ifh Working Paper No. 18/2019

Can APPealing and more informative bills “nudge” individuals into conserving electricity?

Lukas Meub[a], Petrik Runst[a], Kaja von der Leyen[b,*]

a Institute for Small Business Economics at the Georg-August-University Goettingen, Heinrich-Düker-Weg 6, 37073 Göttingen, Germany
b Economics Department, University of Goettingen, Platz der Göttinger Sieben 3, 37073 Goettingen

Abstract:

We use a field experiment on energy billing in a German region to evaluate the effect of two behavioral nudges (consumption feedback and social comparison) on electricity consumption. Similar experiments have revealed significant treatment effects, yet the individual variance has proven to be substantial. On grounds of these heterogeneous treatment effects and the possibility of cross-country behavioral differences, additional experiments are warranted. For our German participants with low pre-treatment consumption compared to many other countries, we find no treatment effects. From this, we deduce that the effect of consumption feedback and social comparison is highly context dependent.

JEL: K32, P18, Q58

Keywords: Energy consumtion, Electricity, Consumtion feedback

* Corresponding author. kaja.vonderleyen@wiwi.uni-goettingen.de

1. Introduction

Household energy bills communicate information that consumers can use to adjust their energy consumption, whether through short- (e.g. switching on/off lights) or longer-term behavior (e.g. investing in more energy-efficient appliances). A large body of literature suggests that various information mechanisms can induce energy conservation (see e.g. Schultz, 2007; Allcott, 2011; Asensio and Delmas, 2015; Allcott and Kessler, 2019). These so-called behavioral nudges are carefully-crafted psychological cues that can affect behavior without changing prices or restricting choice sets and in theory they are a good companion to or replacement of traditional climate policy instruments. They avoid important shortcomings, namely the difficulty of reaching international implementation agreements, competitiveness penalties associated with unilateral introduction, low public acceptability due to higher energy prices and regressive distribution effects.

Current German electricity billing practices clearly do not exploit their potential as carriers of information. Utilities only provide their residential customers with a bill for actual consumption once a year. While many German households pay for electricity monthly, the sum is estimated based on last year's consumption. As Wilhite and Ling (1995) highlight, in such systems there is no relationship between the cost paid in a given month and actual consumption.

Upon first glance, it therefore appears as if optimizing billing practices by taking into account insights from behavioral economics is an opportunity to spark energy efficiency in Germany. Given the German government's ambition to reduce electricity consumption by 10% by 2020 compared with 2008 levels (BMWi and BMU, 2010), one might be tempted to speculate about the potential benefits of a mandatory monthly billing cycle. However, the existing literature shows large inter- and intra-study variation in the treatment effects of interventions in the field of home energy use, which casts doubt on the universality of the findings. In the face of heterogeneous effects, it seems vital to test the effectiveness of nudges in multiple contexts. Given that there may be a wide array of cultural and socio-economic differences in preferences and behaviors, one should be careful in adopting a "one-size-fits-all" approach.

Very few studies examine the effects of behavioral nudges in household electricity bills on electricity consumption in Germany (Dünnhoff and Duscha, 2007; Mack and Hallmann, 2004). These studies are small scale (e.g. one residential block) and rely on low-level interventions. They also find small effect sizes in comparison with results found in other countries (Fischer, 2008).

In order to fill this research gap, we set up a field experiment in collaboration with six local energy utilities in the "Sued-Niedersachsen" region. We introduce monthly billing cycles based on a smart phone app that was designed and programmed for this experiment. There are two graduated treatments, the first one relying on monthly (as opposed to annual) feedback and the second one adding a consumption comparison feature with similar households. In contrast to many existing studies, we do not find treatment effects. We conclude that the effect of consumption feedback and social comparison is highly context dependent. In particular, we suspect that behavioral nudges require comparatively high pre-treatment consumption levels to be effective.

The paper is structured as follows: Part 2 provides the reader with the necessary background and details relevant existing studies. Part 3 outlines our experimental design. In part 4, we present and discuss our results, before part 5 concludes our undertaking.

2. Background

Psychological research in the 1970s and early-1980s contributed to understanding the determinants of energy use and energy conservation particularly for households (see Stern (1992)). These ideas have been pursued by economists interested in energy policy and picked up by utility companies and policy-makers, who have implemented policies based on behavioral insights.

We focus on two specific interventions, namely consumption feedback and consumption comparison among individuals. Households' feedback about their energy savings may encourage them to reduce energy use because their level of self-efficacy (i.e. their perceived capacity to conserve energy) has increased (see Abrahamse et al, 2005). In terms of social comparisons of energy consumption levels, according to Andor and Fels (2018), the potential effect on consumption can be triggered by three phenomena. First, because many people exhibit reference-dependent preferences (see Kahneman 2003), and one such reference point is social norms, deviating from a provided norm (in this case, the electricity consumption of other people) creates disutility and can therefore lead to behavioral change. Second, in situations of uncertainty, individuals may use other people's behavior as focal points by implicitly assuming that others have more information about the socially-desired behavior (see e.g. Allcott and Mullainathan (2010); and Delmas et al. (2013)). Third, social comparisons can evoke competition-like behavior (see Abrahamse et al. 2005), which is especially important when a household's consumption level lies above that of others.

The existing empirical literature on the effects of similar interventions is substantial, covers multiple disciplines and comprises various methodological approaches. In the following, we summarize the findings of a selection of such studies, categorized as meta-studies, small-scale intervention studies and large-scale intervention studies. Overall, the literature finds that feedback and social comparison can spark energy conservation behavior, although the effect sizes obtained substantially vary.

Starting with a selection of literature reviews and meta-Studies, Darby (2006) reviews 38 studies and concludes that feedback mechanisms can induce residential energy savings of 0% to 15%, with a possible upper limit of 21%. In line with this, Fischer (2008) reviews existing studies on the consumption effects of feedback and notes that estimates range from 1% to 20%, with some studies showing no effects at all. The author does not find an effect related to social comparison in her review.

EPRI (2009) reviews 31 studies and find average treatment effects on electricity consumption for different categories of feedback mechanisms ranging from 4% to 12%, with the strongest treatment effects for studies with more direct feedback mechanisms. Erhardt-Martinez et al. (2010) review 57 feedback studies from developed countries undertaken between 1974 and 2009. They also categorize them according to the directness of the feedback mechanism and find average effects ranging from 5% to 14%.

Delmas (2013) conduct a meta-analysis of 156 published field trials from 1975 to 2012 and find that consumption feedback yields an average effect on energy consumption of -7%. Robust studies designed to contain a control and treatment group only find savings of less than 2% on average. Finally, Karlin and Zinger (2015) examine 42 feedback studies published between 1976 and 2010. Again, on average feedback is found to reduce energy consumption by 7%, albeit again with a wide range of effect sizes. When controlling for publication bias, the average effect size falls to about 5%. The authors state that social comparison does not reduce energy consumption across their sample of studies.

Turning to selected small-scale studies, Schultz et al. (2007) investigate the effect of consumption feedback and social comparison on energy conservation in 287 Californian households. The authors find that their feedback induced energy conservation among participants who had above-average consumption. By contrast, for participants with below-average energy consumption, the nudge actually led to an increase in energy consumption. In a second treatment, the authors used happy and sad smiley faces to signify above- or below-average consumption. In this instance, the high consumption group reduced their consumption, whereas consumption did not change for the low consumption group.

For Germany, Dünnhoff and Duscha (2007) investigate the effects of supplementing the usual electricity bill with a normative comparison as well as written advice in partial combination with personal consultancy. The sample comprises 4,500 individuals, distributed over three treatment groups and one control group in the city of Heidelberg. While many of the participants reported finding the additional information interesting and motivating, it did not spark significant electricity conservation. Furthermore, Mack and Hallmann (2004) examine the effects of weekly written feedback on electricity consumption among 30 German households (of which 19 received the treatment). They found that their treatment reduced consumption by 2.9% during the time of the intervention. Among 28 international studies on the effect of feedback mechanisms on electricity consumption, both studies are categorized as lying in the bottom third in terms of effect sizes (Fischer, 2008).

Finally, a number of large-scale interventions are also available. In a research report on nudging commissioned by the Swedish Environmental Protection Agency, Mont et al. (2014) note a large discrepancy between the results in small- and large-scale field trials, with the latter displaying much lower effect sizes. Similarly, the meta-analysis by Karlin and Zinger (2015) finds a strong relationship between effect size and sample size (p.1215) and the authors suspect that publication bias is at work (i.e. a bias against smaller-scale studies with lower effect sizes).

Feedback on electricity bills providing social comparison information was tested in Helsinki, Finland from 1989 to 1992, whereby it was found to reduce consumption by 1-1.5% (Mont et al., 2014). Simultaneously, in coordination with the Finnish trial, a similar experiment was conducted in Oslo, Norway. Interestingly, the treatment had much stronger energy conservation effects in this case, as the sample of 1,450 household saved 10% of their energy consumption on average, primarily due to increased billing frequency (Wilhite and Ling, 1995).

Tedenvall and Mundaca (2016) use an experimental setup to test the impact of real-time consumption feedback. The authors find a decline in consumption by about 1.4-1.9%. However, they interpret the changes to small and conclude that "the implementation of real-time feedback per se is likely to be insufficient to foster increased energy efficiency" (Tedenvall and Mundaca, 2016). Gleerup et al. (2010) analyze the effects of providing consumption feedback via text messages and email on the use of electricity. With a sample of 1,452, households they find a significant reduction in annual electricity use of 3%. Schleich et al. (2013) find that providing feedback to 1,500 Austrian households produces electricity savings of 4.5%.

Perhaps the large-scale intervention study to have gained the most traction is Allcott's (2011) study on nudging and energy billing, exploiting a program run by the utility company Opower. He looks specifically at the effectiveness of giving feedback in a contest setting on energy use using energy reports mailed to as many as 600,000 treatment and control households in the US. Specifically, the treatment comprised a comparison of the household's energy use with that of similar neighbors as well as specific energy conservation tips. These reports are sent to the treatment group monthly, bimonthly, or quarterly, depending on the utility company. Allcott (2011) finds a negative effect on energy consumption of 2% and concludes that the cost-effectiveness compares favorably to that of traditional energy conservation programs.

Ayres et al. (2013) also use the Opower home energy reports to examine their effects on the customers of two utilities (170,000 American households). Their studies confirm the findings in other studies using this data, namely that peer comparison feedback reduces residential energy usage.

From this extensive literature on consumption effects of feedback and social comparison, it is not only clear that treatment effects vary across studies, but also many studies find heterogeneous treatments within their samples. In particular, pre-treatment consumption levels appear to be important. Allcott (2011) finds that treated households in the highest pre-treatment consumption decile achieved an average energy conservation resulting in the considered nudge of 6.3%, whereas those in the lowest pre-treatment consumption decile conserved only 0.3%, with the baseline consumption ranging from 19 to 60 kWh per day. The average electricity savings found in Schleich et al. (2013) were only significant for the 30th-70th consumption percentile. Allcott and Kessler (2019) find that while most participants appreciated receiving the home energy reports, 34% still have a weakly negative willingness to pay for receiving the report. More generally, in their encompassing literature review of studies on nudges and energy conservation, Ayres et al. (2013) note that the effect of social comparison feedback was strongest in households with the highest energy consumption.

Hence, while the findings from the existing literature suggest that consumption feedback and social comparison can indeed serve to reduce household energy consumption, the strong variance of effect sizes points to the context-dependent nature of the hypothesized mechanisms. In particular, previous results offer reasons to suspect that consumption feedback is particularly effective in high consumption contexts as there are many remaining opportunities for saving energy. Once the low-hanging fruit has been picked, the consumption-lowering potential within a household is likely to taper off. Consequently, when conducting a field experiment in Germany – where, according to World Bank data[1], electricity consumption per person is about half of the US value – we expect to find a lower treatment effect than those based on North-American samples.

3. Quantitative Analysis

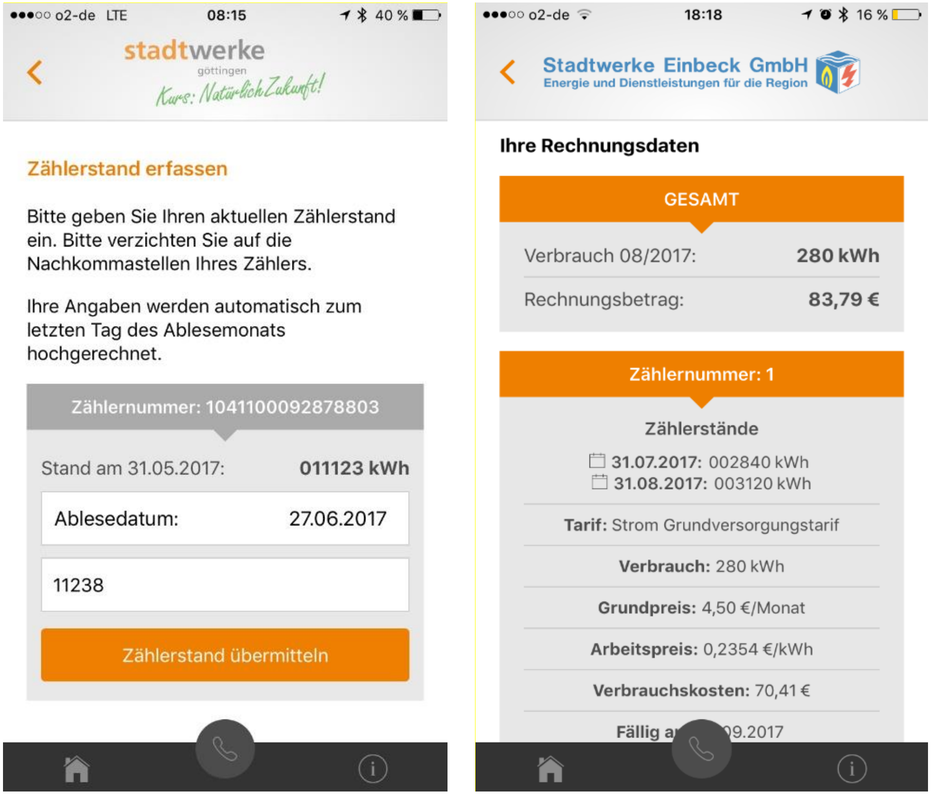
a. General properties of the experiment

Our app-based field experiment investigates the effects of non-monetary incentives (“nudges”) on energy conservation, where we recruited participants from the customer pool of local utility companies. The usual electricity billing procedure requires customers to pay a lump sum monthly rate calculated based on their prior consumption, in addition to a premium for potentially increasing consumption. These payments are balanced with the actual consumption at the end of the year, after which the rate might be adjusted for the next billing cycle. This is by far the most common billing procedure in Germany.

Our experiment manipulates this procedure by having subjects self-report their monthly energy consumption (meter reading) and billing them accordingly. A smartphone app – programmed for the purpose of the experiment only – represents a suitable instrument to implement this alternative billing design. Thus, subjects become aware of their actual consumption and receive monthly feedback on their energy usage over time (see figure 1).

[1] https://data.worldbank.org

Figure 1. Billing app design



Note: On the left, the screen for entering monthly consumption is shown. Here, customers entered the date and their meter reading. They are informed that meter readings will be projected onto the last day of the month. On the right, the billing screen is shown. Here, customers are informed about their consumption in kWh and the resulting payment that is due. Furthermore, they received information about all of their meter readings, their monthly base payment and the cost per kWh. The banner on the top of both screens depicts the respective corporate design of the respective customer's supplier.

One of the treatment groups received additional information on how their consumption fared compared with the average consumption of participating households of equal size. In order to strengthen this feature and distinguish the two treatment groups, this group also received explicit feedback (thumbs up or thumps down) depending on whether their consumption was above or below that of comparable households.

We introduce a control group of participants registering for the study who received no alternative invoice, but rather continued to pay the monthly lump sum charge based on last year's electricity consumption. Their interface only displayed a countdown stating when all app features would be available to them.

The experiment was carried out in two phases. During phase 1, the participants were randomly assigned to one of our treatment conditions and the control group. During phase 2, all participants were able to use all features of the app. The latter allows us to analyze treatment effects based on a differences-in-differences regression and investigate longer-term effects. Furthermore, this design choice reflects an attempt to avoid dissatisfying the customers in the control group by excluding them from the benefits and features of the app. However, our primary analysis refers to the between-subject comparison of phase 1 of the experiment. Table 1 summarizes the treatment conditions, phases and number of participants.

Table 1. Treatment overview

	Phase 1					Phase 2
	App	Feedback*	Social**	App	Feedback	Social
T_0	yes	no	no	yes	yes	Yes
T_1	yes	yes	no	yes	yes	Yes
T_2	yes	yes	yes	yes	yes	Yes

Note: T_0, T_1 and T_2 denote the control group, treatment group 1 and treatment group 2.
* 'Feedback' refers to monthly feedback on consumption.
** 'Social' refers to the comparison of participants' own consumption with that of similar households (same number if household members).

b. Procedure

The field experiment was developed in cooperation with six local energy suppliers, two of which actively participated in the experiment itself and enabled their customers to participate. In order to set up the experiment, a number of technical and organizational issues had to be resolved. In particular, it was necessary to harmonize the app-based billing procedure of our experiment with the suppliers' accounting system and customer support was established for the participants of the experiment. The overall billing procedure was designed in a way that conforms with German legal electricity billing requirements.

In May 2017, we invited 10,000 customers from energy supplier #1 and 2,000 from #2 to participate via a letter by mail.[2] The letter provided customers with a download link and log-in information (see invitation letter in appendix A). The log-in information referred to a unique identifier for each customer that is used by the supplier. Thus, by logging in, each registered customer could be identified and moved to the app-based payment scheme. The invitation letter contained minimal information regarding the experiment to avoid influencing participant behavior, e.g. through a demand effect by hinting at energy conservation. It succinctly stated that the study planned to investigate electricity consumption in private households. In order to encourage participation, the invitation letter emphasized the novel aspect of the study, namely the use of the smartphone application to transmit electricity meter readings. As mentioned above, using apps for this purpose is – apart from a few exceptions – not practiced in Germany.

In order to register for the study, the participants provided the following information in a short questionnaire: age, household size (including number of children), education, dwelling size (number of rooms and square meters), whether they are owners or tenants, whether they have electrical heating, how they think they compare with others in terms of electricity consumption, and whether they would like to consume less electricity. Finally, the participants were asked to respond to statements regarding their interest in technological developments.

Moreover, during the registration, the participants also entered an initial meter reading at the end of May 2017 and received an invoice prior to the start of the experiment that balanced their account from January to May 2017. Accordingly, all participants started with a clean sheet and the treatment groups received the first alternative monthly billing via the app at the end of June 2017. From this point onwards, invoices were available through the app and the subjects were informed by a push notification that they could access their invoice information. All monthly invoices were stored in an archive and available to participants.

Please note that for the last month in phase 1 – i.e. December – we use the customary end-year meter readings for all participants. This is fortunate, as we are able to reconstruct the consumption over the treatment period (beginning of June to end of December) for all initial participants and hence circumvent problems that would have otherwise arisen due to sample attrition. Furthermore, participant readings or entries might be incorrect, which we could thus identify by comparing their own readings with the official reading. Finally. this procedure also allows us to exclude lengthy technical information from the monthly bills via the app that are required by law, as this information is given by the final invoice at the end of the year.

At the end of phase 1, we conducted a survey via the app, collecting information on the number of days that they were absent from their dwelling during the treatment period, whether their electricity consumption changed during the treatment period and – if so – whether their quality of life had been affected, a number of questions to assess their price sensitivity (e.g. income level, changes of energy offers/suppliers over time) and preferences for the environment (e.g. whether they have donated money to environmental organizations and whether they are concerned with different negative events including climate change).

From January 2018 onwards, all participants were able to use all features of the app, i.e. they were billed monthly according to their meter readings and were shown the social comparison. The experiment ended in May 2018 and the app was deactivated. To inform them of this, participants were sent a letter with a thank-you note in advance.

c. Design specifics

In order to guarantee a high quality of our data, participants not entering their monthly meter reading twice were excluded from the study and reassigned to the traditional non-app-based billing procedure with monthly lump sum payments. As a result, we excluded 35 and 25 participants of T_1 and T_2, respectively. If participants failed to report their reading once, their consumption was assumed to be equal to the previous month and they were billed accordingly. In order to prevent participants from dropping out due to temporary absence from their home, they were able to enter their meter readings within a certain time period, namely from three days prior to until three days after the last day of each month. They were informed by a push notification when their meter reading was due. Given that not all meter readings were conducted on the last day of the month, we corrected the reported

[2] Commercial customers, customers with a history of unpaid bills and other unusual cases – like customers only relying on electricity to heat their houses, who are charged at a special rate – were not invited.

consumption, which may – for example – be based on a 27-day consumption period to a full month. In order to calculate the averages for the social comparison treatment, we relied upon averages conditional on household size, i.e. the number of persons in each dwelling.

Furthermore, in order to encourage participants to remain in the study, a monthly lottery was established, in which five individuals would win 100 Euros. Again, a push notification informed subjects about the lottery and whether they had won or not, after which winning participants received the money by bank transfer. The subjects were (truthfully) informed that the probability of winning the lottery was independent of their electricity consumption.

The app was programmed and supervised by a third-party service provider, who collaborated closely with the energy suppliers to guarantee an efficient data transfer. The design of the app was developed to adhere to experimental economics standards, whereby it was kept plain and simple. We integrated the corporate design of the suppliers to induce trust regarding the handling of personal information and payments through the app. In order to further enhance trust, the research nature of the enterprise was emphasized. Hence, the app – presented in the Apple and Google Play Store – was named "University of Göttingen energy app". The programming of the app and the rewards from the lottery where financed by contributions from all six energy suppliers.

d. Sample size and sample attrition

As can be seen in table 2, not all individuals who signed up for the experiment in May 2017 entered their electricity meter information at the end of the first month of the experiment. Of course, none of the participants in the control groups were permitted to do so. After that, the number of participants who entered their meter readings remained relatively stable in both treatment groups over the six-month phase 1 treatment period. In fact, the number of meter readings slightly increases in the fall as individuals are less likely to be absent due to vacations. There is a higher number of meter readings in December as the utility company provided us with meter readings based on the end-of-year invoice for all remaining participants. These end-of-year meter readings were particularly important in the case of control group participants, most of whom had deleted the app at this point because they were not permitted to use any of its interesting features.

Participation declined at the beginning of phase 2, after participants had received a letter thanking them for their participation and informing control group individuals that they were now able to use all app features. As expected, the majority of control group participants were no longer interested in the experiment, as they had not been permitted to use the full functionality of the app during phase 1. Nevertheless, as previously stated, we obtained the meter data for most control group individuals for December 2016 and 2017 from the energy suppliers, thus allowing us to compute their electricity consumption during phase 0 and phase 1.

In May 2017, prior to the start of the experiment, there were 138, 150 and 146 participants in T_0, T_1 and T_2, respectively. Of those individuals, we have meter readings for 90, 85 and 100 participants, respectively for December 2016 to compare pre-experiment consumption levels across groups for the period from January to May 2016 (phase 0). At the end of phase 1, in December 2017, there were 138, 115 and 121 participants in each group. While the sample size is small compared with the large intervention studies referred to in part 2 of this paper, it is not insignificant seen in light of the entire literature on consumption feedback, given that many field experiments contain no more than 10-50 households, which are again often divided into sub-groups (Fischer, 2008).

Table 2. Number of participants with meter readings (by treatment group over time)

	2016	2017	2017							2018				
	(phase 0, pre-experiment)		(phase 1)							(phase 2)				
	Dec	May	June						Dec					May
	0	1	2	3	4	5	6	7	8	9	10	11	12	13
T_0	90	138							138	31	31	31	31	31
T_1	85	150	109	111	114	114	114	118	115	91	91	91	91	91
T_2	100	146	108	112	110	112	116	113	121	90	90	91	91	91

Note: The number of participants in December 2017 includes individuals who failed to enter their meter readings via the app for that month but who remain in the sample. The utilities provided meter information for these participants based on the end-of-year invoice 2017. There are 35 and 25 individuals in T_1 and T_2, respectively who failed to enter their monthly meter readings more than once in phase 1. These individuals have been excluded from the sample. We obtained a number of meter readings for December 2016 from the end-of-year invoice of that year. However, for various reasons we could not obtain this information for all participants (e.g. some of them may have had a different meter number or they may not yet have been customers of the utility company).

[*]The May columns contain all individuals who initially signed up for the experiment.

e. Empirical methods

First, we compare mean consumption values between T_0 and T_1, in addition to T_0 and T2 by applying simple t-tests. Second, we perform an ordinary least squares regression (OLS) analysis, which allows us to include a number of control variables. In the OLS specification, we use cross-section data covering phase 1 only. Finally, we run difference-in-differences (DiD) specifications. Since the utility companies also supplied us with the official meter readings of the participants in December 2016, our data contains meter readings (and therefore consumption values) from the pre- and post-treatment period and we can use DiD to allow for time-invariant unobserved heterogeneity across groups. This ensures that our results are not biased by less-than-perfect randomization.

The first DiD specification covers two time periods for each participant – phase 0 and phase 1 – and has the following functional form:

$$\mathrm{Consumption}_{it} = \alpha + \beta_1 \mathrm{T\_1}_{it} + \beta_2 \mathrm{Post}_{it} + \beta_3(\mathrm{T\_1}_{it} \times \mathrm{Post}_{it})$$

$$+\gamma_1 \mathrm{T\_2}_{it} + \gamma_2 \mathrm{Post}_{it} + \gamma_3(T_{2it} \times \mathrm{Post}_{it}) + \pi \mathrm{X}_{it} + \varepsilon_{it}$$

where the dummy variables 'T_1' and 'T_2' are equal to one if the participant is assigned to the first or second treatment group, respectively, and zero otherwise. The dummy variable 'Post' is equal to one if the observation belongs to the treatment period in phase 1 and zero otherwise (phase 0). All data from phase 2 has been excluded for the purpose of this first DiD analysis. The coefficient for the interaction term 'T_1' and 'Post' captures the treatment effect (also T_2 and Post). The vector X contains a number of control variables, such as household size, number of rooms, and the level of education. Education is measured as a set of dummy variables, where having obtained a vocational training degree serves as the reference category.

In the same manner, we can perform a second DiD specification by comparing phase 1 vis-à-vis phase 2. All data from phase 0 has been eliminated for this second DiD analysis. In phase 1, T_1 and T_2 received treatment and T_0 did not. However, in phase 2 T_0 also receives treatment. Thus, in this second DiD specification T_1/T_2 serve as the control group and T_0 serves as the treatment group. As there are only 32 individuals in T_0 who actively participated until the end of phase 2, the treatment group is somewhat smaller in this specification.

4. Results

a. Estimates obtained

Table 3 displays various descriptive statistics by treatment group to evaluate the existence of systematic differences between the treatment and control groups. Generally, the presented statistics are similar across the groups, although apartment sizes are somewhat larger in T_0 than in T_1 and T_2. In addition, the share of individuals strongly concerned about the environment and those interested in technological developments is larger in T_1 than in both T_0 and T_2. Consequently, we control for these observable characteristics when we run regressions.

The lower panel of table 3 displays total electricity consumption by group and experimental phase. During phase 1, control group consumption (1085.5 kWh) lies between consumption in T_1 (950.8 kWh) and T_2 (1128.1 kWh). In phase 1, control group consumption is actually lower (1408 kWh) than consumption in T_1 (1440 kWh) and T_2 (1458 kWh). Finally, in phase 2, the consumption in T_0 is equal to 1112 kWh, whereas consumption in T_1 equals 1100.7 kWh and T_2 consumptions equals 1171.4 kWh.

The t-tests do not detect statistically significant differences in mean consumption across groups in phase 1 (nor phase 0/ phase 2). Based on these descriptive results, it is unlikely that the treatment affects consumption. Note that phase 1 last seven months, while the pre-treatment period and phase 2 last two to five months, which explains the level differences between periods.

Table 3. Statistics by treatment group

	T_0	T_1	T_2
Age	52.04	51.22	50.23
Standard deviation	*32.11*	*15.45*	*15.10*
#Children	0.41	0.35	0.45
Standard deviation	*0.78*	*0.76*	*0.76*
Apartment (m^2)	114.93	105.37	107.81
Standard deviation	*69.61*	*47.78*	*44.36*
University degree	46.27%	44.56%	45.26%
Standard deviation	*50*	*50*	*50*
Very concerned about environment	18.01%	25.82%	18.83%
Standard deviation	*38*	*44*	*39*
Owner	52.85	51.75	52.53
Standard deviation	*49.9*	*49.98*	*49.95*
Interested in techn. developments	89.45	94.33	89.48
Standard deviation	*30.73*	*23.13*	*30.69*
N	137	115	121
Consumption (kWh)			
Phase 0 (January - May, 2017)	1085.5	950.8	1128.1
Standard deviation	651.0	680.5	540.5
t-test for mean differences		Not significant	Not significant
p-value		*0.17*	*0.62*
N	100	85	90
Phase 1 (June - December , 2017)	1408.3	1440.2	1458.1
Standard deviation	*792.1*	*717.3*	*703.5*
t-test for mean difference vis-à-vis T_0		Not significant	Not significant
p-value		*0.73*	*0.59*
N	137	115	121
Phase 2 (January - May, 2018)	1112.1	1100.7	1171.4
Standard deviation	*589.9*	*560.9*	*580.0*
t-test for mean difference vis-à-vis T_0		Not significant	Not significant
p-value		0.92	0.62
N	32	91	91

Table 4 displays the OLS results. There is no evidence of a treatment effect in T_1 or T_2 as the coefficients are positive but not significantly different from zero. Estimated coefficients for all the control variables coefficients have the expected sign. The number of rooms as well as the number of people living in the household are positively associated with electricity consumption, increasing consumption by 109.9 and 304 kWh per increment, respectively. Similarly, university graduates consume 208.5 kWh less than people in the vocational training category, whereas all other educational categories are not significantly different from the comparison group.

There is evidence for the existence of the owner/tenant dilemma as the coefficient for the owner dummy variable is negative and statistically significant. This should nevertheless be interpreted with caution, as tenants pay (either directly or indirectly over additional costs related to heating, lighting and services) for their electricity consumption. However, due to the shorter dwelling time, tenants' incentives to invest in more energy-efficient appliances are lower.

Table 4. OLS regression (cross section, phase 1)

Dep. Var.	Consumption
T_1	63.33
	(0.348)
T_2	9.261
	(0.903)
University graduation	-208.5***
	(0.005)
University of applied sciences	22.38
	(0.786)
Advanced voc. training / Technician training	-0.309
	(0.997)
No occupational training	-76.18
	(0.698)
Other	71.51
	(0.663)
#Rooms	109.9***
	(0.001)
Household size	304.0***
	(0.000)
Owner	-267.0***
	(0.000)
Interested in technological developments	-31.48
	(0.721)
Constant	765.9***
	(0.000)
Observations	372
r2	0.443

Note: P-values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Robust standard errors have been specified.

Table 5 displays the results of the first DiD specification, where we focus on phase 0 and phase 1. Again, there is no evidence in favor of a negative relationship between treatment and electricity consumption. The interaction term between T_1 and the 'Post' dummy is equal to 17.42 but not statistically significant. The interaction term between T_2 and the 'Post' dummy is equal to 30.72, and again not statistically significant. Control variable coefficients and levels of significance are of similar magnitude as in the OLS regression. Table 5 also displays the results of the second DiD specification, where we focus on phase 1 and phase 2, estimating the treatment effect after T_0 switches from receiving no treatment to receiving treatment. Again, there is no evidence of an effect. The interaction term coefficient (T_0 and 'Post') is equal to -16.73 but statistically insignificant.

Table 5. Difference-in-differences regression (panel data)

	(1) **Phase 0 & 1**	**(2)** **Phase 1 & 2**
Treatment (a)	47.47	-44.49
	(0.370)	(0.473)
Post	287.1***	-347.6***
	(0.000)	(0.000)
Treatment (a) x Post	17.42	-16.73
	(0.641)	(0.711)
Treatment (b)	-7.636	
	(0.895)	
Treatment (b) x Post	30.72	
	(0.473)	
University graduation	-189.0***	-203.5***
	(0.004)	(0.002)
University of applied sciences	3.232	10.37
	(0.966)	(0.892)
Advanced voc. training / Technician training	-7.930	-17.02
	(0.920)	(0.834)
No occupational training	-98.40	-99.18
	(0.570)	(0.566)
Other	76.67	60.78
	(0.583)	(0.693)
# Rooms	106.3***	92.09***
	(0.000)	(0.004)
Household size	281.8***	284.3***
	(0.000)	(0.000)
Owner	-259.9***	-252.0***
	(0.000)	(0.000)
Interested in technological developments	-29.37	-49.14
	(0.711)	(0.549)
Constant	519.9***	914.6***
	(0.008)	(0.000)
Observations	619	586

Note: In specification (1), treatment (a) refers to T_1and treatment (b) refers to T_1. In specification (2), treatment (a) refers to T_0.
P-values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Robust standard errors have been specified.

b. Discussion

Our study thus echoes those existing studies that cast doubt on the efficiency of behavioral interventions in the field of home energy use. In order to gain a better understanding of the possible reasons for why we find no significant treatment effect, it is useful to look at how the literature explains the large variance in treatment effects found in this field of study.

According to Fischer (2008), the determining factor is the design of the nudge itself. Consumption feedback is most efficient when it is frequent, involves interaction and choice for households, includes a breakdown of consumption by appliances, is given over a long period and is presented in an understanding and appealing way. Regarding the latter, Herrmann et al. (2017) find that people experience home energy data as being difficult to understand and link to everyday behaviors. Table 6 assesses our nudge according to Fischer's (2008) criteria for nudge effectiveness. Overall, we find that our nudge sufficiently fulfills the criteria and we reject the notion that our nudge was simply not sufficiently strong to produce any behavioral changes.

Regarding frequency, our nudge took place once a month, which is a large increase compared with the pre-treatment situation in which the participants only received information about their electricity consumption once a year. Regarding consumer-interaction, the design of the nudge meant that consumers not only passively received a bill, but they also had to go to their meters once a month and type the new reading into the application. However, consumer choice was more limited: although receiving the nudge was voluntary, once consumers opted in they were unable to influence the design of the nudge.

Our nudge scores low on the criterion of information specificity, as the households only received information about their total consumption and were not informed regarding which appliances/behaviors generated the highest costs. By contrast, the nudge was put in place for a year, which is longer than most other studies in the literature[3]. Finally, the nudge was presented in an innovative way to consumers through the use of an electricity billing application, which is a novelty in the German context. Fischer (2008) highlighted that computerized nudges appear particularly effective.

Table 6. Assessment of the nudge according to Fischer's criteria

Criteria	**Frequent feedback**	**Interaction**	**Choice**	**Consumption by appliance**	**Time period**	**Design**
Study design	Once per month	Consumers type in meter reading	Voluntary participation, predetermined nudge	Total electricity consumption	12 months	Smartphone application
Fulfillment	+/-	+	+/-	-	+	+

Delmas et al. (2013) state that the variation in effects seen across field experiments on home energy use is due to differences in methodological quality. More specifically, they find that studies of the highest methodological quality (studies with a control group as well as weather and demographic controls) obtain the smallest effect sizes. Our study can be said to belong to this group, given that we randomly assigned participants.

As highlighted in the introduction, several studies find heterogeneous treatment effects, with the effects being stronger among households with higher pre-treatment consumption levels. In the background chapter, we noted the interesting fact that in a similar field experiment consumption feedback reduced consumption among Finnish households by 1-1.5% whereas among Norwegians it was reduced by 10%. In their evaluation of the Norwegian case, Wilhite and Ling (1995, p. 145) highlight that "there is ample evidence of excessive energy use and wasteful energy-use behaviour in Norwegian households".

To explain our results, it is therefore interesting to know the consumption level of our treatment group compared with that of other studies. In our sample, the average daily household electricity consumption in the pre-treatment period was about 7 kWh, during phase 1 it was about 6 kWh and in phase 2 it was about 8 kWh. By comparison, the US sample in Allcott (2011) consumed on average 19 to 60 kWh per day. For the external validity of our results and hence their policy relevance, we compare our in-sample consumption with that of the whole German population, which was 9 kWh in 2017.[4] We thus conclude that our sample in this regard is representative for Germany.

Cross-country variations in electricity consumption not only arise from differences in behavior; moreover, differences in the energy-mix also play an important role. In the case of Germany, more than 70% of total household energy use is devoted to heating, which is mostly generated from gas and oil. Electricity is only used for heating to a very small degree, whereby one might expect that the use of electricity in Germany is low compared with other countries. Indeed, official data confirms this: according to the World Bank, the annual electricity consumption per capita in Germany in 2014 was equal to about 7,000 kWh, whereas in the US this figure was equal to roughly 13,000 kWh.[5]

Given these differences in consumption levels, we should not be surprised to find different consumption changes in response to behavioral treatments across countries, with very low treatment effects in Germany. This conclusion is in line with the findings of Dünnhoff and Duscha's (2007) field experiment in the city of Heidelberg. The authors explain their non-significant treatment effect by the short-term nature of their nudge (Fischer, 2008). In light of the findings of the present study, an interpretation in terms of prior consumption levels is more likely, where the potential for picking low-hanging fruit is greater when electricity consumption is high.

---

[3] The maximum intervention time is often 4-6 weeks, Fischer (2008).

[4] Own calculations based on a collection of energy-data published by the German Federal Ministry for Economic Affairs and Energy. Online (download 31.05.2019): https://www.bmwi.de/Redaktion/DE/Binaer/Energiedaten/energiedaten-gesamt-xls.xls

[5] https://data.worldbank.org

5. Conclusion

This study provides evidence against the universal effectiveness of behavioral nudges on energy consumption. In particular, we show that increased consumption feedback and consumption feedback combined with a social comparison of consumption values does not reduce the electricity use in a sample of German households. As such, although the German billing system has strong scope for improvement in terms of the information delivered to customers, introducing nudges does not appear to be a fruitful climate policy measure.

Our conclusions are related to Gillingham et al. (2009), who acknowledge the raison d'être of behavioral nudges yet deny their practical usefulness for policy purposes. In a similar vein, Loewenstein and Ubel (2017) cautioned against relying on behavioral nudging as a policy tool due to the small energy savings that they cause. Our findings suggest that nudges are not always effective climate policy tools, at least in contexts where energy consumption is relatively low. Thus, governments should be cautious in their widespread implementation of behavioral nudges. As Andor and Fels (2018) mention, an interesting idea is therefore to use limited-scope policy experimentation before a general implementation. Our results provide additional evidence that such a more cautious, piece-meal policy approach is warranted.

However, Fischer (2008) highlight that the household sector has the fastest-growing end energy consumption in Germany and electricity consumption is the largest contributor to this trend. Hence, the efficacy of energy conservation measures – including nudges – may increase in the future. It is also important to emphasize that the effects of including a behavior nudge in e.g. the gas bill in Germany – which contains most of the households heating expenses – remains unknown and is left for future research. Finally, irrespective of the effects on total consumption, there may be other reasons to provide German consumers with more information about their energy consumption and its price, such as increasing demand flexibility, which reduces peak loads in the electricity distribution net and thus infrastructure costs.

References

Abrahamse, W., Steg, L., Vlek, C., Rothengatter, T., 2005. A review of intervention studies aimed at household energy conservation. Journal of environmental psychology, 25(3), pp. 273-291.

Allcott, H., Mullainathan, S., 2010. Behavior and energy policy. American Association for the Advancement of Science, 327 (5970), pp. 1204-1205.

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

Allcott, H., Kessler, J., 2019. The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons. American Economic Journal: Applied Economics, 2019, 11(1), pp. 236-276.

Andor, M.A, Fels, K.M., 2018. Behavioral Economics and Energy Conservation – A Systematic Review of Non-price Interventions and their Causal Effects. Ecological Economics, 148, pp. 178-210.

Asensio, O.I., Delmas, M.A., 2015. Nonprice incentives and energy conservation. Proceedings of the National Academy of Sciences of the United States of America, 2015, 112(6), pp. 510-515.

Ayres, I., Raseman, S., Shih, A., 2013. Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. Journal of Law, Economics and Organization, 29(5), pp. 992-1022.

Benartzi, S., Beshears, J., Milkman, K.L., Sunstein, C.R, Thaler, R.H., Shankar, M., Tucker-Ray, W., Congdon, W.J., Galing, S. 2017. Should Governments Invest More in Nudging? Psychological Science, 28(8), pp. 1041-1055.

BMWi and BMU, 2010. Energiekonzept für eine umweltschonende, zuverlässige und bezahlbare Energieversorgung [Energy concept for an environmentally-friendly, reliable and affordable energy supply]. Federal Ministry for Economic Affairs and Energy (BMWi) and Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU). More information (in German, downloaded 31.05.2019): https://www.umweltbundesamt.de/daten/energie/stromverbrauch

Costa D.L., Kahn M.E., 2010. Energy conservation “nudges” and environmentalist ideology: Evidence from a randomized residential electricity field experiment. NBER Working Paper, 15393.

Darby, S. 2006. The Effectiveness of Feedback on Energy Consumption. A Review for DEFRA on the Literature on Metering, Billing and Direct Displays. Environmental Change Institute. University of Oxford, Oxford.

Delmas, M.A., Fischlein, M., Asensio, O.I., 2013. Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. Energy Policy, 69, pp. 729-739.

Dünnhoff, E., Duscha, M., 2007. Effiziente Beratungsbausteine zur Minderung des Stromverbrauchs in privaten Haushalten. Endbericht [Efficient building blocks for energy counseling aimed at reducing electricity consumption in private households]. Study by the ifeu (Institut für Energie-und Umweltforschung), Heidelberg.

Ehrhardt-Martinez, K., Donnelly, K.A., Laitner, J.P. 2010. Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities (Report No. E105). American Council for an Energy-Efficient Economy, Washington, D.C.

EPRI. 2009. Residential electricity use feedback: a research synthesis and economic framework (report 1016844). Electric Power Research Institute, Palo Alto, CA.

Fischer, C., 2008. Feedback on household electricity consumption: a tool for saving energy? Energy efficiency, 1 (1), pp. 79-104.

Gillingham K., Newell R., Palmer K., 2009. Energy efficiency economics and policy. NBER Working Paper, 15031.

Gleerup, M., Larsen, A., Longo, A. 2013. The effect of feedback by text message (SMS) and email on household electricity consumption: experimental evidence. Energy Journal, 31(3), pp. 111-130.

Herrmann, M.R., Brumby, D.P., Oreszczyn, T., 2007. Watts your usage? A field study of householders’ literacy for residential electricity data. Energy Efficiency, 11(7), pp. 1703-1719.

Hirst, Eric; Brown, Marilyn, 1990. Closing the efficiency gap: barriers to the efficient use of energy, Resources, Conservation and Recycling, 3, pp. 267-281.

Jaffe, A., Stavins, R., Newell, R., 1995. Dynamic incentives of environmental regulations: The effects of alternative policy instruments on technology diffusion. Journal of the American Statistical Association, 105:490, pp. 493-505.

Kahneman, D., 2003. Maps of bounded rationality: Psychology for behavioral economics. The American Economic Review, 93 (5), pp. 1449-1475.

Karlin, B., Zinger, J.F., 2015. The Effects of Feedback on Energy Conservation: A Meta-Analysis. Psychological Bulletin, 141 (6), pp. 1205-1277.

Mack, B., Hallmann, S., 2004. Strom sparen in Lummerlund—eine Interventionsstudie in einer Passiv- und Niedrigenergiehaussiedlung. [Conserving electricity in Lummerlund. An intervention study in a passive and low energy house residential area]. Umweltpsychologie, 8(1), pp. 12–28.

Mont, O., Lehner, M., Heiskanen, E., 2014. Nudging: A tool for sustainable behavior? Report commission by the Swedish Environmental Protection Agency (6643/2014).

Loewenstein, G., Charter, N., 2017. Putting nudges in perspective. Behavioral Public Policy, 2017, 1(1), pp. 26-53.

Sanstad A., Hanemann M., Auffhammer M., 2006. End-use Energy Efficiency in a “Post-Carbon” California Economy: Policy Issues and Research Frontiers. The California Climate Change Center at UC Berkeley.

Schleich, J., Faure, C., Klobasa, M. 2017. Persistence of the effects of providing feedback alongside smart metering devices on household electricity demand. Energy Policy, 107, pp 225-233.

Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., Griskevicius, V., 2007. The constructive, destructive, and reconstructive power of social norms. Psychological Science, 18(5), pp. 429-434.

Sleich, J., Klobasa, M., Gölz, S., Brunner, M., 2013. Effects of feedback on residential electricity demand- Findings from a field trial in Austria. Energy Policy, 61, pp- 1097-1106.

Stern, P.C., 1992. What psychology knows about energy conservation. American Psychologist, 47(10), pp. 1224-1232.

Tedenvall, M., Mundaca, L., 2016. Behaviour, context and electricity use: Exploring the effects of real-time feedback in the Swedish residential sector. Paper presented at the 39th IAEE International Conference ‘Energy: Expectations and Uncertainty’ Bergen, Norway, 19-22 June 2016.

Wilhite, H., Ling, R., 1995. Measured energy savings from a more informative energy bill. Energy and Buildings, 22, pp. 145-155.

World Bank. Electric power consumption (kWh per capita) (downloaded on the 12th of June 2019): https://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC

Annex: Invitation letter

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

stadtwerke
göttingen AG
Kurs: Natürlich Zukunft!

Postanschrift:
Postfach 3834
37028 Göttingen
www.stadtwerke-goettingen.de

Stadtwerke Göttingen · Hildebrandstr. 1 · 37081 Göttingen

Herrn
Rolf Mustermann
Musterstraße 1
37085 Göttingen

Ihr Ansprechpartner:
Max Mustermann
Telefon: (0551) 301-xxx
Telefax: (0551) 327 68
kundenservice@swgoe.de

Göttingen, xx.xx.2017

Werden Sie mit Ihren Stadtwerken und der Uni Göttingen Teil einer innovativen Studie.
Testen Sie die neue GöStromApp!

Lieber Herr Mustermann,

Sie sind Nutzer eines Smartphones? Dann testen Sie unsere neue App und nutzen Sie die Chance auf einen monatlichen Geldgewinn.

Gemeinsam mit der Universität Göttingen starten wir am 01.06.2017 eine einjährige wissenschaftliche Studie zum Stromverbrauch in Privathaushalten. Für Ihre Teilnahme nutzen Sie die kostenlose App der Stadtwerke Göttingen, mit der Sie ganz einfach regelmäßig den Zählerstand Ihres Stromzählers erfassen. Jetzt zahlen Sie monatlich nur noch für Ihren tatsächlichen Stromverbrauch und haben eine kontinuierliche Übersicht über Ihre Stromkosten. Die eventuelle Überraschung einer Nachzahlung am Jahresende bleibt Ihnen also erspart!

Ihre über die App erfassten Verbrauchsdaten werden innerhalb eines geschlossenen Systems anonymisiert, streng vertraulich und ausschließlich im Rahmen dieser wissenschaftlichen Studie der Universität Göttingen verwendet.

Unter allen Studienteilnehmern verlosen wir monatlich x mal xxx Euro als Dankeschön dafür, dass Sie uns und die Universität Göttingen auf unserem innovativen Weg unterstützen.

Neugierig? Dann laden Sie sich bis zum TT.MM.JJJJ die App der Stadtwerke Göttingen im App- oder Google Play-Store herunter und registrieren sich mit

Kundenummer Stadtwerke Göttingen: 123456 (Abb. App)
Registrierungs-Code für die Studie: XyZG

Im Anschluss erhalten Sie über die App detaillierte Informationen zu Studienbeginn und Ablauf.

Herzlichen Dank für Ihre Teilnahme, willkommen in unserem innovativen Team und viel Erfolg bei unserem monatlichen Gewinnspiel!

Mit herzlichen Grüßen
Ihre Stadtwerke Göttingen AG

Max Mustermann
Musterabteilung

Stadtwerke Göttingen AG · Sitz Göttingen · Amtsgericht Göttingen, HRB-Nummer 1439 · Steuer-Nr. 20/200/27172 · USt-IdNr.: DE115304695
Vorsitzender des Aufsichtsrates: Oberbürgermeister Rolf-Georg Köhler · Vorstand: Dr. Gerd Rappenecker, Dipl.-Kfm. Frank Wiegelmann
Sparkasse Göttingen · IBAN: DE74 2605 0001 0015 1002 90 · BIC: NOLADE21GOE

[English translation]

Stadtwerke Goettingen and the university of Goettingen are starting an innovative study.
Be part of it and test the new UniGöEnergieApp!

Dear Mr./Mrs./Ms. Xxx

Do you use a smartphone? Then test the new **UniGöEnergieApp** and get the opportunity to win cash every month.

Together with the university of Goettingen, we are starting a one year scientific study on household electricity use on the 1st of June 2017. By participating you will use the free UniGöEnergieApp to easily and regularly record your electricity meter. In this way, you pay every month only for your actual consumption und have a complete overview of your electricity costs. This saves you possible additional payments in the future.

Your consumption information gathered via the App are stored within a closed, anonymized data system. They will only be used in connection with the scientific study conducted together with the university of Goettingen.

Among all participants in the study, we lot out **5 times 100 euro** monthly as a 'thank you' for helping us and the university of Goettingen to follow innovative paths.

Curious? Then do download the **UniGöEnergieApp** before the **31th of May** either in App- or in Google Playstore. To register, please use:

> **Customer number Stadtwerke Göttingen: xxx**
> **Registration code for the study: yyy**

Following your registration, you will receive detailed information about the study via the App. Thank you for participating, welcome in our innovative team and good luck for the monthly lotteries!