Social Norms and Energy Conservation Beyond the US

Mark A. Andor,* Andreas Gerster,* Jörg Peters† and Christoph M. Schmidt‡

August 14, 2018

The seminal studies by Allcott and Mullainathan (2010), Allcott (2011), and Allcott and Rogers (2014) show that social comparison-based home energy reports (HER) are a cost-effective climate policy intervention in the US. Our paper demonstrates the context-dependency of this result. In most industrialized countries, average electricity consumption and carbon intensity are well below US levels. Consequently, HER interventions can only become cost-effective if treatment effect sizes are substantially higher. For Germany, we provide evidence from a large-scale randomized controlled trial that effect sizes are in fact considerably lower than in the US. We conclude by illustrating that targeting highly responsive subgroups is crucial to reach cost-effectiveness and by identifying the few countries in which HER are promising policy instruments.

JEL codes: D12, D83, L94, Q41.

Keywords: Social norms; energy demand; external validity; randomized field experiments; non-price interventions.

Correspondence: Mark Andor, RWI – Leibniz Institute for Economic Research, Hohenzollernstr. 1-3, D-45128 Essen. E-mail: andor@rwi-essen.de.

Acknowledgements: We gratefully acknowledge financial support by the Stiftung Mercator, without which this project would not have been possible. Furthermore, this work has been partly supported by the Collaborative Research Center *Statistical Modeling of Nonlinear Dynamic Processes* (SFB 823) of the German Research Foundation (DFG), within Project A3, *Dynamic Technology Modeling*. Peters acknowledges the support of a special grant (Sondertatbestand) from the German Federal Ministry for Economic Affairs and Energy and the Ministry of Innovation, Science, and Research of the State of North Rhine-Westphalia. We are grateful for valuable comments and suggestions by Hunt Allcott, Stefano Carattini, Fredrik Carlsson, Paul Ferraro, Manuel Frondel, Kenneth Gillingham, Timo Goeschl, Lorenz Götte, Martin Kesternich, Andreas Lange, Bruno Lanz, Grischa Perino, Michael Price, Stephan Sommer, and Colin Vance as well as participants of the 32nd Annual Congress of the European Economic Association, the 6th Mannheim Energy Conference, the 23rd Annual Conference of the European Association of Environmental and Resource Economists, the 6th World Congress of Environmental and Resource Economists, and the seminars at ZEW, University of Passau, and University of Hamburg. We thank Vera Huwe and Lukas Tomberg for excellent research assistance.

*RWI – Leibniz Institute for Economic Research.

†RWI – Leibniz Institute for Economic Research and University of Passau.

‡RWI – Leibniz Institute for Economic Research and Ruhr-University Bochum.

1. Introduction

A large literature has shown that social norms can affect individual choices in a variety of domains such as water use (Brent et al. 2015, Ferraro et al. 2011, Ferraro and Price 2013, Jaime Torres and Carlsson 2016) and charitable giving (Frey and Meier 2004, Shang and Croson 2009). A prominent example are the social comparison-based home energy reports (HER) that the company *Opower* mailed to millions of US households in order to reduce their electricity consumption. Evaluations of *Opower's* HER have documented considerable reductions in electricity consumption of 1.4-3.3% that are also persistent over time (Allcott 2011, Allcott and Rogers 2014). Given their low intervention costs, HER are said to be a cost-effective policy instrument to combat climate change (Allcott 2011) and often serve as a prime example for behavioral interventions that can be scaled to "millions of consumers" (Allcott and Mullainathan 2010).[1]

In this paper, we assess the cost-effectiveness of HER interventions in industrialized countries other than the US. While our paper does not question the internal validity of the Allcott (2011) and Allcott and Rogers (2014) findings, we complement their work by testing the external validity and transferability to other contexts. We thereby also contribute to the growing literature on the context-dependency of causal effects measured in a particular policy population and on the scalability of experimental findings (e.g. Allcott 2015, Al-Ubaydli et al. 2017a;b, Dehejia et al. 2015, Gechter 2016, Peters et al. 2018, Vivalt 2015).

In Section 2, we first show that US electricity consumption and carbon intensity levels exceed those in virtually all other OECD countries. We demonstrate that this particular context decisively co-determines the cost-effectiveness of HER in the US. Consequently, treatment effects would have to be substantially larger in other countries in order to make HER a cost-effective climate policy intervention. We show that at effect sizes observed in the *Opower* studies, HER can only become cost-effective in the US and Australia.

In a next step, we test for the effectiveness of HER in Germany, a country with electricity consumption levels that compared to the US match the OECD average more closely (Section 3). Based on data from a randomized controlled trial (RCT) among 11,630 households, we find that HER reduce electricity consumption by 0.7%, less than half of the average reductions observed in the *Opower* studies. HER are thus clearly not cost-effective. We furthermore

[1]Combined, Allcott and Mullainathan (2010) and Allcott (2011) have already 605 + 1,407 = 2,012 citations (Google Scholar, 07/24/2018).

examine effect heterogeneity and find that higher consumption households are considerably more responsive, which indicates that the lower effect size is partly due to the lower baseline consumption levels in Germany. Hence, for the majority of industrialized countries exhibiting similarly low consumption levels, effect sizes can be expected to be much lower than in the US, as well.

In Section 4, we build on our heterogeneity analysis across baseline consumption levels to explore the potential of targeting particularly responsive subgroups, such as high-consumption households. To this end, we calculate the effect sizes required to make HER cost-effective and compare them to effect sizes observed in high-consumption strata. We find that for many countries, such as the United Kingdom, South Korea, Canada, Germany, and Poland, targeting HER can be a promising policy instrument, but only if targeted consumer groups exhibit effect sizes between 5% and 7%. In Germany, we estimate an effect of 3.1% in the highest decile, suggesting that even the top decile is not as responsive as required.

In Section 5, we explore the sensitivity of our results to treatment effect persistence and a broader welfare perspective. HER might not only decrease consumption in the year of the treatment but also beyond. Effect persistence considerably amplifies the benefits of HER, which is likely to make them cost-effective in the US, Australia, and Japan, but not in European countries. A further robustness check examines the broader welfare effects following Allcott and Kessler (2015). This approach suggests taking into account market imperfections beyond externalities from carbon emissions. No clear picture emerges from this extended welfare analysis. We show that assumptions to gauge electricity market distortions critically determine whether the welfare perspective improves or worsens the assessment of HER. In the concluding section, we summarize the findings and identify countries in which targeted HER are particularly promising. In the concluding section, we summarize the findings and identify those countries in which targeted or non-targeted HER are particularly promising.

2. Context-Dependency of HER Interventions

By means of a descriptive analysis, this section shows that the cost-effectiveness of HER as a climate policy intervention crucially depends on two factors, the baseline consumption levels as well as the carbon intensity of electricity generation. Higher average electricity consumption levels translate a given effect size (in relative terms) into higher absolute electricity savings in terms of kilowatt-hours (kWh). In addition, households with higher consumption levels tend to exhibit lower behavioral and technical efficiencies and can thus realize higher effect sizes. The second factor, the carbon intensity of electricity generation, determines the extent to which electricity savings translate into mitigation of carbon dioxide (CO_2) emissions. For example, power sectors with large shares of lignite- and hard coal-fired power plants are much more carbon-intensive than those relying on hydropower and nuclear energy. As a result, even similar absolute electricity savings can yield widely diverging CO_2 abatement effects across countries.[2]

We consider the ten OECD countries with the largest total residential electricity consumption (WEC 2016). In addition, we include Poland, an OECD country with low average electricity consumption but very carbon-intensive electricity generation, and Sweden, a country with high average consumption but low carbon intensity. We project HER's cost-effectiveness in these countries, assuming treatment effects of 1.4-3.3%, which corresponds to the full range of effect sizes observed in the *Opower* studies (Allcott 2011). Average electricity consumption levels and carbon intensities are drawn from official data sets (WEC 2016 and IEA 2016).[3] For simplicity, we assume printing and mailing cost of 1 US$ per letter in all countries (as in Allcott 2011) and neglect any administrative cost. As quarterly reports have achieved the highest cost-effectiveness in previous studies (Allcott 2011), we presume that four reports are sent within one year.

We then compare the HER abatement cost to the avoided social cost of carbon, the usual yardstick to assess the cost-effectiveness of climate change mitigation policies (Greenstone et al. 2013, Nordhaus 2014). For simplicity, we denote an intervention as cost-effective if its abatement cost fall below the social cost of carbon. Nordhaus (2014) estimates the social cost of carbon to be 19 US$ per ton in 2015, while US IAWG (2013) provide an estimate of 38 US$. For our assessments, we use the larger estimate of 38$, which is more favorable to the cost-

[2] There is a literature that discusses whether cap-and-trade schemes, such as the European Union Emissions Trading System, and energy saving policies, like HER, are complements (e.g. Dietz et al. 2009) or rather substitutes (e.g. Goulder 2013). Some even argue that energy saving policies that shift demand away from sectors subject to a cap increase aggregate emissions (Perino 2015). We take the most favorable stance for HER and treat them as complements to cap-and-trade schemes in our cost-effectiveness calculations.

[3] We use average carbon intensities – rather than marginal carbon intensities – to approximate abated carbon emissions, as evidence on the timing of electricity savings from HER is missing. The type of marginal power plant often changes within a day, so that evidence on the timing of savings is a prerequisite for determining marginal carbon emissions. In Germany, for example, coal-fired power plants (822 g CO_2 per kWh, UBA 2018) are price-setting in the off-peak period, while gas-fired power plants (353 g CO_2 per kWh, UBA 2018) are price-setting in the peak period. Furthermore, current marginal emissions can also be misleading as, in the long run, electricity savings ultimately induce capacity adjustments in the electricity market. Using the average carbon intensity reflects an agnostic stance on the induced capacity adjustments.

Table 1: International Cost-Effectiveness Comparison of HER Interventions

	(1)	(2)	(3)	(4)	(5)
Country	Average Electricity Consumption in kWh	CO_2 Emissions in g / kWh	Cost in Cent / kWh Saved	Abatement Cost in $ / t CO_2	CO_2 Abatement Cost Relative to US
Canada	11,379	158	1.1 – 2.6	67 – 162	3.3
United States	12,293	489	1.0 – 2.4	20 – 49	1.0
France	5,859	64	2.1 – 5.0	323 – 779	16.0
Germany	3,304	486	3.7 – 8.8	75 – 182	3.7
Italy	2,542	343	4.8 – 11.5	139 – 335	6.9
Poland	1,935	769	6.3 – 15.1	81 – 196	4.0
Spain	4,040	247	3.0 – 7.2	121 – 293	6.0
Sweden	8,025	13	1.5 – 3.6	1,162 – 2,799	57.6
United Kingdom	4,145	459	2.9 – 7.0	64 – 153	3.2
Japan	5,434	572	2.2 – 5.4	39 – 94	1.9
South Korea	3,489	536	3.5 – 8.4	65 – 156	3.2
Australia	6,959	798	1.7 – 4.2	22 – 53	1.1

Notes: Our calculations assume printing and mailing cost of 1 US$ per report, four reports per year and average electricity reductions of 1.37-3.30%. Average electricity consumption and CO_2 intensities of electricity generation correspond to the most recently available data (for 2013), as documented in WEC (2016) and IEA (2015), respectively.

effectiveness of HER. Table 1 summarizes the descriptive statistics and our cost-effectiveness indicators. Average electricity consumption levels and carbon intensities are depicted in Column (1) and (2). By dividing the cost of HER by the electricity savings for the range of effect sizes, we obtain a range of cost per kWh saved, which can be found in Column (3).

Column (4) of Table 1 depicts the resulting abatement cost ranges. It shows that HER are considerably less cost-effective in most countries other than the US. In fact, our projections suggest that no country except the US and Australia reaches abatement cost levels that would justify the use of HER as a policy instrument to combat climate change when using social cost of carbon of 38 US$ per ton as a yardstick. This finding demonstrates that the potential of HER as a cost-effective climate policy instrument hinges strongly upon both average electricity consumption levels and carbon intensities of electricity generation.

Our interim conclusion holds for the treatment effect sizes reported for the US but might change if HER were more effective in other countries. To explore this possibility, Column (5) presents a summary indicator: the CO_2 abatement cost per country relative to those in the US at a given treatment effect size. In Germany, for example, abatement cost of HER exceed those in the US by a factor of 3.7. In other words, to reach the same abatement cost level as in the *Opower* study, the effect size of HER in Germany would have to be at 6.3% ($3.7 \times 1.7\%$, assuming the estimated average effect size of 1.7% reported in Allcott 2011). For France and Sweden, the abatement costs surpass those in the US by a factor of 16.0 and 57.6, respectively, implying that HER interventions cannot become cost-effective even under optimistic assumptions. These abatement cost ratios exceed 1.9 for all considered OECD countries except Australia so that effect sizes would have to be (at least) twice as high to reach the same CO_2 abatement cost as in the US.

3. Treatment Effects of HER in Germany

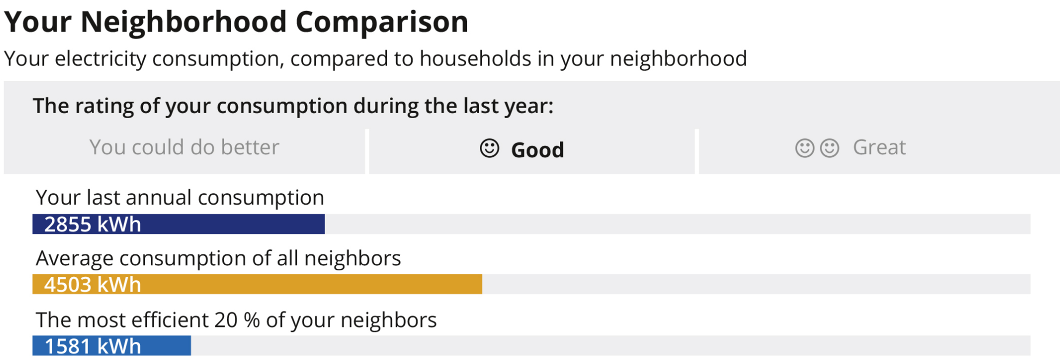
3.1. HER Design in the German Experiment

In cooperation with the German firm *Grünspar*, a service provider for utilities, we designed the HER for our study in a way that matches the *Opower* intervention closely, but not perfectly. We conducted a "natural field experiment" (Levitt and List 2009) to the extent that households were not informed about the experiment. Households in the treatment group received four quarterly letters, while households in the control group did not receive any letter beyond the utilities' regular communication. Just as the *Opower* reports, our HER provided electricity-saving tips and compared the household's consumption with that of its neighbors, as visualized in Figure 1.[4] The four HER are comprehensively documented in Appendix C. Our HER differs from the *Opower* HER with respect to the following features: First, our reports announce an individualized electricity consumption goal for each recipient household (10% less than the previous year) and additionally offer rebates for the purchase of energy-efficient appliances.[5] As goal setting and rebates generally have positive effects on the uptake of energy-efficient durables and energy conservation (e.g. Davis et al. 2014, Harding and Hsiaw 2014), we expect this modification to slightly intensify the HER's effectiveness.

[4]For a comprehensive overview of the literature on the effects of social comparisons to encourage resource conservation see, for example, Abrahamse and Steg (2013) and Andor and Fels (2018).

[5]Table A1 in the appendix provides a detailed comparison between our HER and the one sent out by *Opower*.

Figure 1: Social Comparison Element in HER (Translation of the Original German Version)



Second, while meters in the *Opower* population were read on a monthly or quarterly basis and electricity consumption information could thus be updated in each letter, annual meter reading cycles in Germany do not allow for such intra-year updates. Therefore, we repeated the social comparison based on the most recently available annual electricity consumption in the first three reports. We expect that the absence of intra-year updates might lead to a slight decrease of the energy conservation effect. For example, Gilbert and Zivin (2014) find that receipt of an electricity bill increases households attention and reduces electricity consumption.[6] Third, the last of the four HER contains testimonials, i.e. stylized examples of electricity-saving actions implemented by hypothetical households ("We have recently bought a new energy-efficient refrigerator that saves us around 60 EUR per year", pictures of testimonials can be found in Appendix C).

3.2. Experimental Design and Data

For the implementation of the study, we cooperated with a medium-sized electricity provider located in Kassel, Germany, with around 135,000 residential customers. In total, the trial was implemented in a study sample of 11,630 residential electricity consumers. We randomized the HER intervention among those households that received their annual bill between November 2014 and April 2015. Households in the treatment group received the four HER within one year, while households in the control group did not receive any report or communication other than the business-as-usual correspondence with the electricity provider. We sent the first HER shortly after the households annual meter reading, which serves as the baseline of our analysis.

[6]Yet, intra-year updates could also reduce the conservation effect, as Wichman (2017) finds that customers increase water consumption in response to more frequent billing. The interaction effect of billing and social comparisons might be an interesting topic for future research.

Table 2: Balance of Baseline Characteristics Between Treatment and Control Group

	Treatment Group	Control Group	Difference (Std. Error; p-Value)
Baseline Consumption (2014), in kWh	2,281.3	2,279.3	2.0 (29.62; 0.95)
Baseline Billing Period Length, in Days	362.7	362.6	0.1 (0.12; 0.47)
Number of Households	5,808	5,812	

Endline data is retrieved from the annual business-as-usual metering one year later, which took place in the period between November 2015 and April 2016. Our sample includes only those households that had been with the electricity provider for at least one year so that we can make use of baseline consumption data. In addition, as in the *Opower* programs, households with more than one meter were excluded from the sample. The randomization was stratified by households baseline electricity consumption and billing month.

Table 2 illustrates that baseline electricity consumption and billing period length are perfectly balanced across treatment and control group. Furthermore, the table shows that households in our sample consume on average 2,300 kWh per year and hence far less than the 12,000 kWh consumed by the average US household but also less than the German average of around 3,300 kWh per year (depicted in Table 1). The reason for the latter is that our studys target area, Kassel and its surrounding suburbs, is an urban area where households are typically smaller than the average German household. We examine the representativeness of our study region compared to the rest of Germany by using socio-demographic variables at the regional zip code level, provided by the geomarketing firm *microm Consumer Marketing*. Beyond the lower electricity consumption, we find that our sample provides a fair representation in terms of key socio-demographic variables, such as percentage of retirees, unemployment rate, purchasing power, and non-German citizenship (see Table A2 in the appendix).

Table 3: Average Treatment Effect (ATE) on Households' Electricity Consumption

	(1)	(2)	(3)
ATE	-0.719*	-0.620**	-0.684**
Standard Error	(0.383)	(0.270)	(0.269)
95% Conf. Interval	[-1.469,0.032]	[-1.150,-0.090]	[-1.212,-0.156]
Outliers Removed	-	✓	✓
Time Controls	-	-	✓
R^2	0.0003	0.0005	0.025
Number of Obs.	11,620	11,388	11,388

Notes: The outcome variable is the change in a household's annual electricity consumption between the treatment and baseline period, divided by the average control group consumption in the post period (both in kilowatt-hours), and multiplied by 100 to ease the readability as percentage changes. **,* denote statistical significance at the 5% and 10% level, respectively. Heteroscedasticity robust standard errors are in parantheses.

4. Results

To determine the Average Treatment Effect (ATE) of the HER on electricity consumption, we estimate the following differences-in-differences regression model:

$$\Delta Y_i = \alpha + \delta T_i + \epsilon_i \tag{1}$$

where $\Delta Y_i = (Y_i^{2015} - Y_i^{2014})/Y_{i,c}^{2015}$ corresponds to the change in the annual electricity consumption of household i before (Y_i^{2014}) and after the HER treatment (Y_i^{2015}), T_i is the treatment dummy variable that equals unity for households that received the HER, and ϵ_i designates an idiosyncratic error term. To account for different billing period lengths, we normalize all yearly electricity consumptions to 365 days. Furthermore, we divide the difference between annual consumptions by the average post-period control group consumption, $Y_{i,c}^{2015}$, so that the ATE identified by δ expresses average electricity savings as a percentage of the average consumption level. We use this normalization throughout our analysis as it facilitates cost-effectiveness calculations. In particular, it allows to easily determine the total electricity savings in a population by multiplying the average electricity consumption level by the ATE.

Table 3 presents the results and shows that the average HER treatment effect is a 0.7% reduction and statistically significant at the 10 percent level (Column 1). In Columns (2) - (3) we trim our sample by excluding outliers and include weekly time dummies for both the baseline and the treatment period as further control variables.[7] This change does not markedly alter the effect size but increases precision considerably.

Moreover, throughout the specifications of Table 3, the 95% confidence intervals allow us to exclude average reductions in electricity consumption larger than 1.5% – and hence nearly the entire range of effect sizes that have been documented by Allcott (2011) for the US. Because of the large differences in average consumption levels between German and US households, absolute electricity savings from HER diverge even more strongly. Our ATE of around 0.7% translates into an absolute average electricity reduction of around 16 kWh per year or 0.04 kWh per day, which is equivalent to turning off a 30 Watt light bulb for some 90 minutes every day. For comparison, Allcott (2011) estimates an electricity conservation effect of 1.7% for quarterly reports that translates into absolute savings of 191 kWh per year (0.52 kWh per day) in the US.

Updating our cost-effectiveness analysis from Section 2 shows that our treatment effect estimates imply intervention costs per saved kWh of around 0.25 US($4US/16$ kWh), compared to only around 0.01-0.05 US$ in the US *Opower* samples (Allcott 2011). Dividing the cost estimates by the carbon intensity – which is virtually on par for the US and Germany – yields the cost per mitigated ton of CO_2 of 505 US$ in Germany, compared to 25-105 US$ in the US *Opower* samples. This finding reinforces our assessment that HER are not a cost-effective climate policy instrument in Germany.

Treatment Effect Heterogeneity

The considerably higher absolute consumption levels in the US might also partly explain the higher treatment effects in relative terms. As Figure 4 in Appendix A illustrates, US households consume at least twice the electricity of German households in every domain except for cooking. In particular, space cooling accounts for more than 2,000 kWh in the US but is virtually absent in Germany. Such differences can have important implications for saving potentials. For example, Allcott and Rogers (2014) find that HER are particularly effective in winter and summer when US households respond by reducing electricity consumption for heating and cooling, respectively.

[7] We removed outlier observations when the change in the electricity use from 2014 to 2015 falls below the 1% percentile (-68.1%) or exceeds the 99% percentile (+52.7%). Such large changes in consumption can arise from temporarily uninhabited dwellings, for example, and are very unlikely a consequence of receiving HER. To control for time effects, we include weekly dummies that equal 1 if a week falls into the respective billing period. Weekly dummies are included for both billing periods.

Table 4: Average Treatment Effects (ATE) in Subsamples Based on Households' Baseline Consumption in 2014

	(1) Full Sample	(2) Above Median $Y_i^{2014} > p_{50}(Y_i^{2014})$	(3) Above top Quartile $Y_i^{2014} > p_{75}(Y_i^{2014})$	(4) Above top Decile $Y_i^{2014} > p_{90}(Y_i^{2014})$
ATE	-0.684**	-0.910**	-1.544**	-3.058**
Standard Error	(0.269)	(0.454)	(0.725)	(1.347)
95% Conf. Interval	[-1.212,-0.156]	[-1.801,-0.020]	[-2.965,-0.123]	[-5.702,-0.415]
Outliers Removed	✓	✓	✓	✓
Time Controls	✓	✓	✓	✓
R^2	0.025	0.037	0.067	0.133
Number of Obs.	11,388	5,695	2,847	1,139

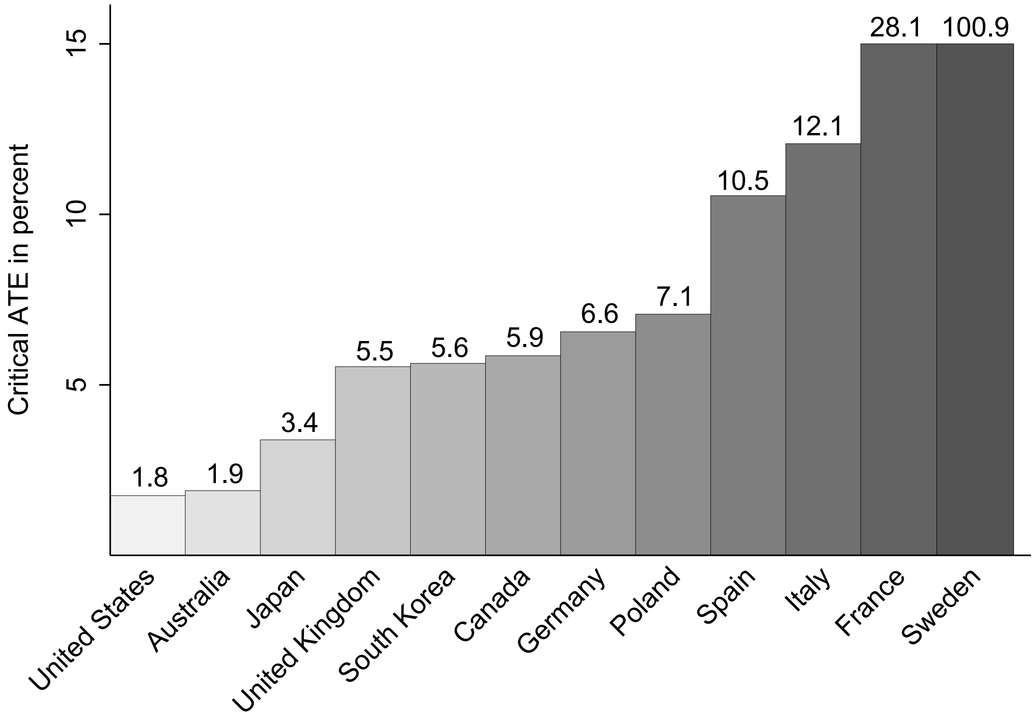
Notes: The outcome variable is the change in a household's annual electricity consumption between the treatment and baseline period, divided by the average control group consumption in the Post period (both in kilowatt-hours), and multiplied by 100 to ease the readability as percentage changes. ** denote statistical significance at the 5% level. Heteroscedasticity robust standard errors are in parantheses. The median of baseline consumption, $p_{50}(Y_i^{2014})$, is 1,883 kWh, the top quartile is 2,794 kWh and the top decile is 3,921 kWh.

To test empirically whether households with high consumption levels have larger saving potentials, we estimate the ATE in subsamples of households that use more than the median, the top quartile, and the top decile of baseline electricity consumption, respectively. Indeed, as Table 4 illustrates, households with larger baseline consumptions realize higher average savings, again expressed as a percentage of the average consumption in the population. For households above the median, we observe a statistically significant reduction of 0.9%. The treatment effect even reaches 1.6% and 3.1% in the top quartile and top decile, respectively, and remains statistically significant in spite of the decreasing sample size.

5. Targeting

The heterogeneity analysis has shown that effect sizes increase considerably with baseline electricity consumption. Top decile households are much more responsive with an effect size of more than 3%. Likewise, Allcott (2011) estimates an effect size of 6% in the top decile of the *Opower* sample in the US, which is substantially higher than the overall ATE. This treatment heterogeneity suggests that there is considerable scope for targeting particularly responsive subgroups, such as households with high baseline electricity consumption.

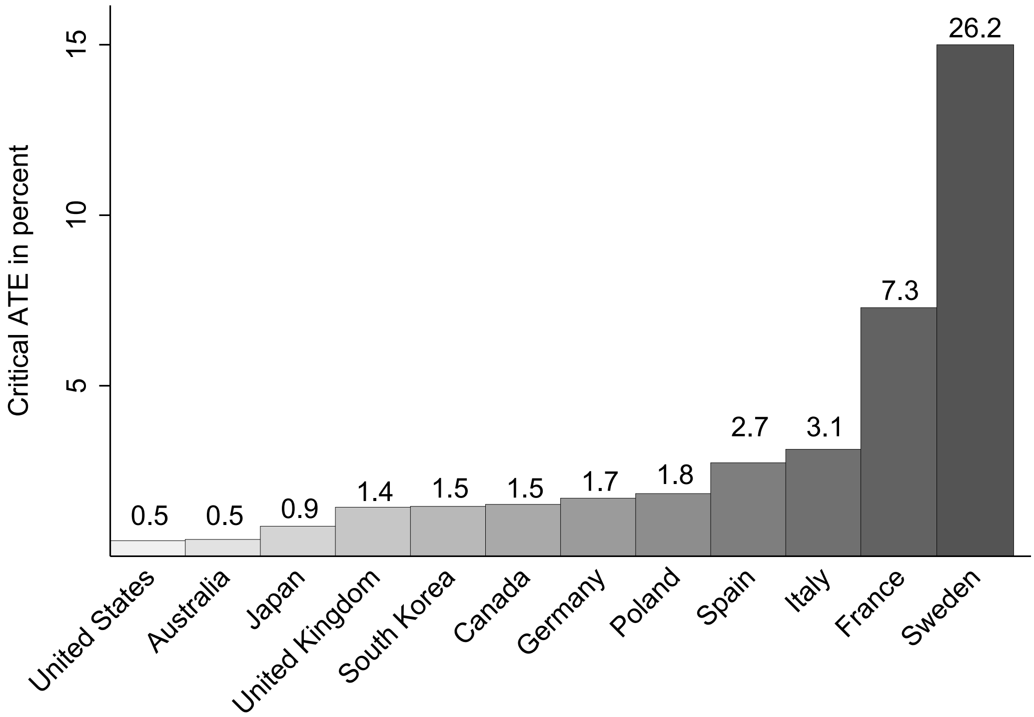
Figure 2: Critical ATE, by Country



Notes: To ease readability, bars are capped at critical ATEs of 15%.

In order to explore the cost-effectiveness of targeted interventions for a set of OECD countries, Figure 2 shows the minimum effect size required to bring down abatement costs of HER below the 38 US$ social cost of carbon yardstick ("critical ATE" in the following). A cost-effective targeting strategy in the respective country would have to identify consumer groups with a responsiveness at least as high as this critical ATE. In Germany, for example, HER would only be cost-effective in groups that exhibit an effect size of at least 6.6%. In our RCT, even the highest decile of electricity consumers reduced their consumption by about 3% (see Table 4). Taking into account that baseline electricity consumption in our sample is below the German average, effect sizes in the top decile might be larger in other parts of the country but would need to double to make targeted HER cost-effective. The UK, South Korea, Canada, and Poland have critical ATEs akin to Germany. Hence, targeting in these countries would have to identify highly responsive consumers that realize about the 6% effect size that has been documented for top decile consumers in the US (Allcott 2011). The critical ATEs in Spain, Italy, and especially France and Sweden reach orders of magnitude where even a well-targeted HER intervention cannot become cost-effective.

Figure 3: Critical ATE in the year of the treatment under persistence (15% Annual Reduction of Effect Size), by Country



Notes: Critical ATEs are calculated based on the assumption that effect sizes attenuate linearly by 15 percentage points per annum and that discount rates are zero. To ease readability, bars are capped at critical ATEs of 15%.

6. Effect Persistency and Welfare

Our cost-effectiveness projections so far have relied on a set of assumptions, most of which favor the cost-effectiveness of HER. For example, we have ignored administrative cost and considered the maximum effect size found in the literature as well as the larger social cost of carbon estimate of 38 US$ per ton. In this section, we examine two further aspects that might alter the HER cost-effectiveness balance: persistency in electricity savings beyond the treatment year and the consideration of a broader set of benefits and costs for welfare analysis.

Hitherto, we have assumed that HER effects only induce electricity savings in the year of the treatment. For the US, though, there is evidence that effects persist over time (Allcott and Rogers 2014 and Brandon et al. 2017). To calculate the HER abatement costs under an optimistic persistence scenario, we assume that effect sizes attenuate linearly by 15 percentage points per annum (Allcott and Rogers 2014) and set discount rates to zero. Electricity savings then increase by a factor of 3.85 implying that abatement costs and thus the critical ATEs decrease accordingly. Referring to the corridor of empirical effect sizes (0.7% in Germany up to a maximum of 3.3% in the US), Figure 3 indicates that HER reach potentially cost-effective levels in all countries but Sweden and France. Especially for the US, Australia, and Japan, the critical ATEs appear to be very promising under this persistence scenario. For the remaining countries, critical ATEs must still reach levels that are substantially higher than the 0.7% observed in our German sample. Given the rather low baseline consumption levels in those countries, substantially higher effect sizes are unlikely in population-wide interventions. Yet, persistence further strengthens the scope for HER targeted at more responsive subgroups as outlined in the previous section.

Following Allcott and Kessler (2015), we now conduct a broader welfare analysis that incorporates more factors than merely the benefits from CO_2 abatement and postage and printing cost. HER can trigger additional welfare effects beyond the reduction of carbon externalities due to demand- and supply-side imperfections, such as a dead-weight loss induced by markups on production costs. On the demand side, consumers benefit, for example, from electricity savings in form of reduced energy costs but bear adjustment costs such as investments for more efficient appliances and loss of utility from forgone energy consumption.

This broader welfare analysis is complex and presented in more detail in Appendix B. To gauge unobserved demand side costs and benefits, we refer to Allcott and Kessler (2015), who estimate the net welfare effect of HER on consumers by eliciting the recipients' willingness-to-pay (WTP) for a continued delivery of HER. They find that recipients' average WTP amounts to 54% of the electricity cost savings from receiving HER. While the remaining 46% reflect the consumers costs from receiving HER, the WTP of 54% of the electricity cost savings implies that HER induce welfare gains by reducing electricity consumption as they are alleviating so-called internalities (see, for instance, Allcott and Kessler 2015, Allcott 2016).

For the supply side, we retrieve country-specific electricity prices as well as production costs and use their difference to approximate the dead-weight loss. As production costs in the electricity sector hinge critically on the considered time horizon, we approximate them by two measures: the average spot price, which measures short-term production cost, and the levelized cost, i.e. per kWh cost of building and operating a power plant over its lifetime. Table A4 in the appendix presents our welfare calculations for both types of production costs. When we use levelized cost to approximate production cost, we find that welfare increases in the US, Canada, France, Poland, Sweden, and South Korea, compared to the results in Section

2. In contrast, when we approximate production cost by spot prices, the welfare assessment becomes only more favorable for South Korea.

7. Conclusion

This paper has examined whether the core result from the *Opower* experiments in the US, the cost-effectiveness of social comparison-based home energy reports (HER) as a climate policy instrument, generalizes to other contexts. In a descriptive analysis, we show that lower electricity consumption levels and carbon intensities of electricity generation reduce the cost-effectiveness potentials of HER considerably in all industrialized countries, except for Australia. In particular for European countries, we demonstrate that treatment effects have to be much larger than those observed in the US to make HER cost-effective.

Moreover, by means of a large-scale randomized controlled trial in Germany, we provide evidence that HER effect sizes are decidedly lower than in the US. Our estimates imply an average reduction in electricity consumption of 0.7%, half of what was observed as minimum effect size in the US. We show that one reason for this deviation is the lower baseline electricity consumption in Germany, which in turn suggests that effect sizes will also be much lower in other industrialized countries with similarly low consumption levels.

We furthermore analyze two aspects that could improve the appeal of HER. First, we show that targeting HER, for instance to high-consumption households, can be a promising policy instrument to reach cost-effectiveness for many countries, such as the United Kingdom, South Korea, Canada, and Poland, but only if highly responsive target groups are identified. For Germany, our empirical estimates imply that even the highest decile is not responsive enough to render HER cost-effective. Second, we illustrate the influence of effect persistence beyond the year of the treatment and show that HER with persistent conservation effects can become cost-effective in all industrialized countries, except for France and Sweden, where almost carbon-free power mixes prevent HERs cost-effectiveness. Yet, in all countries except the US, Australia and Japan, effect sizes would have to exceed 1.4%, which is substantially more than the average treatment effect of 0.7% that we estimate in our German sample. These more optimistic conclusions have to be interpreted with some care, given that in all analyses throughout the paper we used the higher social cost of carbon estimate as a yardstick (38 US$

per ton). Moving more towards the lower estimate (19 US$ per ton) renders HER ineffective again even under generous persistence and targeting assumptions.

From a policy perspective, our results show that HER are not a silver bullet for climate policy in most European countries. Beyond the US, we only identify Australia and Japan as promising countries for a large-scale rollout of HER. Yet, we also show that even in Europe, targeting and effect persistence can render HER cost-effective, making this an important field for future research. Furthermore, more evidence on the marginal abated emissions through HER is important, as emissions vary considerably by location and daytime (cf. Zivin et al. 2014), which can have strong impacts on the evaluation of environmental policies (cf. Holland et al. 2016a;b). In addition, future evaluations should go beyond cost-benefit analyses and provide a comprehensive welfare assessment in the vein of Allcott and Kessler (2015) embracing all benefits and costs of an intervention. Not least, the context variables we have identified – electricity consumption levels and carbon intensities – are also relevant for the assessment of other climate policy instruments targeting residential electricity demand. Other behavioral instruments as well as taxes and subsidies are arguably less cost-effective in European countries, while climate policy instruments that are cost-effective outside the US will likely be cost-effective in the US.

On a more general note, our paper contributes to the emerging literature on the transferability of causal effects across settings and on the scalability of experimental findings (Allcott 2015, Al-Ubaydli et al. 2017a;b, Davis et al. 2017, Deaton and Cartwright 2016, Gechter 2016, Hotz et al. 2005, Leviton 2017, Muller 2015, Peters et al. 2018, Pritchett et al. 2015, Vivalt 2015). Even if a proof-of-concept is furnished with high internal validity for one policy population – as it was done in different evaluations of the *Opower* case for the US – the transferability to other policy populations can prove difficult. We complement the literature by explicitly demonstrating that not only the average treatment effect is context-dependent but also other components of a cost-effectiveness analysis. In particular, we provide an example that differences in important context features can preclude cost-effectiveness of the intervention already a priori in some countries. Yet, prudence in terms of external validity also applies to the present paper: Digitalization of daily routines through smart metering and digital information gadgets might alter both the costs and the effectiveness of information campaigns for energy conservation (see Tiefenbeck et al. 2018, for example).

A. Appendix

Table A1: Comparison of HER Elements

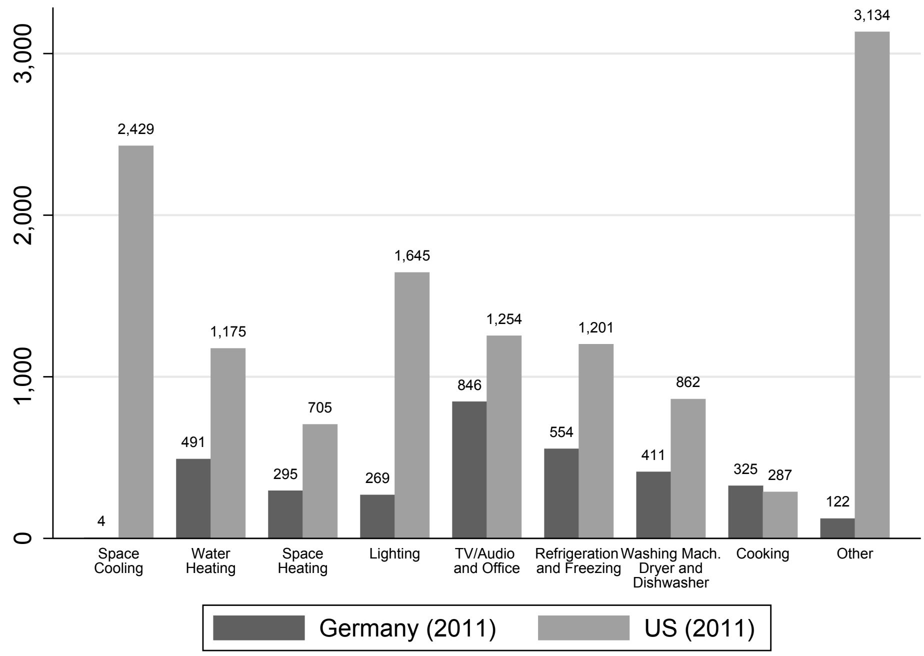
	Our report	*Opower* (Allcott 2011)	Would we expect the differences to increase the effectiveness of our report?
Common Elements			
Social (neighborhood) comparison	x	x	·
Electricity consumption feedback	x	x	·
Electricity saving tips	x	x	·
Diverging Elements			
Possibility to get updated social comparisons and more energy saving tips via an app	x		+ (can trigger continuous engagement with the information from the letters)
Price discounts for energy efficient products in the online shop of the electricity provider	x		+ (Davis et al. 2014)
Frequency of letters	quarterly	monthly - quarterly	- (Allcott 2011)
Update of information (comparison electricity consumption feedback)	yearly	monthly - quarterly	- (Gilbert and Zivin 2014)
Calculation of typical household sizes associated with electricity consumption	x		+ (additional intuitive comparison)
Testimonials (electricity saving actions that other households have implemented)	x		+
Communication of a 10% electricity saving goal within one year	x		+ (Harding and Hsiaw 2014)
Visualization of monthly electricity uses and comparison to last year's consumption		x	-

Table A2: Comparison of ZIP Code Characteristics between the Study Population and German Averages

	Estimation Sample	Germany
Population Density, in Persons per km^2	1,879	224
Percentage of Retirees	23.5%	20.5%
Unemployment Rate	6.8%	6.6%
Purchasing Power per Person, in 1000 EUR	22.0	21.3
Percentage of Foreign Household Heads	7.0%	7.5%

Source: microm (2015).

Figure 4: Composition of Electricity Consumption in Germany and the US



Notes: US data is based on the most recent domestic electricity consumption data from 2011, as documented in EIA (2013). For Germany, we construct the same consumption categories for the year 2011 using data documented in BDEW (2016), UBA (2011) and Destatis (2015).

Table A3: Abatement Cost under the Scenario of a persistent Treatment Effect

	No Persistence	15 Perc. Points Reduction in Effect Size per Year
Country	Abatement Cost in \$ / t CO_2	Abatement Cost in \$ / t CO_2
Canada	67 – 162	17 – 42
United States	20 – 49	5 – 12
France	323 – 779	84 – 202
Germany	75 – 182	20 – 47
Italy	139 – 335	36 – 87
Poland	81 – 196	21 – 51
Spain	121 – 293	32 – 76
Sweden	1,162 – 2,799	302 – 727
United Kingdom	64 – 153	17 – 40
Japan	39 – 94	10 – 24
South Korea	65 – 156	17 – 41
Australia	22 – 53	6 – 14

Notes: Following Allcott and Rogers (2014), the calculations in the table assume linear attenuation rates of 15%. The cost-effectiveness calculations extrapolate electricity reductions until linear decay rates lead to zero reductions. Assumptions about annual electricity uses, carbon intensities of electricity generation and the range of effect sizes are as in Table 1.

B. Welfare Calculations

We follow Allcott and Kessler (2015) and calculate the welfare implications of home energy reports by the following expression:

$$\Delta W = \Delta V - C_n + (\pi_e - \phi_e)\Delta\tilde{e},$$

where:

- ΔW: welfare change per participant induced by the intervention.
- ΔV: private welfare gain from receiving HER. We approximate it by the WTP measure from Allcott and Kessler (2015) that reflects all costs and benefits as perceived by the recipient households. More specifically, households value HER by 54% of the realized electricity cost savings, so that the consumer welfare gain of HER is $\Delta V = \Delta\tilde{e} \cdot 0.54 \cdot p_e$.
- C_n: annual cost of the HER per participant. We assume the cost per quarterly letter to be at 1 US\$ per letter, so that the annual cost are $C_n = 4\$$.

- $\pi_e = p_e - c_c$: mark-up of electricity retail prices p_e over production cost c_c.
- p_e: electricity retail prices that are taken from Eurostat (2016) and the IEA (2016).
- c_c: electricity generation cost. As electricity generation cost depends on the time-of-use of electricity and the time horizon of the analysis, finding a precise measure of generation cost c_c in electricity markets is inherently difficult. We do not want to take a stance on the question of whether levelized generation cost or spot prices are the more suitable electricity generation cost measure and thus present results for both approaches to approximate c_c. First, following Allcott and Kessler (2015) we use levelized generation cost and, second, annual spot price averages. A complete description of all data sources for spot prices and levelized cost is given in the notes to Table A4.
- ϕ_e captures the environmental externalities from electricity generation. We approximate ϕ_e by the social cost of carbon, estimated at around 38 US$ per ton in 2015 (IAWG 2013).
- $\Delta\tilde{e}$ denotes the average treatment effects of HER on electricity consumption (in kWh). To account for the different effect sizes in the *Opower* studies (e.g. Allcott 2011) and in this paper, we use the range of 1.37-3.30%.

As a reference point, Column 6 of Table A4 displays the welfare estimate when we only account for direct cost and climate benefits of HER. This is the welfare analysis that is commonly implemented in classical program evaluations of climate policy interventions and corresponds to our analysis in Section 2. The results show that – under these assumptions – welfare effects of HER are negative in all countries except for the US and Australia, where abatement cost of HER are lower than 38 US$ per ton (Table 1).

Columns 7 and 8 show results for the expanded scope of the welfare analysis as suggested by Allcott and Kessler (2015) for spot prices and levelized cost, respectively. It shows that in a majority of countries HER are not welfare improving, even under optimistic assumptions on effect sizes. Only in South Korea can HER reach positive welfare changes, albeit only when the effectiveness of HER comes close to the upper bound of 3.3%. Using levelized cost in Column 8, the welfare implications of HER improve. Positive welfare effects are possible in Australia, Canada, the US, Japan, and South Korea. In all European countries, welfare effects are still negative, except for Sweden.

Table A4: International Comparison of Welfare Effects of HER Interventions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Country	Average Electricity Consumption in kWh	Retail Price in $ / kWh	Spot Price in $	Levelized Cost of Marginal Plant in $ / kWh	CO_2 Emissions in g per kWh from Electricity Generation	Welfare Change, Considering only Environmental Externalities (as in the Main Text), in $	Welfare Change using Spot Price, in $	Welfare Change using Levelized Cost, in $
Canada	11,379	0.107	0.022	0.071	158	−3.1 – −1.7	−11.9 – −7.3	0.3 – 6.4
United States	12,293	0.127	0.035	0.071	489	−0.9 – 3.5	−5.9 – −4.8	1.2 – 8.5
France	5,859	0.186	0.043	0.101	64	−3.8 – −3.5	−11.8 – −7.2	−2.5 – −0.5
Germany	3,304	0.327	0.035	0.106	486	−3.2 – −2.0	−14.6 – −8.4	−6.8 – −5.2
Italy	2,542	0.269	0.058	0.099	343	−3.5 – −2.9	−8.5 – −5.9	−5.0 – −4.4
Poland	1,935	0.157	0.040	0.083	769	−3.2 – −2.1	−4.2 – −4.1	−2.9 – −1.4
Spain	4,040	0.263	0.056	0.083	247	−3.5 – −2.7	−11.4 – −7.1	−7.8 – −5.6
Sweden	8,025	0.208	0.024	0.114	13	−3.9 – −3.9	−22.8 – −11.8	−1.9 – 0.9
United Kingdom	4,145	0.242	0.062	0.108	459	−3.0 – −1.6	−8.4 – −5.8	−3.2 – −2.1
Japan	5,434	0.225	0.081	0.143	572	−2.4 – −0.1	−4.2 – −4.1	0.6 – 7.0
South Korea	3,489	0.103	0.090	0.130	536	−3.0 – −1.7	−1.0 – 3.3	0.9 – 7.9
Australia	6,959	0.216	0.082	0.108	798	−1.1 – 3.0	−2.7 – −0.9	−0.3 – 5.0

Notes: Our calculations assume printing and mailing cost of one dollar per report, four reports per year and average electricity reductions of 1.37-3.30%. Average electricity consumption and CO_2 intensities of electricity generation correspond to 2013, the most recently available year documented in WEC (2016) and IEA (2015), respectively.

Retail electricity prices for 2015 are drawn from Eurostat (2016) for EU-countries and from IEA (2016) for the remaining countries. Spot prices for electricity are annual spot price averages from the following sources: USA (unweighted mean of 8 price hubs from EIA (2017)); France, Germany and UK (EPEX 2016); Italy (GME 2017); Poland (PSE 2017); Spain (OMIE 2016); Sweden and Norway (unweighted mean of regional prices from Nord Pool (2017)); Japan (JEPX 2017); South Korea (KPX 2017); Australia (unweighted mean of regional prices from AEMC (2016)); Canada (unweighted mean of the Hourly Ontario Energy Price (Class B for residential Consumers) IESO (2015), and the Alberta Average Pool Price, AESO (2016)).

The levelized cost of energy (LCOE) are taken from OECD (2015). We assume combined cycle gas turbine (CCGT) as marginal power plant in the respective country, as for example Allcott and Rogers (2014) and Allcott and Mullainathan (2010). Note that combined cycle gas turbine (CCGT) tend to have larger production cost for electricity so that we tend to overestimate the welfare impacts of HER. Furthermore, we take the largest available production cost estimates, assuming a discount rate of 10%, which increases welfare further. Only in Poland, where natural gas power plants are virtually absent and coal power plants are predominant, we employ levelized cost of a hard coal plant (as we lack LCOE data for Poland, we approximate it by the LCOE of a hard coal power plant in Germany OECD (2015), where fuel cost for hard coal are very similar). In cases where estimates for 2015 are available, we use the latest available estimate from OECD (2015): Sweden (value from 2010), Spain (value from 1998) and Italy (value from 2010). As LCOE are unavailable for Australia in OECD (2015), we take the upper bound of the range provided in WEC (2013)), and approximate the LCOE in Canada by the LCOE in the US, which has been very similar in the past OECD (2015). Where applicable, we use average exchange rates for 2015, as provided by FED (2016). The 2015 average annual exchange rate for Poland (PLN) is provided by NBP (2017).

C. Online Appendix

The online appendix can be found at: https://www.rwi-essen.de/media/content/pages/publikationen/ruhr-economic-papers/rep_714_online_appendix.pdf

References

Abrahamse, W., Steg, L., 2013. Social Influence Approaches to Encourage Resource Conservation: A Meta-Analysis. Global Environmental Change 23 (6), 1773–1785.

AEMC, 2016. Residential Electricity Price Trends, Final Report. Australian Energy Market Commission. URL: `http://www.aemc.gov.au/getattachment/be91ba47-45df-48ee-9dde-e67d68d2e4d4/2016-Electricity-Price-Trends-Report.aspx` (accessed February 23, 2017).

AESO, 2016. AESO 2015 Annual Market Statistics. Alberta Electric System Operator. URL: `https://www.aeso.ca/assets/listedfiles/2015-Annual-Market-Stats-WEB.pdf` (accessed February 24, 2017).

Al-Ubaydli, O., List, J. A., LoRe, D., Suskind, D., 2017a. Scaling for Economists: Lessons from the Non-Adherence Problem in the Medical Literature. Journal of Economic Perspectives 31 (4), 125–44.

Al-Ubaydli, O., List, J. A., Suskind, D. L., 2017b. What Can We Learn from Experiments? Understanding the Threats to the Scalability of Experimental Results. American Economic Review 107 (5), 282–86.

Allcott, H., 2011. Social Norms and Energy Conservation. Journal of Public Economics 95 (9), 1082–1095.

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

Allcott, H., 2016. Paternalism and Energy Efficiency: An Overview. Annual Review of Economics 8, 145–176.

Allcott, H., Kessler, J. B., 2015. The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons. National Bureau of Economic Research (NBER) Working Paper No. 21671 (October).

Allcott, H., Mullainathan, S., 2010. Behavior and Energy Policy. Science 327 (5970), 1204–1205.

Allcott, H., Rogers, T., 2014. The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. American Economic Review 104 (10), 3003–3037.

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, 178–210.

BDEW, 2016. Energie-Info: Stromverbrauch im Haushalt. Bundesverband der Energie- und Wasserwirtschaft. URL: `https://www.bdew.de/internet.nsf/id/0D6D0C20786B2929C1257FE90030A849/$file/BDEW_Stromverbrauch%20im%20Haushalt_Stand_Juli%202016.pdf` (accessed October 28, 2016).

Brandon, A., Ferraro, P. J., List, J. A., Metcalfe, R. D., Price, M. K., Rundhammer, F., 2017. Do the Effects of Social Nudges Persist? Theory and Evidence from 38 Natural Field Experiments. National Bureau of Economic Research (NBER) Working Paper No. 23277 (March).

Brent, D. A., Cook, J. H., Olsen, S., 2015. Social Comparisons, Household Water Use, and Participation in Utility Conservation Programs: Evidence from Three Randomized Trials. Journal of the Association of Environmental and Resource Economists 2 (4), 597–627.

Davis, J., Guryan, J., Hallberg, K., Ludwig, J., 2017. The Economics of Scale-Up. National Bureau of Economic Research (NBER) Working Paper No. 23925 (October).

Davis, L. W., Fuchs, A., Gertler, P., 2014. Cash for Coolers: Evaluating a Large-Scale Appliance Replacement Program in Mexico. American Economic Journal: Economic Policy 6 (4), 207–238.

Deaton, A., Cartwright, N., 2016. Understanding and Misunderstanding Randomized Controlled Trials. National Bureau of Economic Research (NBER) Working Paper No. 22595 (September).

Dehejia, R., Pop-Eleches, C., Samii, C., 2015. From Local to Global: External Validity in a Fertility Natural Experiment. National Bureau of Economic Research (NBER) Working Paper No. 21459 (August).

Destatis, 2015. Umweltnutzung und Wirtschaft. Tabellen zu den Umweltökonomischen Gesamtrechnungen. URL: https://www.destatis.de/DE/Publikationen/Thematisch/UmweltoekonomischeGesamtrechnungen/Querschnitt/UmweltnutzungundWirtschaftTabelle5850007157006Teil_2.pdf?__blob=publicationFile (accessed October 28, 2016).

Dietz, T., Gardner, G. T., Gilligan, J., Stern, P. C., Vandenbergh, M. P., 2009. Household Actions Can Provide a Behavioral Wedge to Rapidly Reduce US Carbon Emissions. Proceedings of the National Academy of Sciences 106 (44), 18452–18456.

EIA, 2013. Annual Energy Outlook 2013. U.S. Energy Information Administration. URL: http://www.eia.gov/outlooks/aeo/pdf/0383(2013).pdf (accessed October 28, 2016).

EIA, 2017. Wholesale Electricity and Natural Gas Market Data. U.S. Energy Information Administration. URL: https://www.eia.gov/electricity/wholesale/history (accessed March 2, 2017).

EPEX, 2016. EPEX SPOT Reaches in 2015 the Highest Spot Power Exchange Volume Ever. European Power Exchange - Press Release. URL: https://www.epexspot.com/document/34250/2016-01-08_EPEX%20SPOT_2015_Annual%20Press%20Release.pdf (accessed March 2, 2017).

Eurostat, 2016. Electricity Price Statistics. URL: http://ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_price_statistics (accessed February 23, 2017).

FED, 2016. Foreign Exchange Rates - G.5A. Board of Governors of the Federal Reserve System. Release Date: January 4, 2016. URL: https://www.federalreserve.gov/releases/g5a/20160104/ (accessed February 24, 2017).

Ferraro, P. J., Miranda, J. J., Price, M. K., 2011. The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment. The American Economic Review 101 (3), 318–322.

Ferraro, P. J., Price, M. K., 2013. Using Non-Pecuniary Strategies to Influence Behavior: Evidence from a Large Scale Field Experiment. Review of Economics and Statistics 95 (1), 64–73.

Frey, B. S., Meier, S., 2004. Social Comparisons and Pro-social Behavior: Testing "Conditional Cooperation" in a Field Experiment. American Economic Review 94 (5), 1717–1722.

Gechter, M., 2016. Generalizing the Results from Social Experiments: Theory and Evidence from Mexico and India. URL: http://www.personal.psu.edu/mdg5396/Gechter_Generalizing_Social_Experiments.pdf (accessed June 6, 2017).

Gilbert, B., Zivin, J. S. G., 2014. Dynamic salience with intermittent billing: Evidence from smart electricity meters. Journal of Economic Behavior & Organization 107, 176–190.

GME, 2017. Summary Data - MPE-MGP Overview . Gestore del Mercato Elettrico. URL: http://www.mercatoelettrico.org/En/Statistiche/ME/DatiSintesi.aspx (accessed March 2, 2017).

Goulder, L. H., 2013. Markets for Pollution Allowances: What Are the (New) Lessons? The Journal of Economic Perspectives 27 (1), 87–102.

Greenstone, M., Kopits, E., Wolverton, A., 2013. Developing a Social Cost of Carbon for US Regulatory Analysis: A Methodology and Interpretation. Review of Environmental Economics and Policy 7 (1), 23–46.

Harding, M., Hsiaw, A., 2014. Goal Setting and Energy Conservation. Journal of Economic Behavior & Organization 107, 209–227.

Holland, S. P., Mansur, E. T., Muller, N. Z., Yates, A. J., 2016a. Are There Environmental Benefits from Driving Electric Vehicles? The Importance of Local Factors. American Economic Review 106 (12), 3700–3729.

Holland, S. P., Mansur, E. T., Muller, N. Z., Yates, A. J., 2016b. Distributional Effects of Air Pollution from Electric Vehicle Adoption. National Bureau of Economic Research (NBER) Working Paper No. 22862 (November).

Hotz, J. V., Imbens, G. W., Mortimer, J. H., 2005. Predicting the Efficacy of Future Training Programs Using Past Experiences at Other Locations. Journal of Econometrics 125 (1-2), 241–270.

IAWG, 2013. Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis. U.S. Interagency Working Group on the Social Cost of Carbon for Regulatory Impact Analysis. URL: https://www.whitehouse.gov/sites/default/files/omb/inforeg/social_cost_of_carbon_for_ria_2013_update.pdf (accessed October 28, 2016).

IEA, 2015. CO2 Emissions from Fuel Combustion - Highlights. International Energy Agency. URL: https://www.iea.org/publications/freepublications/publication/CO2EmissionsFromFuelCombustionHighlights2015.pdf (accessed October 28, 2016).

IEA, 2016. Energy Prices and Taxes - Fourth Quarter 2016. International Energy Agency. URL: http://www.oecd.org/publications/energy-prices-and-taxes-16096835.htm (accessed February 23, 2017).

IESO, 2015. Ontario Energy Report Q4 2015. Independent Electricity System Operator. URL: https://www.ontarioenergyreport.ca/pdfs/5806_IESO_OntarioEnergyReportQ42016_Electricity_EN_FA.pdf (accessed February 23, 2017).

Jaime Torres, M. M., Carlsson, F., 2016. Social Norms and Information Diffusion in Water-saving Programs: Evidence from a Randomized Field Experiment in Colombia. EfD Discussion Paper Series 16-05.

JEPX, 2017. Spot Market Indices - Fiscal Year 2015. Japan Electric Power Exchange. URL: http://www.jepx.org/market/excel/index_2015.csv (accessed February 24, 2017).

KPX, 2017. System Marginal Price 2015. Korea Power Exchange. URL: http://epsis.kpx.or.kr/epsis/ekmaStaticMain.do?cmd=004013&flag=&locale=EN (accessed February 5, 2017).

Leviton, L. C., 2017. Generalizing about Public Health Interventions: A Mixed-Methods Approach to External Validity. Annual Review of Public Health 38, 371–91.

Levitt, S. D., List, J. A., 2009. Field Experiments in Economics: The Past, the Present, and the Future. European Economic Review 53 (1), 1–18.

microm, 2015. Socioeconomic Regional Data. Version: 1. RWI – Rheinisch–Westfälisches Institut für Wirtschaftsforschung e.V. DOI: 10.7807/microm:einwGeAl:V3,10.7807/microm:alq:V3,10.7807/microm:kaufkraft:V3.

Muller, S. M., 2015. Causal Interaction and External Validity: Obstacles to the Policy Relevance of Randomized Evaluations. The World Bank Economic Review 29 (suppl 1), S217–S225.

NBP, 2017. Statistics - Exchange Rates. Narodowy Bank Polski. URL: http://www.nbp.pl/homen.aspx?f=/kursy/kursyen.htm (accessed March 3, 2017).

Nord Pool, 2017. Elspot Prices. Nord Pool. URL: http://www.nordpoolspot.com/Market-data1/Elspot/Area-Prices/ (accessed March 2, 2017).

Nordhaus, W., 2014. Estimates of the Social Cost of Carbon: Concepts and Results from the DICE-2013R Model and Alternative Approaches. Journal of the Association of Environmental and Resource Economists 1 (1/2), 273–312.

OECD, 2015. Projected Costs of Generating Electricity - 2015 Edition. Organisation for Economic Co-operation and Development. URL: https://www.iea.org/publications/freepublications/publication/projected-costs-of-generating-electricity-2015-edition.html (accessed February 23, 2017).

OMIE, 2016. Price Report. Operador do Mercado Ibrico de Energia. URL: http://www.omie.es/files/informe_precios_ing_navegable.pdf (accessed March 2, 2017).

Perino, G., 2015. Climate Campaigns, Cap and Trade, and Carbon Leakage: Why Trying to Reduce Your Carbon Footprint Can Harm the Climate. Journal of the Association of Environmental and Resource Economists 2 (3), 469–495.

Peters, J., Langbein, J., Roberts, G., 2018. Generalization in the Tropics–Development Policy, Randomized Controlled Trials, and External Validity. The World Bank Research Observer 33 (1), 34–64.

Pritchett, L., Sandefur, J., et al., 2015. Learning from Experiments when Context Matters. American Economic Review 105 (5), 471–75.

PSE, 2017. Electricity Prices and Volumes on the Balancing Market. Polskie Sieci Elektroenergetyczne. URL: http://www.pse.pl/index.php?dzid=129&did=&lang_id=2 (accessed February 24, 2017).

Shang, J., Croson, R., 2009. A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods. The Economic Journal 119 (540), 1422–1439.

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., Staake, T., 2018. Overcoming Salience Bias: How Real-Time Feedback Fosters Resource Conservation. Management Science 64 (3), 1458–1476.

UBA, 2011. Klimaschutz durch Reduzierung des Energiebedarfs für Gebäudekühlung. Umweltbundesamt. URL: https://www.umweltbundesamt.de/sites/default/files/medien/461/publikationen/3979.pdf (accessed October 28, 2016).

UBA, 2018. Entwicklung der spezifischen Kohlendioxid-Emissionen des deutschen Strommix in den Jahren 1990 2017. Climate Change 11.

Vivalt, E., 2015. Heterogeneous Treatment Effects in Impact Evaluation. American Economic Review: Papers & Proceedings 105 (5), 467–70.

WEC, 2013. World Energy Perspective - Cost of Energy Technologies. World Energy Council. URL: https://www.worldenergy.org/wp-content/uploads/2013/09/WEC_J1143_CostofTECHNOLOGIES_021013_WEB_Final.pdf (accessed February 23, 2017).

WEC, 2016. Energy Efficiency Indicators via Enerdata, Average Electricity Consumption per Electrified Household. World Energy Council. URL: https://www.wec-indicators.enerdata.eu/secteur.php/household-electricity-use.html (accessed October 28, 2016).

Wichman, C. J., 2017. Information Provision and Consumer Behavior: A Natural Experiment in Billing Frequency. Journal of Public Economics 152, 13–33.

Zivin, J. S. G., Kotchen, M. J., Mansur, E. T., 2014. Spatial and Temporal Heterogeneity of Marginal Emissions: Implications for Electric Cars and Other Electricity-Shifting Policies. Journal of Economic Behavior & Organization 107, 248–268.