkowski et al., 2019), with overall mixed results (See Schubert 2017 or Velez and Moros 2021 for complete summary).

3.4. Social Norms

One of the most prominent cognitive heuristics pro-social and green nudges utilize is a social norm. A social norm is defined as a customary rule of behavior helping guide our interactions with others (Lewis 1969). The study of social norms is multidisciplined in nature and envelops an extensive amount of research, often with conflicting conclusions and ongoing debates (see Legros and Beniamino 2020 for full review).

Critically, social norms are reliant on interdependent behavior; they influence a behavior in that the subject expects others to perform the behavior and thinks others believe they should perform the behavior (Bicchieri and Dimant 2019). More precisely, the behavior influenced by social norms is conditional to social expectations. These social expectations often exist in situations where there is a tension between one's own welfare and the welfare of the group; a tension that is fundamental to collective action problems resulting in negative externalities. Take traffic congestion as an illustration. An individual must decide to drive with their car or ride public transportation to work. Even though driving may be in the individual's best interest (e.g., faster commute, more comfortable), the decision to drive generally results in a larger negative externality of traffic congestion and carbon excess emissions than that of public transportation. In places where there is a social norm of public transportation—one expects their peers to ride public transportation and believes others think they should as well—more people may adhere to the norm and reduce the negative externalities. Indeed, a paper from Bamberg et al., (2007) found social norms in two large urban centers in Germany played a significant role in public transportation-use intentions. In general, it is well-documented, in some situations, individuals will imitate behaviors of their peers or conform to the majority behavior (Cialdini and Goldstein

2004). In these cases, a nudge promoting this social norm dependent behavior may prove to be effective.

3.5. Nudges Utilizing Social Norms

A nudge utilizing a social norm relies on eliciting social expectations by providing information about the group to redirect a behavior. For example, Hallsworth et. al., (2017) nudged UK taxpayers by including a short message on their standard reminder letters indicating most taxpayers pay their taxes on time, accelerating payment for overdue tax and raising £9 million in the first 23 days. In another more extreme example, the local government in Bogota, Colombia, hired 420 mime artists to mock traffic violators in the inner city, postulating the citizens feared being ridiculed more than the standard fine, which may have contributed to decreased traffic fatalities (Caballero 2004). These nudges primarily benefit society and not necessarily the individuals—pro-social in nature.

Nudges utilizing social norms generally can take the form of either descriptive or injunctive norms (Cialdini et al., 1991).[1] Descriptive norms are statements about the prevalence of peers' behaviors (e.g., announcing most taxpayers pay their taxes on time) while injunctive norms communicate approval or disapproval of a behavior (e.g., public mockery for committing a traffic violation). The respective effectiveness of these two types of social norms has been a long-standing debate within the larger literature of social norms, pro-social, and green nudges. A group of seminal papers evaluated a series of programs from the electricity provider OPOWER,

[1] In contrast to Cialdini et al., (1991), Bicchieri and Dimant (2019) make the argument that the traditional definitions for descriptive and injunctive norms are insufficient when considering a nudge intervention since the terms don't differentiate between preferences that are conditional or unconditional on social expectations (interdependent behaviors). They prefer the terms empirical and normative expectations respectively, which inherently rely on the assumption that the target behavior is conditional to social expectations (you expect others to do *x* and you believe that others think you should do *x*). In this paper, I keep to the original and widely used terminology of injunctive vs descriptive while explicitly stating that our target behavior is conditional to social expectations where appropriate.

where home energy report letters were sent to their customers containing various combinations of descriptive and injunctive social norms about energy consumption in their neighborhood (Allcott 2011; Allcott and Mullainathan 2010; Allcott and Rogers 2014). They estimated the average program reduced energy consumption by around 2% or equivalent to a short-run electricity price increase of 11 to 20% and showed the effect did not completely disappear over the course of two years even after the nudge was gradually discontinued (Allcott and Rogers 2014). The OPOWER papers found nudges using injunctive norms had a minimal impact on energy consumption and most of the variation in consumption could be accounted for by the descriptive norms. On the other hand, a paper from Schultz et al., (2007) demonstrated social information solely in the form of descriptive norms (providing information of the household's energy consumption as compared to the average) produced a "boomerang" effect where those households with low energy consumption increased their energy use if they were below the average. However, when the researchers provided an injunctive message (a smiley or frowny face), the boomerang effect was lessened. The distinction between injunctive and descriptive norms was expanded upon in a randomized controlled trial from Bonan et al., (2020), which argued descriptive and injunctive feedback has a more complementary relationship, with the greatest impact on energy conservation occurring through a combination of descriptive and injunctive norms.

An explanation for the heterogenous effects in the literature could lie in contextual and social dynamic effects of the varying studies. For example, a recent replication of the OPOWER programs in Germany, revealed a much smaller treatment effect since energy consumption is already relatively low compared to the US, highlighting the fact that nudges may only be cost-effective towards certain sub-groups or in certain situations (Andor et al., 2020).

Generally, both injunctive and descriptive norm nudges are thought to be a function of their saliency (how prominent or emotional something is) (Cialdini 2003), consistency (Kallgren et al., 2000), and their reflection of the target behavior's frequency. The latter is particularly important when distinguishing between positive and negative descriptive and injunctive norms in nudges; that is, conveying information about the high or low frequency of a target behavior or approval or disapproval of said behavior. In certain cases, a nudge may unintendedly promote a widespread sociably undesirable behavior. For example, a positive descriptive norm in the form of a sign to reduce theft in Arizona's Petrified Forest conveyed that there are high levels of bark theft, leading counterproductively to increases in theft, since people believed theft was more socially acceptable (an injunctive conclusion) (Cialdini et al., 2006). Indeed, as Bicchieri and Dimant (2019) pointed out, it is not easy to separate descriptive and injunctive norm feedback since simply providing positive or negative descriptive norm feedback (the prevalence of a behavior) could lead respondents to draw injunctive conclusions (judgement on that behavior) or that positive or negative injunctive feedback could imply descriptive judgments. In general, previous research has indicated nudges utilizing injunctive norms may be most effective when the socially undesirable behavior is widespread, while descriptive norms are powerful when the sociably desirable behavior is already the majority behavior (Bicchieri and Dimant 2019).

The "backfiring" of descriptive social norms in nudges (or what Stibe and Cugelman (2016) term "reverse norming"), like in Cialdini et al., (2006) or Schultz et al., (2007), can be avoided through a number of methods. One method is to present descriptive norm nudges only to a particular subset of individuals, such as high energy consumers (Kantola et al., 1984). Another method is to provide counteracting injunctive norms (as in Schultz et al. (2007)). Or finally, if the desired behavior is relatively frequent, simply accentuate the rate of behavior (which could result in an additional positive injunctive conclusion) as Goldstein et al. (2008) did with hotel towel reuse signs ("Almost 75% of guests who are asked to participate in our new resource savings program do help by using their towels more than once").

While debate on relative effectiveness of injunctive and descriptive norms in nudges is active, the impact of other aspects of social norms has more consensus. For example, an important aspect of nudges utilizing social norms is the perceived trustworthiness and authority of the messenger. For example, Hallsworth et al., (2016) indicated that when high antibiotic prescribing doctors in England were sent letters from the Chief Medical Officer, a high authoritative figure, with a leaflet describing that their practice was prescribing at a higher rate than 80% of practices in their area, rate of antibiotic prescription significantly decreased compared to those who just received informational material on the dangers of over-prescribing.

The reference network—the group of comparison—of a nudge utilizing a social norm is also critical to its success. In a laboratory setting, Bicchieri et al., (2021) showed individuals discount information about pro-societally when the reference group was too broad or undefined. Moreover, a more local reference network could be associated with higher adherence to the nudge. A wonderful example of this phenomenon comes from Hallsworth et al., (2017), who demonstrated a more local nudge (local area as opposed to country) was more effective at increasing tax compliance. These results are not surprising; after all, social norms are properties of groups and not an individual.

A great number of society's most pressing problems—the climate crisis for example—are the result of negative externalities whose impact will take generations to materialize. The unique situation of the snowballing problem whereby the negative externality of warmer rooms is realized relatively quickly, allows for an opportunity to model a pro-social nudge's effect on a negative externality on a small scale and could provide a meaningful contribution to the literature on pro-social nudging.

4.0. Intervention

4.1. Target Behavior

This study uses a pro-social nudge to address a negative externality in the form of a descriptive social norm nudge. Since understanding the underlying target behavior is critical in designing effective choice architecture (Bicchieri and Dimant 2019; Hauser et al, 2018), this study postulates the factors related to the decision of whether to open or close a window (hereby termed "window behavior") are threefold.

First, window behavior is a product of baseline personal preferences with regard to temperature and sound on a given day; physiological factors play a key role (e.g., if someone gets cold easier or is a deep sleeper with regards to noise). Even if the snowballing problem was non-existent (opening one's window did not result in hotter rooms for other students), there still might be some expected variation in window behavior in the two residential halls due to these factors. It is important to note, in this case, individuals are acting rationally and maximizing their utility based on their personal preferences.

Second, window behavior is influenced by the snowballing problem in the form of excess heat to the rooms. Simply put, as rooms with closed windows become hotter, the likelihood a window is opened/cracked increases as well.

Lastly, this study hypothesizes window behavior can be socially interdependent. That is, an individual's window behavior is conditional on the actions and beliefs of the respective reference network, and it could be, therefore, appropriate to use a nudge utilizing a social norm emphasizing window behavior.

The effects of the snowballing problem create two overall negative externalities. The more local negative externality is the excess heat itself that a third party—namely a room with a closed window—is subject to when another room opens their window. Due to this overproduction of heat supply to the rooms, the snowballing problem results in excessive natural gas usage which contributes to more global externality of excess greenhouse gas pollution. In both cases, individually rational behavior leads to a socially sub-optimal outcome.

4.2. Intervention Design

The intervention is this study is based on insights from previous literature on the optimal design of nudges. The exact wording of the nudge is as follows:

"Hi from Residential Life at [college]! The majority of the students in [building] keep their window closed. This prevents the heating system from running unnecessarily and keeps all rooms comfortable."

As mentioned in section 3, previous research has suggested descriptive nudges are more powerful than injunctive when the majority behavior is already sociably desirable (Bicchieri and Dimant 2019; Cialdini et al., 2006). The majority of students do keep their window closed (the desirable behavior), which suggests a nudge using a descriptive norm may be effective.

To shift behavior with social norms, interventions must create collective expectations within an individual's reference group. Since the locality of the reference group has been shown to be critical to the effectiveness of a social norm feedback (Bicchieri et al., 2021; Hallsworth et al., 2017; Heise and Manji 2016), the treatment residence hall is used as a reference network. The main idea is to provide a local enough reference network so the subjects value the opinions of the group. Additionally, since the perceived authority and authenticity of the messenger has been found to have a strong impact on the magnitude of a social norm nudge's influence (Stibe and Cugelman 2016, Hallsworth et al., 2016, Bicchieri et al., 2021), the nudge is sent from the Office of Residential Life—which is in charge of student life in the residence halls and has close relationships with the students.

Lastly, an informational component to the nudge ("*This prevents the heating system from running unnecessarily and keeps all rooms comfortable."*) is needed to establish a connection between the desired behavior (a closed window) and the outcome of interest (comfortable rooms). This is, in essence, highlighting the desired behavior as socially interdependent which is then reinforced by the earlier descriptive component.

5.0. Experiment Design

This study utilizes a natural field experiment design to assess the impact of the pro-social nudge. There are two freshmen residence halls that have similar heating systems (i.e., where the snowballing problem could occur in a similar way). The treatment building, whose students would receive the nudge, was established randomly. The study design includes a 22-day pre-period and a 28-day experimental period for a total of 50 days of analysis. The pre-period includes 4 days before the majority of students arrived on campus During the pre-period, baseline measures of all relevant outcomes were observed. During the experimental period, those outcomes continued to be measured and the nudge was delivered once a week from Residential

Life as a text message (a total of 4 times) on the same day and time to all students in the treatment building.

6.0. Data

The outcomes of interest are window behavior, room temperature, gas usage, student satisfaction, and student perception of the number of days their rooms were too hot/cold rooms in the last week (Table 1). Window behavior was recorded by visually inspecting the windows from outside each building each day at a similar time (around 8am). Only windows in an occupied student room were included in the analysis, not those in hallways or bathrooms. Windows of rooms with no students (treatment=4, control=19) and rooms with fixed AC-units (treatment=2, control=1) were also excluded. Window behavior was measured on a 0-2 scale (0=fully closed, 1=cracked to half open, and 2=half-open to fully open). A scale is used since there could be a considerable difference in the amount of cold air entering the room between a slightly cracked and a fully open window. Natural gas usage was also measured by visually observing the gas meters outside of the residence halls each day. The units for gas usage are CFF (hundred cubic feet). Natural gas usage was normalized by dividing the change in CFF between each measurement period by the hours elapsed. Room temperature data (F°) was collected by physically placing temperature-recorders (16 in each building) in a sample of rooms in the treatment and control buildings before the start of the semester. The temperature-recorders took measurements every two hours, resulting in 18944 individual measurements. The recorders were installed in identical locations in rooms along with a note indicating their purpose of temporarily measuring temperature levels (Image 1. in Appendix A).

Separate but identical surveys were sent to the treatment and control buildings both before and after the experimental period to determine student satisfaction with temperature comfort in their rooms ("Survey," Appendix B). Respondents who completed the survey were entered into a raffle to win a $25 gift card. The surveys asked the students to first determine how many days in the past week were too hot and too cold and then to indicate their overall satisfaction with the temperature comfort in their rooms (10-point Likert scale). The surveys had response rates of 39.5%, 36.2% ,18.1% and 18.4% for the treatment-before, control-before, treatment-after, and control-after the experimental period, respectively. Average daily outside temperature was obtained from the Wunderground resource bank (Wunderground). Basic summary statistics of the all the variables are found in Table 2 in Appendix B.

Student demographic variables, aggregated at floor level, were obtained from the college. Table 3 displays differences in means and proportion tests between between the treatment and control buildings for the demographic characteristics, with no statistically significant differences reported, leading to the conclusion that the treatment and controls samples are relatively similar on observable characteristics.

Table 1: Variables

Variable	Aggregation Level	Description	Source
Treatment		Treatment or control building (1 being treatment; 0 being control)	N/A
After		After or before first intervention (1 being after; 0 being before)	N/A
Day0-Day6		Number of days after a nudge is sent (i.e., Day0 is day of the nudge)	N/A
Avg. Outside Temp.		Average daily outside temperature (F°)	Wunderground
Room Temperature	Room	Temperature (F°) every 2 hours from room temperature recorders	32 temperature recorders
Window Behavior	Room	Daily window behavior (0 being closed; 2 being >half open)	Visual observation
Gas/hour	Building	Daily per hour change for each building (CCF/hr)	Visual observation
StudSat	Student	Average student satisfaction (1 being very low; 10 being very high)	Survey
RoomHot	Student	Average # hot rooms in the last week	Survey
RoomCold	Student	Average # cold rooms in the last week	Survey

Table 3: Summary of Demographics

	Count	Age (days)	GPA	% Male	% Undec.	% White	% Black	% Hispanic	% Other	% Internat.	% Athlete
Treatment	105	7035.31	3.05	52.40	72.40	61.90	12.40	15.20	10.50	5.70	45.70
		(23.50)	(0.08)	(0.05)	(0.04)	(0.05)	(0.03)	(0.04)	(0.03)	(0.02)	(0.05)
Control	141	7008.57	2.93	50.40	70.20	58.90	14.20	17.00	9.90	2.80	43.30
		(23.43)	(0.07)	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.01)	(0.04)
Difference		26.75	0.12	2.00	2.20	3.00	-1.80	-1.80	0.60	2.90	2.40
		(3.03)	(0.01)	(0.06)	(0.06)	(0.06)	(0.04)	(0.05)	(0.04)	(0.03)	(0.06)
Full Sample	246	7029.59	2.98	51.20	71.10	60.20	13.40	16.30	10.20	4.10	44.30
		(16.73)	(0.05)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)

Difference calculated as [Treatment-Control]. No statistically significant differences between groups were found.

7.0. Methodology and Empirical Results

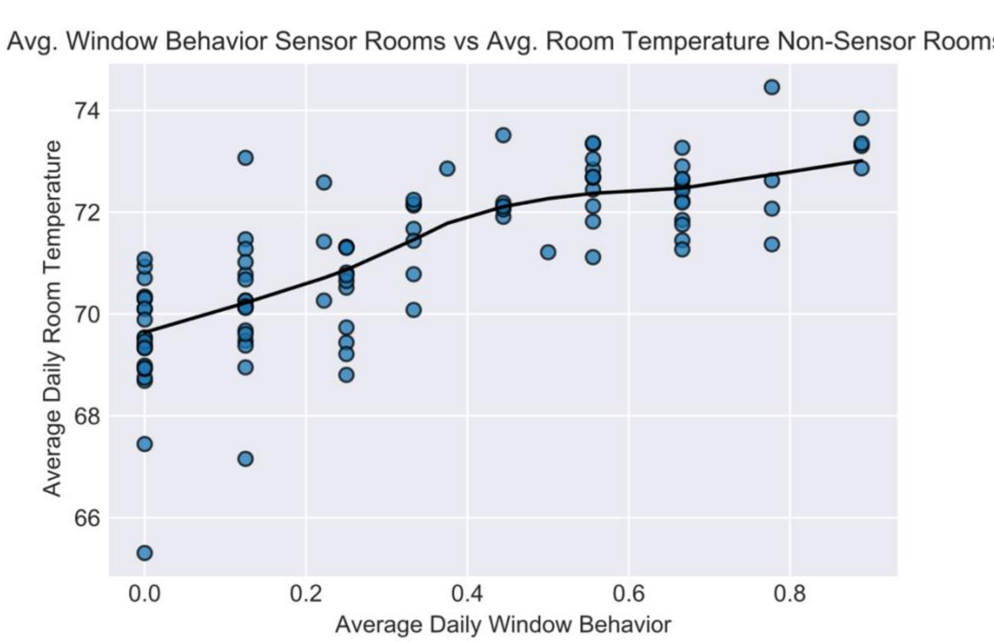
In the first part of the following section, an initial model of the snowballing problem is developed. The remaining section is structured to follow the potential causal pathway of the nudge, accessing the direct impact on window behavior and then the corresponding indirect effect on room temperature and gas usage.

7.1. Modeling the Snowballing Problem

As mentioned in section 2, the snowballing problem is mostly a factor of window behavior of sensor-rooms and results in increased temperatures to rooms in specific zones. With the limitations of the data, only a simplified version of the snowballing problem can be modeled. Namely, the concept that when students in sensor rooms open their windows, there are hotter rooms for the rest of the building.

Figure 1 depicts the average daily window behavior (total window behavior on a given day divided by the number of windows in a building) of sensor rooms against the average daily inner temperature of non-sensor rooms fitted to a LOESS curve. From the graph, there seems to be a positive relationship (r=0.768). That is, as average window behavior for the sensor-rooms increases, so does the average temperature for non-sensor rooms.

Figure 1:



For a more sophisticated analysis controlling for outside temperature, the snowballing problem can be modeled as:

$$temp\_nonsensor \ = \beta_0 + \beta_1 \ winbe\_sensor \ + \ \beta_2 avg\_outtemp$$

Where $temp\_nonsensor$, $winbe\_sensor$, and $avg\_outtemp$ equal the daily average temperature in non-sensor rooms, the daily average window behavior of sensor rooms, and the daily average outside temperature.

In model (1) of Table 4, window behavior of sensor rooms is statistically significant indicating for every 1 unit increase in average window behavior of (more windows are open) sensor rooms, there is an associated 4.510 F° increase in non-sensor rooms. To put this in more real terms, this coeffect indicates if all sensor room windows went from closed to cracked-half open, room temperature would increase by 4.510F°. Models (2) and (3) include average outside temperature which is slightly positively associated with temperature of non-sensor rooms, which is not entirely unexpected since as outside temperature increases, room temperature will also increase, although the main variation in temperature comes from the heating system.

Models (3) and (4) also include the average window behavior of the non-sensor rooms. Both these coefficients are not significant. This simply emphasizes the importance of the window behavior of the sensor rooms in determining the temperature of non-sensor rooms. Finally, model (4) modifies window behavior of sensor rooms, using an inverse hyperbolic sine transformation allowing for a more reasonable assessment of the coefficient. This transformation is very similar to a standard log transformation but is defined for values where [x=0.] Here, the model indicates every 10% increase in window behavior of sensor rooms was associated with a 0.182F° increase in temperature of non-sensor rooms, holding all else constant.

Table 4: Average Temperature Non-Sensor Rooms

Regression Model	(1)	(2)	(3)	(4)
Intercept	69.607***	68.468***	68.352***	67.959***
	(0.160)	(0.545)	(0.546)	(0.558)
Window Behavior Sensor Rooms	4.510***	4.443***	4.237***	
	(0.380)	(0.374)	(0.395)	
Avg. Outside Temp.		0.032**	0.023	0.029*
		(0.014)	(0.015)	(0.016)
Window Behavior Non-Sensor Rooms			1.55	0.785
			(1.007)	(1.049)
Modified In(Window Behavior Sensor Rooms)*				1.182***
				(0.113)
R-squared	0.590	0.609	0.619	0.608
F-statistic	141.172	75.689	51.926	49.658
N	100	100	100	100

Standard error reported in parentheses. *,**,*** denote conventional significance levels of 10%, 5%, and 1% respectively.

*transformed with (10x)—>ln(x+z$\sqrt{x^2+1}$)

Overall goodness of fit of the models is quite high. In model (1) $R^2 = 0.59$, indicating that much of the variation in average room temperature of non-sensor rooms is accounted for by the average window behavior of sensor rooms. In models (2), (3), and (4), including outside temperature only slightly increased the proportion of variation in average room temperature of non-sensor rooms that is captured.

It is important to point out window behavior is measured daily, so temperature must be averaged to that day and the assumption that window behavior is the same for that whole day is made. Naturally, window behavior likely changes more than once a day; any effects on temperature from a different combinations of window behavior throughout the day are thus hidden. Additionally, the exact frequency at which the sensors in the sensor-rooms take an AZT is unknown—though the college indicates it is less than every 2 hours. This would determine the frequency of the feedback loop, holding window behavior constant. Lastly, there are 32 temperature-recorders in the buildings, with only 16 temperature-recorders in non-sensor rooms. This limits the ability to model the full effect of the window behavior of sensor-rooms on the rest of the building.

7.2. Initial Analysis of Intervention

The hypothesized pathway of the intervention's impact is that the nudge decreases window behavior (more windows are closed) lowering room temperatures which would also decrease natural gas usage. This study utilizes a difference-in-difference design, meaning there is an expected greater drop (or less of an increase) in window behavior, room temperature, and gas usage for the treatment group as compared to the control after the experimental period begins.

The difference-in-difference methodology allows for the relaxation of the assumption of exchangeability and removes any biases in the post-intervention period that are the result of permanent differences between treatment and control groups. Nevertheless, to ensure eternal validity of the difference-in-difference model, the parallel trends assumption needs to be satisfied. This means in the absence of the intervention, the difference between the treatment and control group is constant over time. Figures, 2, and 3, and 4 show the average window behavior, room temperature, and gas/hour over the length of the experiment. Pre-trends trends between the treatment and control groups are relatively similar, giving some confidence similar trends would continue given no intervention.

Figure 2

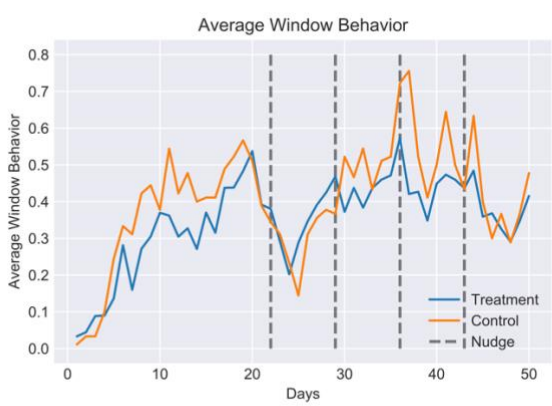


Figure 3

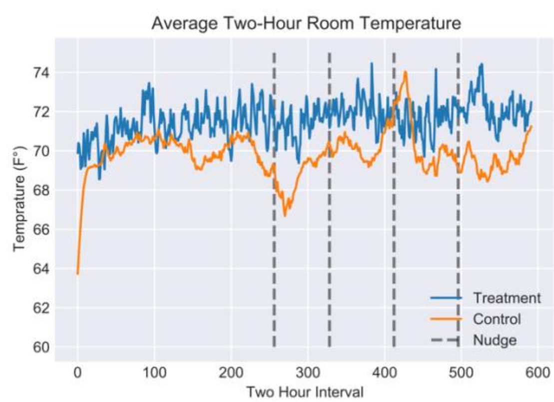
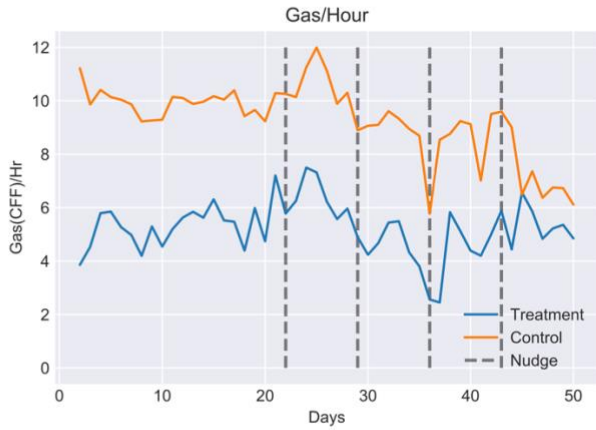


Figure 4



As shown in Figure 2, over the course of the experiment, average window behavior for the control is generally higher than that of the treatment, however, the respective movements in the trends are similar. In Figure 3, the room temperature in the treatment building is generally higher than that of control, being perhaps slightly higher after the intervention than before. Interestingly, around the time of the third nudge, there is a large spike in room temperature for the control. This logically corresponds to a large spike in average window behavior around the similar time. In Figure 4, the daily gas/hour changes (e.g., the 10th day represents the change in the gas meter from the 9th day divided by hours elapsed) are depicted. The control building begins with a much higher gas usage per hour than the treatment. This is expected since the control building has more square footage. However, by the end of the experimental period, gas/hour is much closer to that of the control. In contrast, the trend for gas/hour of the treatment building remains steadier over the entire period.

A first look of the data in a preliminary difference-in-difference table is presented in Table 5. As expected, the mean average window behavior in both the treatment and control groups increase after the intervention date since the pre-period includes days before students arrived on campus. While average window behavior, theoretically the most directly impacted outcome, showed an 0.035 *increase* in the treatment building after the intervention, it is not statistically significant. The difference-in-difference for room temperature is a statistically significant *increase* in room temperature by 0.613F° for the treatment group after the nudge was applied. In line with the direction of room temperature, there is a large statistically significant *increase* in mean gas/hour. Interpretated at face value, this indicates the intervention resulted in a 1.025CCF larger change in hourly gas usage than what would have otherwise occurred with no intervention.

While the survey responses certainly suffer from self-selection bias and relatively small sample size, the responses can be used to establish some students are indeed uncomfortable with the temperature in their rooms. Before the intervention, the mean satisfaction with their room temperature was 4.181 and 6.510 (10-point Likert scale) for the treatment and control respectively. The mean number of hot days in the past week for the treatment and control was 4.143 and 2.373 respectively.

Overall, these results suggest the nudge either had a negligible or a *reverse* effect as the original hypothesis postulated.

Table 5: Difference in Means

		Treatment			Control			
		Before	After	Difference	Before	After	Difference	Difference-Difference
Room Temperature		71.131	71.737	0.606***	69.688	69.681	-0.007	0.613***
	SE	(0.071)	(0.062)	(0.094)	(0.030)	(0.039)	(0.079)	(0.112)
	N	4096	5376		4096	5376		
Avg. Window Behavior		0.276	0.386	0.11***	0.294	0.368	0.074*	0.035
	SE	(0.029)	(0.014)	(0.030)	(0.030)	(0.022)	(0.033)	(0.047)
	N	22	28		22	28		
Gas Usage/Hour		5.334	5.153	-0.181	9.946	8.740	-1.207*	1.025***
	SE	(0.170)	(0.221)	(0.295)	(0.106)	(0.306)	(0.332)	(0.470)
	N	22.000	28.000		22.000	28.000		
Satisfaction		4.810	4.737	-0.073	6.510	7.115	0.606	-0.678
	SE	(0.309)	(0.540)	(0.585)	(0.266)	(0.352)	(0.207)	(0.726)
	N	51	26		42	19		
# Hot Days		4.143	4.158	0.015	2.373	1.231	-1.142*	1.157
	SE	(0.290)	(0.563)	(0.573)	(0.300)	(0.295)	(0.484)	(0.736)
	N	51	26		42	19		
# Cold Days		0.762	0.526	-0.236	0.608	0.346	-0.262	0.026
	SE	(0.228)	(0.269)	(0.385)	(0.210)	(0.207)	(0.332)	(0.505)
	N	51	26		42	19		

Standard error reported in parentheses. Difference calculated as [After-Control]. Difference-in-Difference calculated as [Dif. Treat-Dif. Cont]. *,**,*** denote conventional significance levels of 10%, 5%, and 1% respectively.

Note: "After" defined as all the days after the first nudge was sent

7.3. Regression Analysis

Instead of using average window behavior as the dependent variable, a logit regression model is used to model window behavior, with 0 being a window is closed and 1 being some level of openness. The baseline model is:

(1) $Window\ Behavior = \beta_0 + \beta_1\ Treatment + \beta_2 After + \beta_3 Treatment * After +\ \varepsilon$

Table 6 reports the average marginal effects of the models. Model (1) simply reports the marginal effects of the baseline model with the treatment effect of the nudge as (β_3) , with the cutoff into the experimental period being the day after the first nudge was delivered. Model (2) adds to the previous model by including average outside temperature and the student demographic data as controls described in Table 3.

Both models (1) and (2) show statistically significant treatment effects for the intervention (though only at the $[p = 0.1]$ level) with the average marginal effect changing little with the addition of the controls. As shown in model (2), the windows in the treatment building after the intervention were 3.5 percentage points more likely to be open than what would be expected given no intervention.

Table 6: AME Window Behavior

	Avg. Marginal Effect	
	(1)	(2)
Treatment	-0.047***	0.000***
	(0.016)	(0.000)
After	0.076***	0.062***
	(0.014)	(0.014)
Treatment:After	0.036*	0.035*
	(0.020)	(0.020)
Avg. Outside Temp.		0.005***
		(0.001)
Controls	No	Yes
N	8859	8859

Standard error reported in parentheses. *,**,*** denote conventional significance levels of 10%, 5%, and 1% respectively.

The baseline ordinary least squares model for room temperature is:

(1) $Room\ Temperature = \beta_0 + \beta_1\ Treatment + \beta_2 After + \beta_3 Treatment * After + \varepsilon$

The structures of models (1)-(3) and results are reported in Table 7. Model (1) is the baseline model representing the treatment effect (β_3) of the nudge after the intervention. Model (2) adds to the previous model by adding average daily outside temperature and student demographic controls. As demographics are aggregated at the floor level, the temperature recorders on each floor were matched with the corresponding floor demographics. Model (3) examines the effects of the nudge at a more granular level by comparing temperature levels in the treatment and control in the 6 days after each nudge was delivered (e.g., Day0 is the day the nudge was sent).

Model (1) shows that being in the treatment building after the nudge is associated with a 0.613F° higher room temperature than what would be expected without the nudge. The addition of average outside temperature and demographic controls in model (2) does not change the treatment effect. As expected, average outside temperature is positivity associated with room temperature: every 1F° increase in average outside temperature is associated with a 0.041F° increase in room temperature. Lastly, model (3) shows the effect of the nudge 0-6 days after the nudge was delivered. Being in the treatment building 2-6 days after a nudge is associated with 1.606-0.416F° higher room temperature compared to the control, with a diminishing effect each day.

Table 7: Room Temperature

	(1)	(2)	(3)
Intercept	69.688***	0.000***	0.000***
	(0.060)	(0.000)	(0.000)
Treatment	1.444***	0.064***	0.062***
	-0.084	(0.005)	(0.005)
After	(0.006)	-0.128	
	(0.079)	(0.079)	
Treatment:After	0.613***	0.613***	
	-0.112	(0.111)	
Avg. Outside Temp.		0.040***	0.045***
		(0.004)	(0.005)
D0:Treatment			0.211
			(0.210)
D1:Treatment			0.053
			(0.210)
D2:Treatment			1.606***
			(0.210)
D3:Treatment			1.167***
			(0.210)
D4:Treatment			0.910***
			(0.210)
D5:Treatment			0.712***
			(0.210)
D6:Treatment			0.416**
			(0.210)
Controls	No	Yes	Yes
R-squared	0.055	0.081	0.087
F-statistics	366.484	185.841	84.205
N	18944	18944	18560

Standard error reported in parentheses. *,**,*** denote conventional significance levels of 10%, 5%, and 1% respectively

*Note dummy coefficients for model (3) are not shown and controls are student demographic factors

The baseline ordinary least square equation for gas/hour is:

(1) $GasPerHr = \beta_0 + \beta_1\, Treatment + \beta_2 After + \beta_3 Treatment * After + \ \varepsilon$

Table 8 portrays the structure of the models and regression results. The baseline model (1) estimates the treatment effect (β_3)of the nudge for daily gas per hour respectively. Model (2) includes average outside temperature for the previous day. This is because the measurement was taken in the morning, with outside temperature of the previous day lining more closely to the measurement interval.

Table 8 portrays the regression models for gas/hour for the daily intervals. Model (1) is the baseline model showing a statistically significant treatment effect. This is interpretated as the treatment building having, after the intervention, 1.025CCF/hour higher natural gas usage than the control. The addition of the average outside temperature in model (2) has a negligible change in the treatment effect. Unsurprisingly, since as outside temperature increases the heating system does not have to produce as much heat to reach the Target Zone Temp, average outside temperature is negatively associated with gas usage: for every 1F° increase in average outside temperature there is a corresponding 0.048CCF/hour decrease.

Table 8: Natural Gas/Hour

Regression Model	(1)	(2)
Intercept	9.946***	11.717***
	(0.251)	(0.635)
Treatment	-4.612***	-4.584***
	(0.355)	(0.344)
After	-1.207***	-1.215***
	(0.332)	(0.318)
Treatment:After	1.025**	1.007**
	(0.470)	(0.453)
Lead Avg. Outside Temp.		-0.048***
		(0.016)
R-squared	0.769	0.780
F-statistics	104.329	86.069
N	98	98

Standard error reported in parentheses. *,**,*** denote conventional significance levels of 10%, 5%, and 1% respectively

8.0. Discussion

The most directly impacted outcome of the nudge is window behavior. The original hypothesis predicts window behavior would decrease (more windows are closed) after the nudge in the treatment building. Opposite to this original hypothesis, there is some evidence that window behavior increased slightly after the nudge in the treatment building (3.5 percentage points more likely to be open). The impact of the nudge, however, could be distorted by data that is not granular enough to capture the true effect of the nudge and bias enters the study because of when the measurements were taken (around 8am daily). For example, suppose, because of the nudge, students opened their windows more during the daytime in the treatment building but closed their windows at nighttime to pre-period window behavior rates; since the measurement was taken in the early morning, the nighttime window behavior might mask the daytime behavior.

If the nudge encouraged, at the very least, sensor-room students to keep their windows open, then there should be some measurable impact on room temperature. With more granular data, results for room temperature provide evidence the nudge made the snowballing problem worse. Over the various regression models, there's a treatment effect indicating rooms were 0.04-1.60F° hotter than they otherwise would be after the nudge. Although, with only a sample of 32 rooms, these results must be interpreted with some caution.

The most indirectly affected outcome, gas usage, saw a large statistically significant increase in the treatment building after the intervention as well (about 1CCF/hour more natural gas consumption). As Figure 4 indicates however, the effect was largely driven by a decrease in gas usage from the control. This is a surprising result since the nudge should not have affected the control building. One possible explanation is the trends of the treatment and control building were both decreasing after the pre-period and the nudge prevented gas/hour in the treatment from also decreasing (because of an increase in room temperature). Another possibility is the presence of some confounding factor only affecting the control building and not the treatment. For instance, natural gas is used for both water heating and room temperature in a building. Perhaps water heating in the control building reacts differently to changes in temperature than the treatment building.

In summary, the empirical results of the nudge indicate it either had no effect or made the snowballing problem worse. This is counter to the original hypothesis that the nudge would encourage students to keep their windows closed. In other words, the nudge potentially "backfired."

There are three logical explanations for a null or backfiring effect. First, as mentioned in section 4, this study implemented a descriptive social norm since the desired behavior (keeping windows closed) is already the majority behavior. Nevertheless, on certain days over the experimental period the number of windows with some level of openness neared 45%. An open or closed window is very observable to all students in the residence hall. If students interpreted the nudge's statement "the majority of students keep their window closed" as false, then the authority and trustworthiness of the message and messenger may be called into question, leading to a null effect. On a similar note, if, regardless of the contents of the nudge, it simply made window behavior more salient and since a large percentage of students did have their windows open, than perhaps it was inferred that having one's window open is the socially acceptable behavior, leading to a backfiring effect.

Secondly, previous literature suggests nudges using social norms can sometimes be interpreted as critical statements of behavior, particularly when one's perceived freedom of choice is being limited. This can lead to the frequency of a target behavior moving the opposite direction than desired. Although this is mainly observed when utilizing injunctive norms (judgements about a behavior) (Bicchieri and Dimant 2019). In the field of behavioral psychology, this effect is termed "reactance": a rebelling reaction against a perceived reduction in freedoms that lead an individual to continue behaviors or beliefs opposite to the original intent of the intervention (Brehm 1966). Moreover, reactance can be magnified in situations where the messenger is outside one's social group or seen as an overdemanding authoritative figure (Miller et al., 2006). This study used Residential Life as messenger because of the perceived authority and trustworthiness with students. However, an explanation for the backfiring of the nudge could be that Residential Life is perceived as an authoritative figure outside the student's social group, rather than inside. There is some evidence to support the nudge created reactance behavior;

Residential Life received some student feedback indicating they felt the nudge conveyed a negative injunctive judgment on their window behavior.

Lastly, in a situation that may be more unique to pro-social nudges, it is possible the nudge was simply unable to persuade a pro-social behavior in face of the large individual incentive to keep one's window open and experience more comfortable rooms.

9.0. Limitations

There are several limitations to the current study. Foremost, from an experimental design standpoint, there is the opportunity for spillover effects between control and treatment groups because of the proximity of the residence halls. That is, if students from the control building are exposed to the nudge (perhaps through talking with students in the treatment building), then the true treatment effect could be diminished. Secondly, while there are no statistically significant differences in student demographics between the treatment and control buildings, the lack of student level data limits the ability of the study to control for possible confounding factors more precisely.

From an academic perspective, as pointed out by Szaszi et al., (2018), a common complaint with nudges implemented in field settings is the focus on optimization of outcome while ignoring the opportunity to isolate individual effects to deepen the academic field (e.g., social norms vs. environmental sustainability messaging). While a valid observation, a limited timeframe and student population made the feasibility of, for example, comparing different types of cognitive heuristics or to contrast traditional incentive-based approaches improbable. On a similar note, the impact of the social norm from the more informational component to the nudge or a reminder could not be isolated. Additionally, since the majority of rooms do not have sensors installed, not all students have the ability to contribute to the snowballing problem (and by extension the negative externalities) by opening their windows. However, since students are not aware of this fact, and since the goal of the nudge is to create a social environment in which closing one's window is socially interdependent, the nudge being sent to all students is still valid.

Lastly, the results of this study are likely highly context dependent. While the snowballing problem does model as a negative environmental externality and similar situations may exist on other campuses (Jones 2022), the rather unique circumstances of the problem may make the external validity to a larger population dubious.

10.0. Conclusion

Literature on the use of behavioral interventions, such as pro-social nudges, is quickly developing. This study provides an opportunity to test a pro-social nudge's impact on two negative externalities at a small liberal arts college created by the interaction between the heating system and the tendency of students to open their windows—termed the snowballing problem. The aim of the nudge is to create a social norm to keep one's window closed, incentivizing the student to temporarily experiencing some disutility so all rooms can become more comfortable. The nudge utilizes a descriptive social norm as the underlying cognitive mechanism.

Using a natural field experiment design, the results of this study do not provide evidence the pro-social nudge impacted the target behavior to reduce the externalities of hotter rooms or excess greenhouse gas emissions. In fact, there is some evidence the nudge had a backfiring effect—making students more likely to keep their window open leading to increased temperatures and natural gas usage. A null result is not uncommon within the larger literature of pro-social nudges, with many nudges having heterogenous effects that differ depending on the characteristics of the population and situation (Bao and Ho 201). As Stibe and Cugelman (2016) identify, however, the phenomenon of backfiring within the field of behavioral economics is a growing area of interest and contributions to this field and explanations for backfiring effects are important for policy makers to avoid well-intentioned policy that ultimately backfires. This study highlights that nudges using descriptive social norms, particularly when the sociably desirable behavior does not overwhelmingly match the majority behavior, may be susceptible to unintentionally making the sociably undesirable behavior more socially acceptable. Furthermore, as Sunstein (2017) points out, the examination of behavioral reactance (rejecting an intervention because of the intervention itself), while often observed with mandates or bans, is an understudied area in relation to nudges. This study indicates that social norms in nudges, even when exclusively utilizing a descriptive norm, may produce reactance behavior—leading to a backfiring effect. Additionally, this study suggests reactance behavior may be intensified when the perception of the messenger is outside one's social group.

On a broader scale, altering actions with pro-social nudges may be particularly difficult since they nudge behavior beneficial to a larger group or society—not necessarily the individual. This tension between individual and societal utility is a critical area for policy makers and a ripe area for future study within the field of behavioral economics. This study is thus an important reminder to policy makers and researchers that nudges alone may not be sufficient to promote pro-social behavior.

Citations:

Allcott, Hunt. 2011. "Social Norms and Energy Conservation." *Journal of Public Economics* 95 (9–10): 1082–95.

Allcott, Hunt, and Sendhil Mullainathan. 2010. "Behavior and Energy Policy." *Science* 327 (5970): 1204–5. https://doi.org/10.1126/science.1180775.

Allcott, Hunt, and Todd Rogers. 2014. “The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation.” *American Economic Review* 104 (10): 3003–37.

Andor, Mark A., Andreas Gerster, Jörg Peters, and Christoph M. Schmidt. 2020. “Social Norms and Energy Conservation Beyond the US.” *Journal of Environmental Economics and Management* 103 (September): 102351. https://doi.org/10.1016/j.jeem.2020.102351.

Bamberg, Sebastian, Marcel Hunecke, and Anke Blöbaum. 2007. “Social Context, Personal Norms and the Use of Public Transportation: Two Field Studies.” *Journal of Environmental Psychology* 27 (3): 190–203. https://doi.org/10.1016/j.jenvp.2007.04.001.

Bao, Jiayi, and Benjamin Ho. 2015. “Heterogeneous Effects of Informational Nudges on Pro-Social Behavior.” *The BE Journal of Economic Analysis & Policy* 15 (4): 1619–55.

Bhanot, Syon P. 2021. “Isolating the Effect of Injunctive Norms on Conservation Behavior: New Evidence from a Field Experiment in California.” *Organizational Behavior and Human Decision Processes*, Nudges and Choice Architecture in Organizations, 163 (March): 30–42. https://doi.org/10.1016/j.obhdp.2018.11.002.

Bicchieri, Cristina, and Eugen Dimant. 2019. “Nudging with Care: The Risks and Benefits of Social Information.” *Public Choice*, 1–22.

Bicchieri, Cristina, Eugen Dimant, and Erte Xiao. 2021. “Deviant or Wrong? The Effects of Norm Information on the Efficacy of Punishment.” *Journal of Economic Behavior & Organization* 188: 209–35.

Bonan, Jacopo, Cristina Cattaneo, Giovanna d’Adda, and Massimo Tavoni. 2020. “The Interaction of Descriptive and Injunctive Social Norms in Promoting Energy Conservation.” *Nature Energy* 5 (11): 900–909.

Bruns, Hendrik, Elena Kantorowicz-Reznichenko, Katharina Klement, Marijane Luistro Jonsson, and Bilel Rahali. 2018. “Can Nudges Be Transparent and yet Effective?” *Journal of Economic Psychology* 65: 41–59.

Brehm, Jack W. 1966. “A Theory of Psychological Reactance.”

Caballero, Mara. 2004. “Academic Turns City into a Social Experiment.” *Harvard Gazette* (blog). March 11, 2004. https://news.harvard.edu/gazette/story/2004/03/academic-turns-city-into-a-social-experiment/.

Carlsson, Fredrik, Christina Annette Gravert, Verena Kurz, and Olof Johansson-Stenman. 2019. “Nudging as an Environmental Policy Instrument.”

Carlsson, Fredrik, Christina Gravert, Olof Johansson-Stenman, and Verena Kurz. 2021. “The Use of Green Nudges as an Environmental Policy Instrument.” *Review of Environmental Economics and Policy* 15 (2): 216–37.

Cialdini, Robert B. 2003. “Crafting Normative Messages to Protect the Environment.” *Current Directions in Psychological Science* 12 (4): 105–9.

Cialdini, Robert B., Linda J. Demaine, Brad J. Sagarin, Daniel W. Barrett, Kelton Rhoads, and Patricia L. Winter. 2006. “Managing Social Norms for Persuasive Impact.” *Social Influence* 1 (1): 3–15. https://doi.org/10.1080/15534510500181459.

Cialdini, Robert B., and Noah J. Goldstein. 2004. "Social Influence: Compliance and Conformity." *Annual Review of Psychology* 55 (1): 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015.

Cialdini, Robert B., Carl A. Kallgren, and Raymond R. Reno. 1991. "A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior." In *Advances in Experimental Social Psychology*, edited by Mark P. Zanna, 24:201–34. Academic Press. https://doi.org/10.1016/S0065-2601(08)60330-5.

Czajkowski, Mikołaj, Katarzyna Zagórska, and Nick Hanley. 2019. "Social Norm Nudging and Preferences for Household Recycling." *Resource and Energy Economics* 58 (November): 101110. https://doi.org/10.1016/j.reseneeco.2019.07.004.

Fong, Geoffrey T., David Hammond, and Sara C. Hitchman. 2009. "The Impact of Pictures on the Effectiveness of Tobacco Warnings." *Bulletin of the World Health Organization* 87: 640–43.

Goldstein, Noah J., Robert B. Cialdini, and Vladas Griskevicius. 2008. "A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels." *Journal of Consumer Research* 35 (3): 472–82.

Hagman, William, David Andersson, Daniel Västfjäll, and Gustav Tinghög. 2015. "Public Views on Policies Involving Nudges." *Review of Philosophy and Psychology* 6 (3): 439–53.

Haita-Falah, Corina. 2017. "Sunk-Cost Fallacy and Cognitive Ability in Individual Decision-Making." *Journal of Economic Psychology* 58: 44–59.

Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev. 2017. "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." *Journal of Public Economics* 148: 14–31.

Hallsworth, Michael, Tim Chadborn, Anna Sallis, Michael Sanders, Daniel Berry, Felix Greaves, Lara Clements, and Sally C. Davies. 2016. "Provision of Social Norm Feedback to High Prescribers of Antibiotics in General Practice: A Pragmatic National Randomised Controlled Trial." *The Lancet* 387 (10029): 1743–52.

Hanks, Andrew S., David R. Just, Laura E. Smith, and Brian Wansink. 2012. "Healthy Convenience: Nudging Students toward Healthier Choices in the Lunchroom." *Journal of Public Health* 34 (3): 370–76.

Hauser, Oliver P., Francesca Gino, and Michael I. Norton. 2018. "Budging Beliefs, Nudging Behaviour." *Mind & Society* 17 (1): 15–26.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Macmillan.

Kallgren, Carl A., Raymond R. Reno, and Robert B. Cialdini. 2000. "A Focus Theory of Normative Conduct: When Norms Do and Do Not Affect Behavior." *Personality and Social Psychology Bulletin* 26 (8): 1002–12.

Kantola, Steven J., Geoff J. Syme, and Norm A. Campbell. 1984. "Cognitive Dissonance and Energy Conservation." *Journal of Applied Psychology* 69 (3): 416.

Jones, Janae. Keep campus windows closed during winter months, January 7, 2022. https://www.k-state.edu/today/announcement/?id=79583.

Legros, Sophie, and Beniamino Cislaghi. 2020. "Mapping the Social-Norms Literature: An Overview of Reviews." *Perspectives on Psychological Science* 15 (1): 62–80.

"Wunderground Local Weather Forecast, News and Conditions." 2023. *Weather Underground*. Accessed April 8. https://www.wunderground.com/.

Lewis, David. 1969. *Convention: A Philosophical Study*. John Wiley & Sons.

Miller, Claude H., Michael Burgoon, Joseph R. Grandpre, and Eusebio M. Alvaro. 2006. "Identifying Principal Risk Factors for the Initiation of Adolescent Smoking Behaviors: The Significance of Psychological Reactance." *Health Communication* 19 (3): 241–52.f

Nayar, AMISHI, and S. Kanaka. 2017. "A Comparative Study on Water Conservation through Behavioral Economics Based Nudging: Evidence from Indian City 'a Nudge in Time Can Save Nine.'" *International Journal of Business and Social Science* 8 (11).

Noggle, Robert. 2018. "Manipulation, Salience, and Nudges." *Bioethics* 32 (3): 164–70.

Nyborg, Karine. 2018. "Social Norms and the Environment." *Annual Review of Resource Economics* 10 (1): 405–23. https://doi.org/10.1146/annurev-resource-100517-023232.

Schappert, Mikayla, and Andrew Shelly. 2018. "The Environmental Audit of Lycoming College ." *Lycoming College* ,

Schubert, Christian. 2017. "Green Nudges: Do They Work? Are They Ethical?" *Ecological Economics* 132 (February): 329–42. https://doi.org/10.1016/j.ecolecon.2016.11.009.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius. 2007. "The Constructive, Destructive, and Reconstructive Power of Social Norms." *Psychological Science* 18 (5): 429–34.

Stibe, Agnis, and Brian Cugelman. 2016. "Persuasive Backfiring: When Behavior Change Interventions Trigger Unintended Negative Outcomes." In *International Conference on Persuasive Technology*, 65–77. Springer.

Sunstein, Cass R. 2017. "Nudges That Fail." *Behavioural Public Policy* 1 (1): 4–25.

Szaszi, Barnabas, Anna Palinkas, Bence Palfi, Aba Szollosi, and Balazs Aczel. 2018. "A Systematic Scoping Review of the Choice Architecture Movement: Toward Understanding When and Why Nudges Work." *Journal of Behavioral Decision Making* 31 (3): 355–66.

Thaler, Richard H., and Shlomo Benartzi. 2004. "Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy* 112 (S1): S164–87.

Thaler, Richard H., and Cass R. Sunstein. 2009. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Penguin.

Tyers, Roger. 2018. "Nudging the Jetset to Offset: Voluntary Carbon Offsetting and the Limits to Nudging." *Journal of Sustainable Tourism* 26 (10): 1668–86.

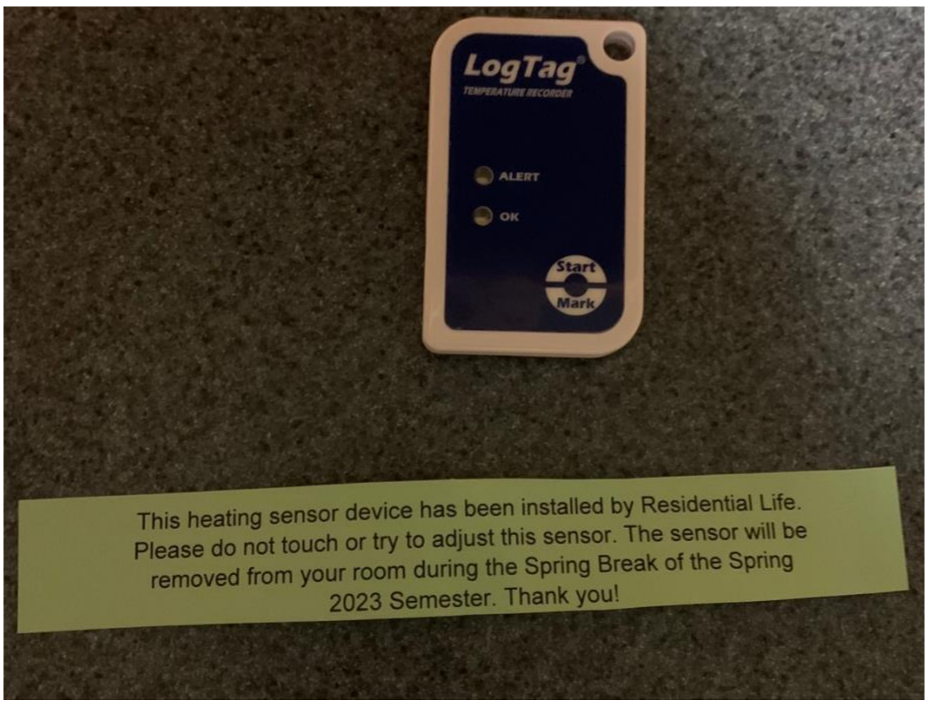
Velez, Maria Alejandra, and Lina Moros. 2021. "Have Behavioral Sciences Delivered on Their Promise to Influence Environmental Policy and Conservation Practice?" *Current Opinion in Behavioral Sciences* 42: 132–38.

Appendix A

Table 1: Summary of Variables

	Treatment				Control			
Variable	Mean	SD	Min	Max	Mean	SD	Min	Max
Avg. Outside Temp.	36.70	7.09	16.60	55.50	36.70	7.09	16.60	55.50
Room Temp.	71.48	4.57	50.10	97.00	69.68	2.91	56.80	79.80
Avg. Window Behavior	0.35	0.12	0.03	0.57	0.40	0.16	0.01	0.76
Gas/hour	5.23	1.02	2.45	7.50	9.26	1.39	5.77	11.99
StudSat	4.79	2.10	1.00	10.00	6.71	1.88	3.00	10.00
RoomHot	4.15	2.06	0.00	7.00	1.99	2.02	0.00	7.00
RoomCold	0.69	1.38	0.00	5.00	0.52	1.36	0.00	6.00

Image 1.



Appendix B

Survey

The survey of overall student satisfaction with the temperature in their rooms was digitally sent to all students in the treatment and control buildings through the same software that sent the nudges. Students were notified that, upon completion of the to the survey, they will have the opportunity to enter their email and be entered into a raffle for a $25 visa gift card, which will be given out approximately 1 week after the survey by Residential Life.

Starter Prompt:

"*Hello! This is a message from Residential Life at Lycoming! Please complete this survey on the temperature comfort in your room to be entered into a raffle for a $25 gift card! The survey takes 60 seconds to complete*"

Survey Questions:

Section

Dear Respondent,

1. What is the purpose of this survey?
This is a survey connected to a study trying to better understand the heating/cooling systems in the Lycoming dormitories.

2. What will I be asked to do as I complete this survey?
You will be asked to rate your level of comfort with the temperature in your room over the past week.

3. What are the benefits for completing the survey?
After completing the survey, you will have the option to enter your Lycoming email to be submitted into a raffle to win a $25 visa gift card to be handed out by Residential Life. Your email is not connected to your previous answers.

4. How will my information be kept private?
Your name and email address are not linked to your responses, allowing full confidentiality. Any further written material as a result of this survey cannot use names or any other identifying information that could identify you. Thus, your name is never connected with your survey information.

5. What are the risks of participating in this survey?
In any research study, there is always a small chance of loss of confidentiality. To protect against this risk, your anonymous information is coded with a number and then filed securely in a password protected file on a secure server.

6. Who can I contact about this survey?
Please feel free to contact Residential Life at reslife@lycoming.edu with any questions.

2. Were there any days in the past week where your dormitory was too cold? *

- Yes
- No

3. How many days in the past week did you think it was too cold in your room? *

1	2	3	4	5	6	7

4. Were there any days in the past week where your dormitory was too hot? *

- Yes
- No

5. How many days in the past week did you think it was too hot in your room? *

1	2	3	4	5	6	7

6. On a scale of 1 to 10, please rate your satisfaction with the temperature in your dormitory room over the past week. *

1	2	3	4	5	6	7	8	9	10

Very Unsatisfied | Very Satisfied

$25 Visa Gift Card Raffle

Winners will be contacted by Residential Life via email. Previous answers are not connected to your email.

1. *Optional*: Please enter your Lycoming College email for a chance to win a $25 Visa Gift Card from Residential Life! *

Enter your answer