Do homo sapiens know their prices? Insights on dysfunctional price mechanisms from a large field experiment

Working Paper

Author(s):
Ritter, Nolan; Bingler, Julia Anna

Publication date:
2021-03

Permanent link:
https://doi.org/10.3929/ethz-b-000477129

Originally published in:
Economics Working Paper Series 21/348

CER-ETH – Center of Economic Research at ETH Zurich

Do homo sapiens know their prices? Insights on dysfunctional price mechanisms from a large field experiment

N. Ritter, J. A. Bingler

Working Paper 21/348
March 2021

Economics Working Paper Series

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Do homo sapiens know their prices? Insights on dysfunctional price mechanisms from a large field experiment

Nolan Ritter[1,2] and Julia Anna Bingler[3]

[1]Mercator Research Institute on Global Commons and Climate Change (MCC), Torgauer Strasse 12 - 15, 10829 Berlin, Germany, ritter@mcc-berlin.net.

[2]Potsdam Institute for Climate Impact Research (PIK), Telegraphenberg A 31, 14473 Potsdam, Germany.

[3]ETH Zurich, Center of Economic Research, Zürichbergstrasse 18, 8092 Zurich, Switzerland, binglerj@ethz.ch.

March 22, 2021

Abstract

We use a large, randomized field experiment that exogenously varies prices to test their effect on consumption. Full information is available at zero costs. However, households state prices that are, on average, ten times larger than actual. But ignorant households cannot react to prices and so the price mechanism becomes dysfunctional. Our results explain small or zero price effects from previous research. We show that researchers must provide evidence for a functional price mechanism before ascribing causal effects or risk biased conclusions. The same applies to price instruments that are often regarded as first best solutions.

Keywords: households; information; price experiment;

JEL Classification: D12, D83, L95, Q41, Q54

Acknowledgements: We would like to thank the many individuals who commented and discussed previous versions of this paper, especially William Greene (NYU Stern), Kenneth Gillingham (Yale University), Laura Grant (Claremont McKenna College), Anna Alberini (University of Maryland), and Massimo Filippini (ETH Zurich) in the context of the Empirical Methods in Energy and Environmental Economics seminar and the 2020 EMEE conference at ETH, Lucas Bretschger (ETH Zurich) and the colleagues at the Chair of Economics/Resource Economics at CER-ETH Zurich, and the participants at the EAERE 2020 and the SURED 2020 conference. All errors are our own.

1 Introduction

In economics, the assumption of functional price mechanisms is ubiquitous. Individuals make optimal decisions when they possess factually correct and full information and optimize accordingly. However, individuals may neither always possess accurate information nor always optimize as assumed. For example, the benefits of fully informed decisions have to be weighed against the costs of gathering information. This may lead to widely recognized phenomena such as rational ignorance (Downs, 1957) or the use of heuristics in decision making (Simon, 1957, 1979; Kahneman and Tversky, 1979; Thaler, 1980; Shiller, 1981; De Bondt and Thaler, 1985; Thaler, 1985; Kahneman et al., 1991; Kahneman, 2003). One assumption underlying the price mechanism is that individuals know prices and choose quantities accordingly. But a sheer endless selection of available goods and services raises the question whether individuals really do.

In this paper, we provide strong evidence that the assumptions underlying the price mechanism do not always hold. By explicitly testing the price mechanism, we are able to explain findings and interpretations of small or zero price effects from previous research using a framework that could inform further research designs and the interpretation of empirical estimates. Our analysis shows that researchers need to provide evidence for a functional price mechanism when estimating price elasticities and before ascribing causal price effects or risk severely biased conclusions. Testing the price mechanism enables the detection of false positives. In addition, it helps to differentiate between analyses lacking power and analyses that do not find effects when there are none.

To our knowledge, this is the first paper that analyses the effect of a randomly administered exogenous price treatment and comprehensively tests the price mechanism underlying the observed effect using the same experimental data. Our paper is inspired by a host of excellent research papers that put great emphasis on careful research design (for example Manski, 1993; Angrist and Pischke, 2010; Deaton, 2010) and by Ludwig et al. (2011) and Lopez de Leon and Rizzi (2014) who advocate for testing causal mechanisms. The basis of our analysis is the Irish Gas Customer Behavioural Trial which has internal and external validity, in which attrition occurs at random, and which has a comprehensive post-treatment questionnaire that allows us to test the fundamental assumptions underlying the price mechanism. While the trials administer a selection of exogenous treatments to participating households, we exclusively focus on the groups that receive price experiment.

There are many advantages for using the Irish Gas Customer Behaviour Trial data. As Davis and Kilian (2011) point out, natural gas is a homogenous good so that we neither have to accommodate for potential differences in product quality nor for product differentiation. Davis and Kilian further point out that there is no secondary market for natural gas. Households can neither store natural gas for later consumption nor for resale. Further, we argue that because natural gas is piped to households, households' purchase of natural gas is not associated with additional private costs for transportation such as fuel, or time costs. In addition, the Irish gas market was a monopoly at the time of the experiment. The monopolist set a single annual natural gas price for all Irish households including households in the control group. This implies that households in the control group only had to know one single price to achieve full price information. Households in the treatment group needed to learn only one single new price every two months to be fully informed. Moreover, the experimenters equipped every control and every treatment households with in-home displays that revealed past and present consumption levels, costs, and the price at the press of a button. It is hard to imagine a setting in which full information is more easily achievable. Nearly every other market is characterized by a vast amount of product varieties, different and constantly varying prices at a similarly vast number of different geographic sales points each associated with individual time and transportation costs. Additionally, with an approximate annual consumption of about $11,000$ Kilowatt hours (kWh), the average household spends about 430 Euros annually on natural gas. This implies that the cost of natural gas is non-negligible. The Gas Customer Behaviour Trials are also remarkable because of their thorough post-trial survey that allows us to comprehensively test the assumptions underlying the price mechanism. Overall, we demonstrate that the field experiment has internal and external validity. We show that attrition occurs at random. Further, we provide evidence that individuals who answered the post-survey questionnaire on which our tests of the price mechanism rely do not differ in their observable characteristics from those who did not.

We provide overwhelming evidence that neither treated nor control households know the actual price, although this information is available from the in-home device at the press of a button. When asked, households indicate, on average, prices in excess of ten fold the actual ones. We calculate a probability of 1 in 4 billion that actual and stated prices are identical. Thus, we are able to undoubtedly show that households are ignorant of the price in our case. Without price information, households cannot react to it. This implies that the price mechanism is dysfunctional and price treatments cannot have any effect on consumption. Further tests provide compelling evidence that households exert limited control over the quantity of natural gas they consume and that they have false beliefs regarding the reasons as to why they adjust their natural gas consumption. Overall, these findings are at odds with the assumptions underlying the price mechanism, i.e. that rational households should be aware as to why they adjust their consumption. Unsurprisingly, when testing the effect of the price treatment on household consumption levels, our frequentist approach finds insignificant treatment coefficients close to zero with small standard errors across all model specifications. Additionally, we calculate Bayes factors that indicate that the null hypothesis is up to 24.9 times more likely than the alternative hypothesis.

Our findings have far reaching implications for a number of different areas. First, we contribute to the literature that aims at improving the credibility of empirical economics (for an overview see Angrist and Pischke, 2010). We fully agree that much progress has been made since Leamer (1983) demanded to take the "con" out of econometrics. A good research design clearly is pivotal for credible research. However, our results clearly show that testing the hypotheses underlying the assumed causal mechanism is equally important, yet hardly ever done. Therefore, an unknown number of published articles might provide false positives because the causal mechanism underlying the research hypotheses' claims remained untested. Similarly, we expect that a large number of insignificant findings were disregarded falsely assuming a lack of statistical power. We argue that researchers who go to great lengths to avoid bias in general (Allcott, 2015; Oster, 2019) and endogeneity in particular (see, i. e. Hahn and Hausman, 2003; Blundell and Powell, 2004; Goldsmith-Pinkham et al., 2020; Harari, 2020), must also provide evidence for a functioning mechanism underlying their hypotheses before ascribing causal effects or risk bias. More specifically, for the present case, evidence for a functioning price mechanism would be the necessary condition to identify causal price effects; while significant coefficients on a price treatment are the sufficient condition. The validity of an assumed causal mechanism cannot be inferred from finding the expected regression coefficients. Instead, it is important to formally test. Had we found a statistically significant price effect, the lack of price knowledge would have indicated to us that our estimated price effect was a false positive. Testing the price mechanism is therefore also a fail-safe that provides us from drawing biased conclusions.

Second, our paper contributes to the literature on price salience (Hossain and Morgan, 2006; Finkelstein, 2009; Chetty et al., 2009; Bundorf et al., 2012; Jessoe and Rapson, 2014a; Kőszegi and Matějka, 2020; Fabra and Reguant, 2020). For many piped goods such as electricity or natural gas, price knowledge is usually less well developed compared to other products although they are homogeneous and home delivered (Jessoe and Rapson, 2014a; Buchanan et al., 2015; Alberini, 2018; Blasch et al., 2019). This is reflected in very low price elasticities (Allcott, 2011; Ito, 2014). To explain these findings, the literature suggests that inattention and context-specific preferences matter. For example, Davis and Metcalf (2016) conduct an experiment to show that specific information leads to better outcomes than generic information. Fowlie et al. (2017) conduct a randomized controlled trial in the residential electricity market to analyze whether consumption decisions are based on heuristics. They find overwhelming evidence for a default effect. When confronted by a choice with a default decision, a large majority of individuals accepts the default. This heuristic simplifies complex choices. A recent paper by Kőszegi and Matějka (2020) explores why individuals simplify when making multi product decisions and the amount of attention that can be given to prices is limited and costly. In a setting in which households only have to know a single price set by a monopolist for an entire year for a homogenous good that is delivered to their homes via pipe whose price can be obtained at the press of a button on an in-home device, we demonstrate that households nevertheless do not know the decision-relevant price. A limited amount of attention may explain why households in our experiment are ignorant of the price and indicate prices that are, on average, ten times the actual one. Our finding clarifies that households may simplify to the point of ignorance because their attention is fully allocated to dealing with problems perceived as more important.

Third, our paper contributes to the literature on dynamic pricing in general and on optimal natural gas consumption in particular. Retail prices for natural gas are still mostly fixed and do not reflect varying levels of scarcity revealed in the wholesale market. Davis and Kilian (2011) show circumstances in which a price ceiling for natural gas prevents the allocation of natural gas to households who value it most in the U.S. residential market. In another study, Davis and Muehlegger (2010) find that natural gas prices are considerably above their marginal level which leads them to conclude that households consume too little natural gas. Jessoe and Rapson (2014a) conclude that the combination of dynamic pricing and improvements in metering technology (Borenstein, 2002, 2005; Joskow and Wolfram, 2012) could improve efficiency in the short- and in the long-run. In the absence of evidence for dynamic pricing for natural gas, we might draw on the literature concerning dynamic electricity prices. Holland and Mansur (2006) simulate the PJM electricity market and find that real-time prices improve efficiency. Allcott and Rogers (2014) provide another important study for the electricity market which focuses on short- and long-term effects of behavioral interventions. Using a synthetic control method, Bretschger and Grieg (2020) corroborate that prices have limited effect on driving, too. Our findings may explain why some of the previous mentioned analyses find that prices have little to no effect.

Last, our findings highlight that without testing causal mechanisms researchers run the risk of drawing false conclusions or remain ignorant of highly heterogeneous effects caused by a price mechanism that is dysfunctional in a significant part of the population. While in our case hardly anyone indicated a price close to the actual one, there certainly are cases in which a significant share of individuals know the price. Then, a mean estimate overstates the effect a price based instrument has on the uninformed while it underestimates the effect for those who are informed. While on average such a price instrument achieves its intended goal, the instrument's distributional effects are most certainly different than anticipated. Examples of heterogenous levels of information abound in the literature. Better informed agents choose different retirement plans (Duflo and Saez, 2003), apply more often for social benefits (Bhargava and Manoli, 2015), and make more efficient decisions (Jensen, 2007; Grubb and Osborne, 2015). Fabra and Reguant (2020) analyze consequences if buyers differ with respect to their willingness to search for information. Reiss and White (2005) analyze household reactions to price increases in electricity and find that most of the corresponding change in aggregated demand is caused by a small fraction of informed households. The findings from these studies underline how important it is to test causal mechanisms.

This is the outline of our paper. The following section describes the experimental design. In Section 3, we test the price mechanism, present descriptive statistics, conduct tests for internal and external validity, and present and discuss the treatment effects. We also compare our findings to previous studies using the same data set. Section 4 summarizes and concludes.

2 Experimental design

Between December 2009 and May 2011, the Irish Commission for Energy Regulation (CER), together with the Irish main gas supplier Bord Gáis Energy, conducted the *Gas Customer Behavioural Trial* (CBT). To test the influence of information and price treatments on natural gas consumption, a sample of $1,892$ households was created. The trial had three phases. First, smart-meters for monitoring natural gas consumption were installed in all participating households. Second, between December 1, 2009 and May 31, 2010, the experimenters verified the accuracy of the smart-meters and the reliability of data transmission. During this period, no treatments were administered. Third, between June 1, 2010 and May 31, 2011, households were randomly allocated to control or treatment groups. Socio-economic and dwelling characteristics were surveyed in the pre-trial period (CER, 2011b, p.35). To incentivize participation, households received 25 Euros upon completing the survey (CER, 2011b, p. 60).

2.1 Treatment details

Households were randomly allocated into 5 groups. Although we exclusively focus on the effect of exogenous price variations on natural gas consumption, we describe the trial setup and analyze its internal validity for all treatments because it strengthens our analysis of the internal and the external validity. Group 0 continued to receive bi-monthly bills as is customary in Ireland and was not subjected to any kind of treatment. Table 1 provides an overview of the characteristics of the different groups.

Households allocated to Group 1 received bi-monthly bills like the Group 0, but were provided additional information in the form of a written energy usage statement (EUS). This statement included a comparison of the actual consumption with that of the same billing period in the previous year and a comparison to the average consumption of natural gas consuming households. Households were also informed at what time of the day most of their natural gas consumption occurred. The experimenters hypothesized that the provision of information via the energy usage statement reduces con-

Table 1: *Treatment groups in the final data set (N=1,275)*

group	N	bi-monthly bill	EUS	IHD	price variation
0	454	yes	no	no	no
1	199	yes	yes	no	no
2	196	no	yes	no	no
3	211	yes	yes	yes	no
4	215	yes	yes	yes	yes

This table provides an overview of the individual treatments different groups receive in the field experiment. Abbreviations: N = number of participating households. Measurements indicates the total number of metering observations. EUS = energy usage statement with the bill, IHD = in-home device. Group names: 0 = control group, 1 = bi-monthly bill, 2 = monthly bill, 3 = IHD 4 = price variation.

sumption, especially in households with higher than average consumption levels. With the energy usage statement being the only systematic difference to Group 0, a comparison between Group 0 and Group 1 reveals the effect of the energy usage statement.

While households in Ireland receive bills on a bi-monthly basis, households in Group 2 were billed on a monthly basis. Similar to Group 1, households in Group 2 were also provided with an energy usage statement. The only systematic difference between Group 1 and Group 2 was the frequency with which households receive bills for their natural gas consumption. Thus, a comparison between Group 1 and Group 2 indicates the effect of variations in the billing frequency.

Group 3 received bi-monthly bills, an energy usage statement, and an in-home device (IHD). An in-home device displays information collected by the smart-meter. Although there is a smart-meter installed in every household to monitor natural gas consumption, only households with an in-home device (IHD) have access to the collected information. IHDs make it possible to display natural gas consumption in real time with respect to consumption levels, the natural gas tariff, and the cost of consumption. Because information on prices and costs is readily available compared to monthly or bi-monthly billing, the experimenters expected households' consumption levels to differ. The only systematic difference between Group 1 and Group 3 is the in-home device, hence comparing the consumption levels of both groups returns the effect of the in-home device. A comparison between Group 3 and Group 0 returns the combined effect of an energy usage statement and an in-home device.

Similar to Group 3, Group 4 was given an energy usage statement, bi-monthly billing, and an in-home device. In addition, it was also subjected to a varying natural gas price. Table 2 indicates the tariffs that the groups were subjected to. In February and March of 2011, both groups were charged the same price for natural gas. Thus, the experiment includes a placebo treatment for which we should find no significant treatment effects on natural gas consumption if randomization was overall successful. In June and July of 2010, households in Group 4 paid 3.3 Euro cents per kilowatt hour of natural gas, which is about 15.4% less than households in Group 3. Group 3 was charged a flat tariff of 3.9 Euro cents per Kilowatt hour similar to all other households in Ireland including the other treatment groups. In December and January, when the weather was coldest and natural gas consumption was highest, households in Group 4 paid a 17.9% higher tariff compared to households in Group 3. Hence, a comparison between households in Group 3 and Group 4 indicates the effect of price changes on consumption. Altogether, the price treatments are of considerable magnitude with variations between −15.4% and +17.9%.

Table 2: *Natural gas prices by month and treatment group (CER, 2011b, p.6)*

Tariff groups	**Months**					
	Jun/Jul	Aug/Sep	Oct/Nov	Dec/Jan	Feb/Mar	Apr/May
I: 0, 1, 2, 3	3.9	3.9	3.9	3.9	3.9	3.9
II: 4	3.3	3.3	3.8	4.6	3.9	3.4
(II - I)	−0.6	−0.6	−0.1	0.7	0.0	−0.5
(II - I) / I	−15.4%	−15.4%	−2.6%	17.9%	0.0%	−12.8%

This table indicates the prices that households in different treatment groups pay for natural gas at different points in time. All prices are in Euro cents per kWh. Price differences between the groups are shown in absolute (II-I) and relative (II-I)/I terms. Group names: 0 = control group, 1 = bi-monthly bill, 2 = monthly bill, 3 = IHD, 4 = price variation. Months June through December observed in 2010. Months January through May observed in 2011.

Table 3 provides an overview of how the exogenous price variations are hypothesized to influence consumption under a functional price mechanism. The first column assigns a number for each hypothesis tested for ease of reference. Columns 2 and 3 indicate which groups we compare. The fourth column indicates that each hypothesis is tested using 426 observations which is a comparatively large number of observations for field experiments.

Table 3: *Hypotheses regarding information and price treatments*

No.	groups control	groups treatment	N	period	hypothesized effect on consumption
1	3	4	426	June 2010	$\beta > 0$
2	3	4	426	July 2010	$\beta > 0$
3	3	4	426	August 2010	$\beta > 0$
4	3	4	426	September 2010	$\beta > 0$
5	3	4	426	October 2010	$\beta > 0$
6	3	4	426	November 2010	$\beta > 0$
7	3	4	426	December 2010	$\beta < 0$
8	3	4	426	January 2011	$\beta < 0$
9	3	4	426	February 2011	$\beta = 0$
10	3	4	426	March 2011	$\beta = 0$
11	3	4	426	April 2011	$\beta > 0$
12	3	4	426	Mai 2011	$\beta > 0$

This table indicates the hypotheses we test. Group 3 receives no price treatment and is our control group. Group 4 receives price treatment and is our treatment group. No. indicates the number of the hypothesis. N is the number of observations. $\beta > 0$ indicates an expected increase in consumption when the treatment group (Group 4) pays a lower price than the control group (Group 3).

The two ultimate columns indicate the treatment under scrutiny and the hypothesis regarding the impact of the treatment on consumption. For example, $\beta > 0$ indicates that the exogenous price reduction is hypothesized to increase natural gas consumption. When the price in the treatment group (Group 4) exceeds that of the control group (Group 3) in December and January, a coefficient of $\beta < 0$ indicates the expectation that consumption is lower in treated than in control households.

3 Empirical analysis

Because our focus is on dysfunctional prices, we begin our empirical analysis scrutinizing the price mechanism first. After we present overwhelming evidence that it is dysfunctional (Section 3.1), we continue with updating our hypothesized effects (Section 3.2). Afterwards, we present descriptive statistics (Section 3.3), show that pre-trial consumption was identical across treated and control households (Section 3.4), provide evidence that randomization was successful (Section 3.5), and that attrition occurs at random

(Section 3.6). We scrutinize the external validity in Section 3.7 before presenting our estimates of the treatment effects in Section 3.8. We show that our results hold irrespective of whether we apply frequentist or Bayesian statistics (Section 3.9). We carry out an heterogeneity analysis in Section 3.10. We compare our findings with previous analyses of the same data (Section 3.11) before concluding in Section 4.

3.1 Mechanism

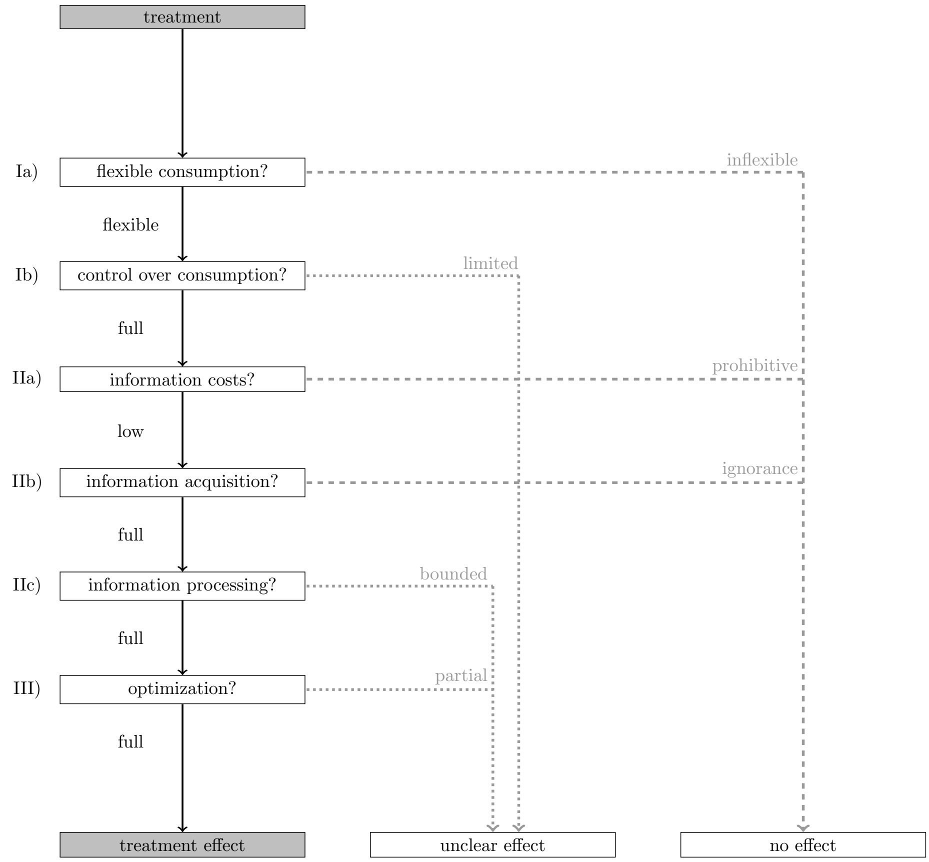
Table 3 summarizes the expected effects of exogenous price variations if this assumption also holds in the field experiment at hand. In the following, we test the price mechanism comprehensively to clarify what expectations we should rationally harbor regarding the price treatment. Figure 1 illustrates how we put individual aspects of the price mechanism to the test.

Our main findings are that (1) households have control over their level of consumption although with limited precision, and (2) that households are ignorant of the price but not rationally so. Although acquiring accurate price information is possible at virtually zero cost and although we show that households find their in-home device easy to use, (3) they can neither indicate the price they pay nor give a ball park figure. Moreover, (4) households have false beliefs as to why they reduced their consumption between the pre-trial and the trial period. (5) We find some evidence for behavioral costs. Overall, the complete lack of price knowledge rules out that households react to prices.

3.1.1 Aspect I: Control over consumption

Aspect I) focuses on whether individuals have control over consumption. In the context of energy consumption and heating, this issue has been highlighted by Buhl et al. (2017), Du et al. (2018), and Gianniou et al. (2018). Aspect Ia) pursues the question whether consumption is flexible. Should households lack discretion about their natural gas consumption, exogenous price increases cannot influence consumption. It is common knowledge that households consume more natural gas for heating in winter than in summer.

Figure 1: *Mechanism*



This figure reconciles individual aspects that have been identified in previous research as important for a functional price mechanism. We analyze each element in turn. Testing the mechanism may reveal false positives on the price effect estimates. It also helps to differentiate between analyses lacking power and analyses that do not find effects when there are none.

This is confirmed by Figures 2 and 3 we present at a later stage of our analysis. The finding that households have control over their consumption is in favor of a functioning price mechanism.

Aspect Ib) explores whether individuals are able to precisely regulate their consumption. 274 of the 426 households (64.3%) underlying our analysis of the price effects indicate they have a thermostat at home. This device allows them to set the in-home temperature level. Although this does not allow them to precisely determine the level of natural gas they consume, they can at least determine the level of heat they produce in their home. 149 out of 426 (35.0%) do not have such a thermostat which makes controlling consumption harder. The remaining 3 households do not know whether such a device is installed in their home. Overall, the majority of households has at least a certain degree of control over their consumption which is again in favor of a functioning price mechanism.

3.1.2 Aspect II: Information acquisition and processing

Aspect IIa) raises the question whether information costs are prohibitive. If households arrive at the conclusion that the costs outweigh the benefits, households remain rationally ignorant (Downs, 1957; Nyborg, 2011; Jessoe and Rapson, 2014b; Lopez de Leon and Rizzi, 2014; Kim and Yoon, 2019). However, as indicated in the introduction, the natural gas price is probably the easiest and least costly to be fully informed about. Moreover, the post-trial survey reveals that the overwhelming majority of households find that the in-home device that provides price information is either easy or very easy to use (139 of 164 participants or 84.8%). This indicates that we did not exaggerate when saying that accurate information was available at the press of a button. With information costs close to zero, we can rule out rational ignorance.

Aspect IIb) pursues whether households acquire price information when costs are close to zero. Table 15 in the Appendix shows individuals' responses to being asked about the natural gas price. The actual price was between 3.3 and 4.9 Euro cents per kWh while individuals' responses are usually much higher. The large share of households indicated prices of 20 Euro cents and more. In the control group, 12 of 172 individuals or about 7% indicated the correct price.

Table 4 shows that households in the control group (Group 3) indicated on average a price of 58 Euro cents per kWh when the actual price was 3.9 Euro cents. A formal test whether stated and actual prices were identical rejects this hypothesis with $t = 7.610$. The Bayes factor indicates that the probability that actual and stated prices are equal is about 1 in 4 billion, an extreme level of evidence against price knowledge. Treated households (Group 4) indicated on average that the price in September was 60.2 Euro cents per kWh while the actual one was 3.3 Euro cents. With $t = 3.104$ we reject that actual and stated prices are identical. Because of the lower number of observations compared to the control group, the Bayes factor indicates that the alternative hypothesis is 0.104 times as likely as the null. For December, treated households indicated on average a price of 49.8 Euro cents per kWh, while the actual price was 4.6 Euro cents. With $t = 3.056$, we again reject that actual and stated prices are identical. The Bayes factor is 0.119. According to Jeffreys (1961), these two Bayes factors offer substantial evidence against the hypothesis that treated households know their natural gas price or the changes that occur over the course of the experiment (see Table 16 in the Appendix). Overall, these tests indicate that households do not know the price. In fact, it is apparent that households are unable to even give a ball park figure. This is in stark contrast to textbook theory but is in line with the literature on price salience that often finds that individuals are less well informed than they theoretically should be, for example Bordalo et al. (2013), Buchanan et al. (2015), Sexton (2015), Blake et al. (2018), or Alberini et al. (2019).

We test whether respondents who answered questions regarding their natural gas price differed from those who did not. To this end, we created a binary variable that takes on a value of 1 in case that respondents answered the questions regarding their natural gas price in the post-treatment survey and 0 otherwise. We regressed the same variables that we used to test the internal validity on that binary variable. Table 17 in the Appendix shows that observable characteristics do not explain participation in the post-treatment survey. Furthermore, those who receive the price treatment do also not differ with respect to home improvements such as additional insulation of the house, new double glazed windows, and others. To this end, we added these variables as additional controls to the regression we used for testing the internal validity. However, none of these controls are statistically significant (Table 18 in the Appendix). Last, there is no evidence to suggest that higher education levels correlate with higher levels of price knowledge.

Table 4: *Differences between actual and stated prices*

group	N	indicated price	actual price	t	BF_{10}
control	172	58.0	3.9	7.610^{**}	0.000
treatment (September)	32	60.2	3.3	3.104^{**}	0.104
treatment (December)	29	49.8	4.6	3.056^{**}	0.119

This table returns the results from testing whether indicated and actual prices are identical. The first column indicates whether responses are from the control or the treatment group. The second column indicates the number of observations N. Column 3 is the natural gas price households indicated on average. Column 4 indicates the actual price. We provide t-values in the penultimate column. The final column indicates Bayes factors. For example, in the second row, the alternative hypothesis is 0.104 times as likely as the null hypothesis. ** (*) indicates statistical significance at the 1% (5%) level. Prices in Euro cents per kWh.

Our findings are in stark contrast to the pervasive assumption that households know and react to prices. The evidence we uncovered unambiguously shows that the price mechanism is dysfunctional in the case at hand. Consequently, when the experimenters exogenously vary the price, households remain oblivious of these changes and, hence, have no reason to alter their consumption. At this point, it is clear that we have to update the hypotheses from Table 3 to reflect that without price knowledge price effects must be zero.

One possible explanation for the total lack of price knowledge is that individuals have only a limited amount of attention. Kőszegi and Matějka (2020) argue that households simplify decision making because the amount of attention they can allocate is limited. Our data implies that households simplify to the point of total ignorance. Without price knowledge, it remains unclear how price or total cost increases trigger long-term reactions such as home improvements. The post-trial survey offers some evidence supporting this hypothesis. About two-thirds of all households spent on average less than two and a half minutes to review a gas bill. In Ireland, all households receive bi-monthly gas bills by letter. Households who were subjected to price variation indicated in the post-trial survey that they either had not considered the tariff or their bills at all (18%) or had spent less than 15 minutes over the course of the 12 months trial to familiarize themselves with the tariff or the six natural gas bills they had received (49%). Overall, the data reveals that a majority of individuals neither pays attention to natural gas prices nor their overall bills. Further, a limited amount of attention may also explain why households are unaware of their consumption. Only 31 out of 103 post-trial survey respondents agreed that they knew how much natural gas their individual appliances consumed (see Table 19 in the Appendix).

Aspect IIc) analyzes whether households process and fully understand the information they acquire. Informed households who respond rationally to prices and price changes should be able to indicate the reasons as to why they adjusted their natural gas consumption. But if households make mistakes when processing information, for example because of not paying attention or because of bounded rationality, the treatment effect is unclear. Evidence of households' low energy- and financial literacy has been presented for example by Blasch et al. (2018) and Laes et al. (2018). When treated households were asked whether the price was the reason they reduced their consumption, 73 out of 182 (40%) agreed. However, we showed that households have no knowledge of the prices they pay for natural gas. Yet, we also find that all households considerably reduced their natural gas consumption between the six months pre-trial and the same six months in the trial period (Table 20 in the Appendix). This overall reduction occurred because the winter in the pre-trial period was the coldest since 1962/1963 (Met, 2010) and households used more natural gas in the pre-trial period accordingly. Overall, households falsely believe that they lowered their consumption because of price increases when in fact they know nothing about them.

3.1.3 Aspect III: Optimization

Aspect III) is about optimization behavior, an issue prominently highlighted for example in Simon (1957), Simon (1959), Brown and Deaton (1972), Simon (1978), Verhallen and Fred van Raaij (1986), and Kahneman (2003), and as analyzed in the context of energy consumption for example in Barr et al. (2005), Owens and Driffill (2008), Yohanis (2012), De Meester et al. (2013), Bogliacino et al. (2015), or Frederiks et al. (2015). Optimizing behavior includes two aspects: convenience, and comfort.

With regards to convenience, there may exist behavioral change costs, for example the inconvenience of having to adjust the thermostat manually at different times of the day. The post-trial survey indicates that a sizable share of participants may face such costs. When asked whether it was too inconvenient to reduce natural gas consumption, 43 out of 103 post-trial respondents agreed. This is at odds with rational optimization where marginal price increases in a normal good trigger reductions in its marginal consumption to maximize utility. Not adjusting marginal consumption causes sub-optimal levels of utility.

We are aware that households need a certain amount of heat as a basic need. However, the residential natural gas consumption in Ireland is certainly above this level. While about 10% of households indicate thermostat settings below 18 degree Celsius, 61% use settings between 18 and 20 degrees, while 19% indicate temperatures above 20 degrees.

The aspect relates to the reduced comfort from heating less. The post-trial survey indicates that individuals' optimization behavior may differ from what is usually assumed. When asked whether using less gas meant a less comfortable home, 46 out of 103 disagreed. From a utility optimizing perspective, these individuals consume too much natural gas. If a reduced consumption did not imply less utility, one should instead consume more of other goods and services or increase savings.

However, this may be related to intra-household conflict (Castilla and Walker, 2013; Ashraf, 2009; Bjorvatn et al., 2020). In any household, individual members may have different preferences. Consequently, the actual consumption level may not reflect any individual's preference but a compromise following conflict among the household members. When individuals were asked in the post-trial survey whether they would like to alter their household's natural gas consumption, more than half of all respondents indicated that they could not influence the natural gas consumption of other household members. This may indicate intra-household conflict which is a potential confounder and may introduce bias from endogeneity. With intra-household conflict the treatment effect is unclear. Overall, our findings indicate further impediments to a functioning price mechanism.

3.2 Updated hypotheses

Overall, the evidence we presented highlights that key assumptions of the price mechanism are not met. The most serious piece of evidence against a functioning price mechanism is the absence of price knowledge. The causal mechanism is therefore dysfunctional and we must formulate our expectations regarding the effect of prices on consumption accordingly. Given a total lack of price knowledge, all coefficients for the price treatment should be zero (see Table 5).

Table 5: *Updated hypotheses regarding price treatments*

No.	groups control	groups treatment	N	effect	updated hypothesis
1	3	4	426	June 2010	$\beta = 0$
2	3	4	426	July 2010	$\beta = 0$
3	3	4	426	August 2010	$\beta = 0$
4	3	4	426	September 2010	$\beta = 0$
5	3	4	426	October 2010	$\beta = 0$
6	3	4	426	November 2010	$\beta = 0$
7	3	4	426	December 2010	$\beta = 0$
8	3	4	426	January 2011	$\beta = 0$
9	3	4	426	February 2011	$\beta = 0$
10	3	4	426	March 2011	$\beta = 0$
11	3	4	426	April 2011	$\beta = 0$
12	3	4	426	Mai 2011	$\beta = 0$

This table presents our updated hypotheses after the formal tests of the causal mechanism. Finding that households do not know the natural gas price, the price mechanism cannot work. Accordingly, all hypotheses regarding exogenous price variation have to indicate the absence of price effects. No. indicates the number of the hypothesis. N is the number of observations.

3.3 Descriptive statistics

Table 6 presents descriptive statistics. The Irish Social Science Data Archive (ISSDA) provides the data free of charge upon request. Altogether, the data comprises a rich selection of socio-economic variables and dwelling characteristics for each household. For example, in addition to the treatment status, the data set controls for whether households use natural gas for hot water. Across all experiments, 8 out of 10 households do. There is also a large degree of similarity in the remaining variables. Overall, this may serve as a first indicator that randomization was successful.

Table 6: *Descriptive statistics*

Variable	EUS	monthly	IHD	all info	price
observations	653	394	410	665	426
1 if allocated to treatment group	0.305 (0.461)	0.495 (0.501)	0.515 (0.500)	0.317 (0.466)	0.505 (0.501)
1 if gas is used for hot water	0.796 (0.403)	0.799 (0.401)	0.817 (0.387)	0.802 (0.399)	0.822 (0.383)
1 if household size: 2 persons	0.329 (0.470)	0.325 (0.469)	0.315 (0.465)	0.301 (0.459)	0.291 (0.455)
1 if household size: 3 persons	0.179 (0.384)	0.193 (0.395)	0.190 (0.393)	0.191 (0.393)	0.216 (0.412)
1 if household size: 4+ persons	0.326 (0.469)	0.343 (0.475)	0.341 (0.475)	0.329 (0.470)	0.336 (0.473)
1 if house is semi-detached	0.538 (0.499)	0.528 (0.500)	0.527 (0.500)	0.555 (0.497)	0.577 (0.495)
1 if house is detached	0.176 (0.381)	0.193 (0.395)	0.193 (0.395)	0.168 (0.375)	0.160 (0.367)
1 if house is terraced	0.217 (0.413)	0.221 (0.415)	0.210 (0.408)	0.208 (0.406)	0.202 (0.402)
1 if house is bungalow	0.046 (0.210)	0.048 (0.215)	0.054 (0.226)	0.045 (0.208)	0.035 (0.185)
1 if type of house was not revealed	0.002 (0.039)			0.002 (0.039)	
1 if property is owned	0.355 (0.479)	0.343 (0.475)	0.351 (0.478)	0.370 (0.483)	0.366 (0.482)
1 if property is owned with mortgage	0.565 (0.496)	0.594 (0.492)	0.580 (0.494)	0.552 (0.498)	0.575 (0.495)
1 if property status was not revealed	0.003 (0.055)	0.003 (0.050)		0.003 (0.055)	
1 if house built between 2001 and 2005	0.147 (0.354)	0.155 (0.362)	0.151 (0.359)	0.141 (0.349)	0.153 (0.360)
1 if house built between 1981 and 2000	0.302 (0.459)	0.297 (0.457)	0.317 (0.466)	0.344 (0.476)	0.347 (0.477)
1 if house built between 1936 and 1980	0.412 (0.493)	0.429 (0.496)	0.383 (0.487)	0.373 (0.484)	0.359 (0.480)
1 if house built before 1935	0.116 (0.321)	0.091 (0.289)	0.107 (0.310)	0.111 (0.315)	0.094 (0.292)
1 if number of bedrooms is 2	0.081 (0.273)	0.084 (0.277)	0.076 (0.265)	0.072 (0.259)	0.087 (0.282)
1 if number of bedrooms is 3	0.521 (0.500)	0.475 (0.500)	0.490 (0.501)	0.546 (0.498)	0.523 (0.500)
1 if number of bedrooms is 4	0.328 (0.470)	0.360 (0.481)	0.354 (0.479)	0.317 (0.466)	0.324 (0.469)
1 if number of bedrooms is 5 +	0.057 (0.231)	0.074 (0.261)	0.071 (0.257)	0.051 (0.220)	0.056 (0.231)
1 if number of bedrooms was not revealed	0.003 (0.055)			0.003 (0.055)	

This table indicates means (shares) and standard deviations in parentheses of the variables underlying our analysis. This table is continued below. EUS = energy usage statement, IHD = in-home device. Group names: 0 = control group, 1 = bi-monthly bill, 2 = monthly bill, 3 = IHD, 4 = price variation. Blank cells indicate empty categories.

3.4 Pre-trial consumption

The Irish Gas Customer Behavioural Trials provide data on natural gas consumption for a six months trial period between December 1, 2009 and May 31, 2010 prior to the start of the experiment. Figure 2 compares natural

Descriptive statistics (continued)

Variable	EUS	monthly	IHD	all info	price
observations	653	394	410	665	426
1 if income is between 15,001 and 30,000 Euro	0.126 (0.332)	0.157 (0.365)	0.146 (0.354)	0.123 (0.329)	0.143 (0.351)
1 if income is between 30,001 and 50,000 Euro	0.221 (0.415)	0.234 (0.424)	0.205 (0.404)	0.205 (0.404)	0.207 (0.405)
1 if income is between 50,001 and 75,000 Euro	0.162 (0.369)	0.162 (0.369)	0.156 (0.363)	0.153 (0.361)	0.150 (0.358)
1 if income is over 75,000 Euro	0.193 (0.395)	0.206 (0.405)	0.210 (0.408)	0.192 (0.395)	0.207 (0.405)
1 if income was not revealed	0.237 (0.426)	0.198 (0.399)	0.229 (0.421)	0.257 (0.437)	0.242 (0.429)
1 if household has no alternative heating source	0.571 (0.495)	0.530 (0.500)	0.551 (0.498)	0.561 (0.497)	0.561 (0.497)
1 if share of double glazed windows is up to 0.25	0.020 (0.140)	0.013 (0.112)	0.020 (0.138)	0.023 (0.149)	0.023 (0.152)
1 if share of double glazed windows is up to 0.50	0.040 (0.196)	0.028 (0.165)	0.034 (0.182)	0.039 (0.194)	0.042 (0.201)
1 if share of double glazed windows is up to 0.75	0.026 (0.159)	0.036 (0.185)	0.032 (0.175)	0.024 (0.153)	0.023 (0.152)
1 if all windows are double glazed	0.845 (0.362)	0.863 (0.344)	0.854 (0.354)	0.851 (0.356)	0.859 (0.348)
1 if house has no wall insulation	0.366 (0.482)	0.365 (0.482)	0.351 (0.478)	0.356 (0.479)	0.359 (0.480)
1 if no knowledge about wall insulation	0.167 (0.373)	0.155 (0.362)	0.161 (0.368)	0.164 (0.370)	0.155 (0.362)

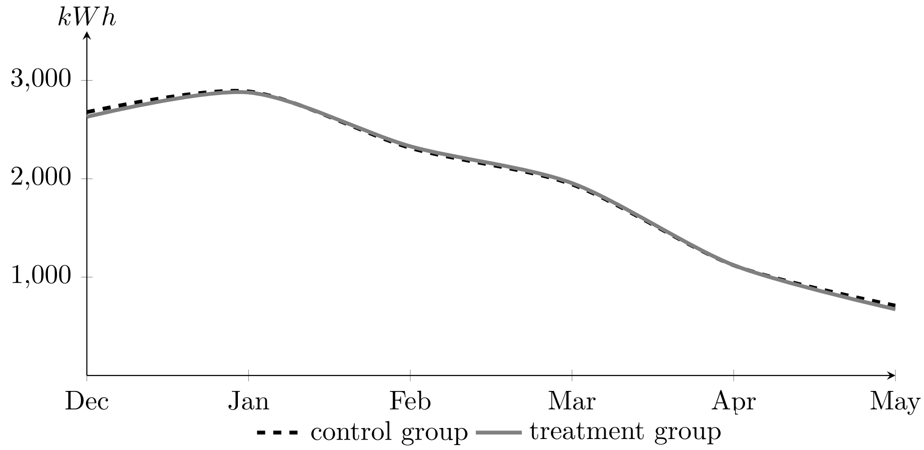
This is the continuation of Table 4. EUS = energy usage statement, IHD = in-home device. Group names: 0 = control group, 1 = bi-monthly bill, 2 = monthly bill, 3 = IHD, 4 = price variation. For further information, see Table 3.

gas consumption across our treatment and control households for the price treatment (Groups 3 and 4) at the monthly level. A cursory visual inspection reveals virtually identical consumption patterns in the pre-treatment period. With $t = -0.132$, a formal test fails to find evidence for significantly different consumption patterns across treatment and control households. Using weak informative priors, we find a Bayes factor of $BF_{10} = 0.077$. The factor describes that the probability of the alternative hypothesis is about 0.077 times as likely as the null or that the null is $1/0.077 \approx 13$ times more likely than the alternative.

3.5 Randomization

With strong evidence that there are no differences in pre-trial consumption, we now turn to analyze whether household characteristics explain treatment status. To this end, we estimate linear probability models (Table 21 in the Appendix) by regressing the binary treatment indicator on household characteristics. For most of the treatments, all coefficients are statistically

Figure 2: *Pre-trial natural gas consumption*



This figure compares the monthly natural gas consumption in kWhs between treated and control households in the pre-treatment period between December 2009 and May 2010.

insignificant. However, there are 2 coefficients which are statistically significant at the 1% level. This raises the question as to whether there exist systematic differences between the control and the treatment group. However, the probability of finding at least 2 statistically significant false positive coefficients by chance at the 5% level in 162 coefficients, is 0.998. At the 1% level, the probability of finding 2 false positives is lower compared to the 5% level but remains high with 0.482. To us, the frequentist results present little to no evidence of systematic differences between the control and the treatment groups. We could potentially control for testing multiple hypotheses, for example by the procedures suggested by Benjamini and Hochberg (1995) or Benjamini and Yekutieli (2001). However, adjusting the p-values to accommodate for a large number of tests would turn all statistically significant differences insignificant.

However, we would like to be able to assess how strongly the data is in support of similar groups. Again, we turn to Bayes factors that can find evidence for similarity. Table 7 informs about the Bayes factors from regressions using mixtures of g-priors (Liang et al., 2008). For all treatments, the probability that the data was generated under the alternative hypothesis of systematic differences across groups is less than 0.000 times as probable as it is under the null hypothesis. This is an extreme level of evidence for highly similar treatment and control groups and internal validity.

Table 7: *Bayesian results on internal validity*

Variable	EUS	monthly	IHD	all info	price
observations	653	394	410	665	426
BF_{10}	0.000	0.000	0.000	0.000	0.000
interpretation	extreme H_0	extreme H_0	extreme H_0	extreme H_0	extreme H_0

This table holds the Bayes factors for our tests on the internal validity. The factor indicates how likely the alternative hypothesis is compared to the null hypothesis. A result of 0.000 indicates an extreme level of evidence in favor of internal validity. Interpretation of Bayes factors according to Table 16.

3.6 Attrition

Initially, the trials aimed at a sample size of 1,927 households determined by a power study accommodating assumptions regarding natural gas consumption levels. This sample size allowed for a 35% attrition rate. Ultimately, 1,892 households volunteered to participate in the trial (Table 8).

Table 8: *Number of observations by experimental stage*

	experimental stage				
	1	2	3	4	5
Control	681	543	524	458	454
bi-monthly bill	303	257	236	204	199
monthly bill	303	248	228	198	196
IHD	303	263	251	217	211
price variation	302	265	254	219	215
Total	1,892	1,576	1,493	1,296	1,275

This table indicates the number of observations by experimental stage. The difference between an earlier and a later stage indicates the number of attritors. Stages: 1 = trial setup, 2 = allocated to treatment groups, 3 = metering data, 4 = pre-trial information, 5 = estimation sample of households using natural gas for heating.

In the time between the contracting of households (stage 1) and the allocation into treatment groups, attrition reduced the number of households in the sample from 1,892 to 1,576. Households dropped out of the trial because they either moved home before the trial started or because the smart-meter was not operational (CER, 2011b, p. 38). Between stages 2 and 3, attrition claims another 83 households, so that the sample includes $1,493$ households for which metering information is available. The sample consisting of households for whom treatment allocation, metering-information, and household characteristics are available comprises $1,296$ observations (see Table 9). Another 21 households were removed from the sample because they indicated to not use natural gas for heating. Therefore, the sample of natural gas using households who use natural gas for heating, who have filled out the pre-trial questionnaire, and for whom metering information is available comprises $1,275$ households. Altogether, the combined rate of attrition and drop-out is lower than the anticipated and accommodated for rate of 35%.

Table 9: *Number of observations by attrition and availability of pre-trial data*

	attrition		Total
pre-trial questionnaire	no	yes	
yes	1,296	69	1,365
no	197	0	197
Total	1,493	69	1,562

This table indicates the number of observations by the availability of pre-trial data and attrition. Note that the final sample excludes 21 households which do not use natural gas for heating.

We continue to test whether observable characteristic explain attrition. To this end, we combine the information on those 68 households who dropped out after having completed the pre-trial questionnaire with the $1,275$ households of the final sample. For the 197 households who dropped out before completing the questionnaire, no household information is available. We estimate a linear probability model that tries to explain the drop-out from the sample (see Table 22 in the Appendix). All coefficients are insignificant and the fit is poor with $R^2 = 0.0297$. With $1,342$ observations, the corresponding Bayes factor is 0.000. This implies that the data under the alternative hypothesis of systematic attrition is less than 0.000 times as likely as it is under the null hypothesis. Overall, there exists an extreme level of evidence for the absence of systematic differences between attritors and non-attritors. We also tested whether the natural gas consumption of the 197 households who dropped out and for whom we do not have any household information differed from those who remained in the final sample. These 197 households annually consumed on average 337 kWhs less than non-attritors. But with $t = -0.68$ the difference is statistically insignificant.

3.7 Household recruitment and external validity

At the time of household recruitment, Bord Gáis Energy represented about 98% of the Irish natural gas market (CER, 2011b, p. 26) or about $430,000$ households in total (CER, 2011a). Prior to the trial, Bord Gáis Energy was privy to certain information such as contact details, meter locations, house types (semi-detached, detached, terrace, flat, bungalow), payment information, and the level of natural gas consumption for the preceding years. Based on these variables, a number of restrictions were implemented to safeguard the representativeness of the sample.

First, at the outset of the recruitment process households had to be at their address for at least 12 months. This ensured that the natural gas consumption in the previous year could be used for stratified sampling. Second, households who used prepayments were excluded. Third, meter and billing address had to be identical to rule out that the metered location was a second home which might only be occupied at certain times of the year.

The recruitment process resulted in a sample with a natural gas consumption profile that is highly similar to that of the general population (CER, 2011b). The same applies to the distribution of the age of the home. There is also general agreement between the sample and the population with respect to the geographical distribution and the house type (see Table 10). The upper panel indicates that the geographical distribution of sampled households is virtually identical compared to the population. The lower panel shows that house types are also very similar between the sample and the population.

Because participation in the trial was voluntary, a non-response survey was carried out to track whether there were systematic differences between participants and non-responders. Altogether, differences across groups are again small (Table 11). However, because the data from the non-response

Table 10: *Location and house type*

	sample	population
location		
Clonmel	0.01	0.01
Cork	0.11	0.13
Dublin	0.64	0.64
Kilkenny	0.01	0.01
Limerick	0.04	0.05
Non Dublin	0.16	0.13
Waterford	0.02	0.03
house type		
semi-detached	0.62	0.57
bungalow	0.00	0.00
detached	0.20	0.18
flat	0.02	0.07
terrace	0.15	0.17

This table compares location and house type between the our final sample and the population. Source: (CER, 2011b, p. 125).

survey is unavailable, we cannot conduct formal tests and have to rely on the analysis by CER (2011c).

Table 11 presents household and home characteristics of trial participants and non-responders. Overall, there are no significant differences between both samples. The similarity is particularly close for the type, the age, and the energy efficiency of the home. The number of household members in terms of adults and children is also highly similar. The type of home and the number of bedrooms are virtually identical in both samples. About 4% to 5% of households live in a bungalow, about 21% in a terraced house, 18% to 20% in detached houses, 51% to 55% in semi-detached houses, while between 2% and 4% live in apartments. Altogether, the comparisons between the sample and the population indicate a high level of similarity. The same is true for the comparisons between participants and the sample of non-respondents. Overall, we provide evidence that the sample is highly representative of the population, so that the gas metering trials have external validity.

3.8 Treatment effects

As a first step in assessing treatment effects, we visually compare the average natural gas consumption across control (Group 3) and treatment households (Group 4). Figure 3 indicates that the mean natural gas consumption is

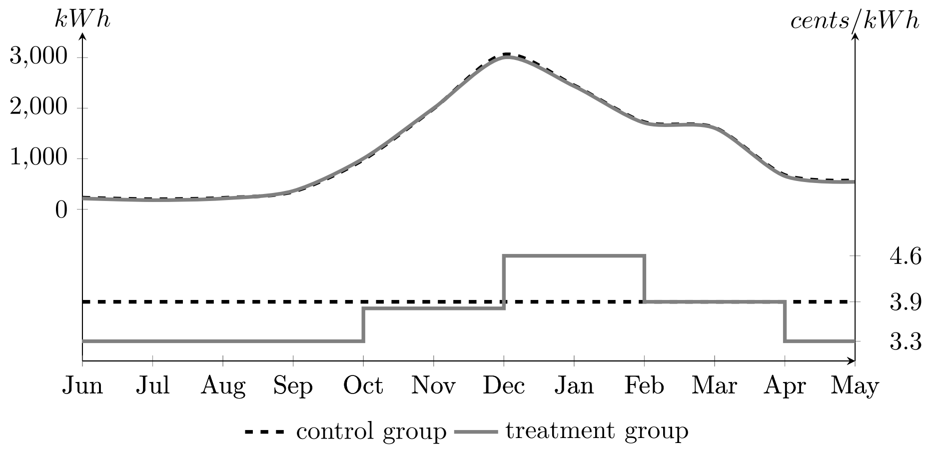
Table 11: *Responders vs. non-responders*

	non-responders	participants
occupation		
carer	0.01	0.01
retired	0.22	0.33
unemployed	0.07	0.10
self-employed	0.11	0.08
employed	0.59	0.48
adults in household		
1	0.20	0.19
2	0.47	0.54
3	0.21	0.16
4	0.06	0.07
5+	0.06	0.03
children in household		
0	0.67	0.63
1	0.11	0.16
2	0.16	0.14
3	0.05	0.05
4	0.01	0.01
ownership		
rent from private landlord	0.01	0.02
rent from local authority	0.07	0.05
own outright	0.48	0.36
own with mortgage	0.42	0.56
other	0.00	0.00
home type		
bungalow	0.04	0.05
terraced house	0.21	0.21
detached house	0.18	0.20
semi-detached house	0.55	0.51
apartment	0.02	0.04
number of bedrooms		
1	0.02	0.01
2	0.10	0.09
3	0.48	0.51
4	0.31	0.33
5+	0.10	0.06
age of home in years		
less than 5	0.00	0.04
less than 10	0.08	0.15
less than 30	0.32	0.31
less than 75	0.49	0.40
more than 75	0.11	0.10
energy efficiency of home		
lagging jacket	0.84	0.89
attic insulated in last 5 years	0.21	0.28
attic insulated more than 5 years ago	0.65	0.62
external walls insulated	0.48	0.48
boiler serviced every year	0.63	0.55
boiler serviced every 2 to 3 years	0.29	0.36

This table compares socio-economic characterisitics of responders and non-responders. Source: (CER, 2011b, p. 125).

virtually identical across the treatment groups and hypotheses tested. The price, however, varies considerably by design. Given that consumption levels were also virtually identical in both groups before the onset of the treatment (see Figure 2), this is a first indicator that the treatment has no effect.

Figure 3: *Natural gas consumption and prices by treatment*



This figure compares the control to the treatment group in terms of consumption (left y-axis) and natural gas prices in Euro cents per kWh (right y-axis).

For formal tests, we estimate regression models following

$$y = \gamma \cdot \text{treatment} + \beta X + \epsilon \ , \tag{1}$$

where y is the monthly natural gas consumption, treatment is a binary treatment indicator and X is a matrix that holds the same control variables that we used to determine random allocation. For causal identification, we assume unit homogeneity (Holland, 1986) and stable unit treatment values (Cox, 1958). This is reasonable because the participants of the trial are dispersed across Ireland and do not know each other. Moreover, the number of treated households and the size of the treatment are of insufficient magnitude to affect the local economy or influence households in the control group in any other way.

Table 12 presents the treatment effects at the monthly level. The first column indicates the time period of comparison. The second column holds the expected effect, for example $\beta > 0$ indicates that treatment should increase consumption. The final columns indicate the coefficients and the standard errors.

Table 12: *Treatment effects (N=426)*

Treatment	hypotheses	coef.	std. err.
June 2010	$\beta > 0$	−7.397	20.279
July 2010	$\beta > 0$	−9.160	20.661
August 2010	$\beta > 0$	−2.247	21.676
September 2010	$\beta > 0$	26.668	21.611
October 2010	$\beta > 0$	48.963	31.772
November 2010	$\beta > 0$	34.656	38.665
December 2010	$\beta < 0$	3.647	52.042
January 2011	$\beta < 0$	21.120	45.985
February 2011	$\beta = 0$	13.103	38.690
March 2011	$\beta = 0$	20.209	42.880
April 2011	$\beta > 0$	5.532	35.530
May 2011	$\beta > 0$	−4.269	33.471

This table provides an overview of the treatment effects. The first column indicates the month for which the test is carried out. The second column holds the expected effect, for example $\beta > 0$ indicates that treatment should increase consumption. In February and March both groups have the same natural gas price. These two months may serve as placebo treatment. The final columns indicate the coefficients and the standard errors.** (*) indicates statistical significance at the 1% (5%) level. Coef. is for coefficient. Std. err. is for standard error.

In February and March, treated and control households paid the same price for natural gas. These two months may serve as placebo tests. Therefore, we expect to find similar consumption levels. Although the treatment group consumed additional 13 kWhs in February and 21 kWhs in March of natural gas, the difference is statistically insignificant. For all other months, with the exception of December and January, treated households pay less than control households. Nevertheless, observed differences in consumption are minor and statistically insignificant. In August and September when the treated pay 15.4% less the point estimates remain close to zero and insignificant. We also do not find any differences in consumption for the heating period during the cold months of the year. In December and January when the treated pay 17.9% more than control households, point estimates are positive, when we would expect negative signs. The 95% lower bounds for single-sided t-tests is 82 kWh in December and 72 kWh in January. With consumption levels in the control group of $3,060$ and $2,450$ kWh, maximum relative reductions are between 2.6% and 2.9% for a 17.9% price increase.

This implies maximum elasticities between -0.15 and -0.16. Overall, we tested 12 hypotheses out of which 10 are one-sided. The point estimates for 6 out of these 10 have the wrong sign. Because all estimates are statistically insignificant, we do not adjust p-values to account for the testing of multiple hypotheses. Overall, the wrong signs, the small standard errors, and the lack of statistical evidence leads us to conclude that the price does not inform household natural gas consumption - at least not in the short-run.

3.9 Bayesian analysis

Frequentist statistics indicated no statistically significant effects. For readers who are interested in a Bayeisan analysis, we now present results from Bayesian regressions on the same models as in the preceding frequentist analysis. Figure 4 in the Appendix presents trace plots that provide evidence that the MCMC processes are stable for all treatments and neither need burning nor thinning.

We use weak informative priors to achieve completely data driven results. Consequently, our Bayesian results are virtually identical to those presented before. We are able to calculate Bayes factors that indicate the odds that the price treatment has no effect on natural gas consumption. Table 13 holds the results along with their interpretation. For February and March, when control and treatment households have the same price, Bayes factors indicate strong evidence in favor of no systematic differences. This is what we should expect from a placebo treatment. When the prices are lowest for the treated in August and September (-15.5%), there is substantial evidence for no systematic differences. The Bayes factors indicate that the alternative hypotheses of systematic differences are 0.131 and 0.183 times as likely as the null. In other words, the null hypotheses are between 5.5 and 7.6 times more likely. When prices are highest for treated households in December and January, the alternative hypotheses are 0.034 and 0.042 times as likely as the nulls. Overall, Bayes factors provide further substantial to very strong evidence that the price treatment had no effect on natural gas consumption.

Table 13: *Bayes factors using regression designs* ($N = 426$)

Treatment	coef	error	BF_{10}	interpretation
June 2010	−7.689	19.556	0.121	substantial H_0
July 2010	−9.464	19.851	0.139	substantial H_0
August 2010	−1.672	20.946	0.112	substantial H_0
September 2010	25.553	20.939	0.184	substantial H_0
October 2010	46.618	30.919	0.206	substantial H_0
November 2010	34.822	37.819	0.071	substantial H_0
December 2010	5.292	50.672	0.014	very strong H_0
January 2011	23.098	44.781	0.019	very strong H_0
February 2011	13.523	37.561	0.021	very strong H_0
March 2011	18.817	41.736	0.068	strong H_0
April 2011	1.162	34.541	0.092	strong H_0
May 2011	−8.420	32.499	0.101	substantial H_0

This table provides an overview of the treatment effects. The first column indicates the month of treatment. The second column indicates the coefficient, while the third column indicates the error. Column 4 holds the Bayes factor that indicate how likely the alternative hypothesis is compared to the null hypothesis. For example, $BF_{10} = 014$ indicates that the alternative hypothesis is 0.014 times as likely as the null. The final column interprets the Bayes factors according to Table 16.

3.10 Heterogeneity analysis

We tested whether treatment effects depend on household sizes by interacting household size with the treatment variable:

$$y = \gamma \cdot \text{treatment} + \varphi \cdot H + \phi \cdot \text{treatment} \cdot H + \beta X + \epsilon \ , \tag{2}$$

where H is a categorical variable that measures household size. Table 14 shows that out of 60 treatment coefficients three are statistically significant at the 5% level. Controlling for multiple hypothesis testing renders these coefficients insignificant.

We also tested whether the treatment depended on the level of natural gas consumption itself. To this end, we estimated unconditional quantile regressions following Firpo et al. (2009). Table 23 in the Appendix shows that the impact of the treatment does not depend on the level of natural gas consumption.

Table 14: *Treatment effects by household size*

Month	household size			
	1	2	3	4+
June 2010	−22.040	−53.808	48.719	−16.006
	(51.404)	(37.659)	(43.960)	(34.873)
July 2010	7.840	−51.342	37.781	−9.896
	(52.449)	(38.425)	(44.854)	(35.582)
August 2010	14.201	−76.499	61.919	14.252
	(54.809)	(40.154)	(46.872)	(37.183)
September 2010	30.826	−46.436	100.142*	42.126
	(54.616)	(40.012)	(46.707)	(37.052)
October 2010	−10.044	18.718	98.947	71.556
	(80.751)	(59.159)	(69.057)	(54.782)
November 2010	−153.862	−35.477	147.092	113.373
	(97.491)	(71.423)	(83.373)	(66.138)
December 2010	−186.142	−51.783	−12.513	150.310
	(131.667)	(96.461)	(112.599)	(89.323)
January 2021	−217.499	−26.969	169.167	82.052
	(116.023)	(84.999)	(99.220)	(78.710)
February 2021	−170.590	−40.816	178.995*	42.283
	(97.504)	(71.432)	(83.383)	(66.147)
March 2021	−166.811	−75.059	199.194*	78.509
	(107.996)	(79.119)	(92.356)	(73.265)
April 2021	−61.463	−81.151	112.594	45.076
	(89.950)	(65.898)	(76.924)	(45.076)
May 2021	−42.576	−123.400*	77.258	65.784
	(84.525)	(61.924)	(72.284)	(57.342)

This table returns treatment coefficients for different household sizes. The numbers 1-5+ indicate the number of individuals living in a household. ** (*) indicates statistical significance at the 1% (5%) level. Standard errors in parentheses.

3.11 Explaining the results of previous analyses of the Irish Gas Customer Behavioural Trials

The Irish Gas Customer Behavioural Trials have been analyzed before. The technical reports by the original experimenters (CER, 2011a,b,d) indicate that when comparing those who received any kind of treatment compared to those who did not receive any kind of treatment, the treated consumed about 2.9% less natural gas. This finding is statistically significant at the 10% level. However, their conclusion when analyzing individual treatments is the following: “When the stimuli are compared with each other, the statistical tests fail to find evidence of a particular stimulus being superior from any other in terms of reduction in gas consumption.” This means that there are no significant effects when comparing the outcomes of Group 3 and Group 4 who are identical in every respect but the price treatment. In short, the original experimenters found no evidence that increasing prices reduce natural gas consumption. However, the experimenters do not conduct formal tests to explain why the price variation does not yield any effect.

Harold et al. (2018) also analyze the Irish Gas Customer Behavioural Trials. They find statistically significant effects when comparing combinations of treatments to a control group that receives no treatments. But similar to the original study, they fail to find any statistically significant effects for the price treatment. The authors believe that the reason for the absence of significant results is "most likely [due] to a sampling size issue with just around 200 households per individual treatment." While a lack of statistical power in small samples is often the reason for insignificant results, the insignificant price coefficients in this analysis are caused by a dysfunctional price mechanism.

We scrutinize the mechanism by which prices are hypothesized to influence the consumed quantity. By providing strong evidence that households do not know the price of natural gas, we are able to explain the insignificant results of the previous analyses. Our main contribution is to highlight that it is crucial to test the assumptions underlying the hypothesized mechanism which is the basis for the expected treatment effects. Otherwise, treatment effects cannot be ascribed a causal interpretation. Testing the price mechanism prevents drawing false conclusions from insignificant price effects such as a flawed research design or a lack of statistical power. It also makes false positives less likely.

4 Conclusion

In this piece, we examined the effect of exogenous price treatments on residential natural gas consumption. By explicitly testing the price mechanism, we are able to explain findings and interpretations of small or zero price effects from previous research using a framework that could inform further research designs and the interpretation of empirical estimates. Our analysis shows that researchers need to provide evidence for a functional price mechanism when estimating price elasticities and before ascribing causal price effects, or risk biased conclusions. Testing the price mechanism enables the detection of false positives. In addition, it helps to differentiate between analyses lacking power and analyses that do not find effects when there are none.

The data underlying our analysis is from the Irish Gas Customer Behavioural Trials. The data exhibits internal and external validity. The treatment is of considerable magnitude with treated households facing exogenous price variations between -15.4% and $+17.9\%$. Control households pay a flat tariff of 3.9 cents per kWh. Overall, the Irish Gas Metering Trials have all the hallmarks of a well-designed and well-executed large field experiment that stretches over an entire year with a six month pre-trial period and a post-trial survey. However, there are no significant treatment effects. Point estimates close to zero in combination with small standard errors suggest that the exogenous price changes did not inform natural gas consumption.

Our empirical analysis hosts a battery of tests. Our frequentist analysis finds no evidence that randomization was unsuccessful. Further, there is no evidence to suggest systematic attrition. There is no evidence for systematically different pre-trial consumption levels of natural gas. However, we find highly significant differences between actual and stated prices. This leads us to conclude that exogenous variations in the actual natural gas price cannot inform consumption decisions in the short-run. Long-term effects are beyond the scope of our analysis. However, long-term reactions to price increases in natural gas, for example home improvements, also require knowledge of price increases. Treatment effects are insignificant with point estimates close to zero with small standard errors and with 6 out of 10 having the wrong sign. Therefore, we conclude that the treatment has no effect on natural gas consumption which is in line with the finding that households no not know the actual price of natural gas.

Applying Bayesian analysis to cross-check our results, we arrive at the same conclusions. Those who accept Bayes factors as valid tools for analysis, may draw stronger conclusions. We find an extreme level of evidence in favor of successful randomization. We also find strong evidence in favor of similar pre-trial consumption levels. We find an extreme level of evidence in favor of random attrition.

When testing the price mechanism, we show that its underlying assumptions are not met in our case. Despite the fact that all relevant information to optimize individual consumption is available at zero information costs, we find a probability of 1 in 4 billion that control households know the actual price of natural gas. This is an extreme level of evidence against price knowledge. We find substantial evidence that neither the treated know the price. In addition, half of the participants would find it inconvenient to adjust their behavior, which may indicate the presence of non-negligible behavioral adjustment costs. Further, 30% of households indicate that they do not know how much natural gas individual appliances consume. Overall, our findings highlight that it is necessary to test the price mechanism carefully to avoid drawing biased conclusions.

Our findings have important implications. In particular, the validity of the mechanism cannot be inferred from estimation results that are in line with hypothesized effects. Our setting shows that households in the control group who only had to know a single price set by a monopolist for an entire year failed to achieve full price information. There is no reason to assume why households should be better informed when time-varying gas price are implemented that fluctuate at a higher frequency making it harder to follow them. Also, the provision of information may not have the desired effects because households pay it little to no attention. From the evidence we provided we conclude that researchers must provide formal evidence in favor of their hypothesized causal mechanism before ascribing causal effects, particularly when assuming that observation units know prices. In the absence of formal proof, researchers should be suspicious of false positives.

Of course, our warning also extends to the application of price instruments which are typically regarded as first best solutions. However, price instruments cease to be solutions if prices are dysfunctional in a large share of households. For example, burning natural gas to generate heat leads to carbon dioxide emissions which, in turn, contribute to climate change. If policy makers decided to implement a carbon charge to internalize this externality, our analysis shows that households likely will not even realize that the cost of using natural gas has increased. The charge would, first, fail to reduce natural gas consumption, and, second, lower available households' budgets and thereby their utility. Substituting dynamic for fixed pricing likely has the same effect in our context. In general, first best solutions may be vastly inferior to other forms of regulation depending on the level of price functionality.

References

Alberini, A. (2018). Household energy use, energy efficiency, emissions, and behaviors. *Energy Efficiency 11*, 577–588.

Alberini, A., O. Khymych, and M. Scasny (2019). The Elusive Effects of Residential Energy Efficiency Improvements: Evidence from Ukraine. *SSRN Electronic Journal*.

Allcott, H. (2011). Rethinking real-time electricity pricing. *Resource and Energy Economics 33*(4), 820–842.

Allcott, H. (2015). Site selection bias in program evaluation. *The Quarterly Journal of Economics 130*(3), 1117–1165.

Allcott, H. and T. Rogers (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review 104*(10), 3003–37.

Angrist, J. D. and J.-S. Pischke (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives 24*(2), 3–30.

Ashraf, N. (2009). Spousal control and intra-household decision making: An experimental study in the philippines. *American Economic Review 99*(4), 1245–77.

Barr, S., A. W. Gilg, and N. Ford (2005). The household energy gap: examining the divide between habitual-and purchase-related conservation behaviours. *Energy Policy 33*(11), 1425–1444.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological) 57*(1), 289–300.

Benjamini, Y. and D. Yekutieli (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 1165–1188.

Bhargava, S. and D. Manoli (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an irs field experiment. *American Economic Review 105*(11), 3489–3529.

Bjorvatn, K., T. D. Getahun, and S. K. Halvorsen (2020). Conflict or cooperation? experimental evidence on intra-household allocations in ethiopia. *Journal of Behavioral and Experimental Economics 85*, 101508.

Blake, T., S. Moshary, K. Sweeney, and S. Tadelis (2018). Price salience and product choice. Technical report, National Bureau of Economic Research.

Blasch, J., N. Boogen, C. Daminato, and M. Filippini (2018). Empower the consumer! Energy-related financial literacy and its socioeconomic determinants. *CER-ETH Working Paper 18*(289).

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

Blundell, R. W. and J. L. Powell (2004). Endogeneity in semiparametric binary response models. *The Review of Economic Studies 71*(3), 655–679.

Bogliacino, F., C. Codagnone, and G. A. Veltri (2015). The behavioural turn in consumer policy: Perspectives and clarifi cations. *Intereconomics 50*(2), 108–114.

Bordalo, P., N. Gennaioli, and A. Shleifer (2013). Salience and consumer choice. *Journal of Political Economy 121*(5), 803–843.

Borenstein, S. (2002). The trouble with electricity markets: Understanding California's restructuring disaster. *Journal of Economic Perspectives 16*(1), 191–211.

Borenstein, S. (2005). The long-run efficiency of real-time electricity pricing. *The Energy Journal 26*(3), 93–116.

Bretschger, L. and E. Grieg (2020). Exiting the fossil world: The effects of fuel taxation in the uk. *Economics Working Paper Series 20*.

Brown, A. and A. Deaton (1972). Surveys in applied economics: models of consumer behaviour. *The Economic Journal 82*(328), 1145–1236.

Buchanan, K., R. Russo, and B. Anderson (2015). The question of energy reduction: The problem(s) with feedback. *Energy Policy 77*, 89–96.

Buhl, J., M. Hasselkuß, P. Suski, and H. Berg (2017). Automating Behavior ? An Experimental Living Lab Study on the Effect of Smart Home Systems and Traffic Light Feedback on Heating Energy Consumption. *Current Journal of Applied Science and Technology 22*(4), 1–18.

Bundorf, M. K., J. Levin, and N. Mahoney (2012). Pricing and welfare in health plan choice. *American Economic Review 102*(7), 3214–48.

Castilla, C. and T. Walker (2013). Is ignorance bliss? the effect of asymmetric information between spouses on intra-household allocations. *American Economic Review 103*(3), 263–68.

CER (2011a). *Cost-Benefit Analysis (CBA) for a National Gas Smart Metering Rollout in Ireland.* Commission for Energy Regulation.

CER (2011b). *Gas Customer Behavioural Trial Findings Report.* Commission for Energy Regulation.

CER (2011c). Gas customer behavioural trial findings report.

CER (2011d). *Gas Customer Behavioural Trial Findings Report Appendices.* Commission for Energy Regulation.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and Evidence. *American Economic Review 99*(4), 1145–77.

Cox, D. R. (1958). *Planning of experiments.* Wiley.

Davis, L. W. and L. Kilian (2011). The allocative cost of price ceilings in the US residential market for natural gas. *Journal of Political Economy 119*(2), 212–241.

Davis, L. W. and G. E. Metcalf (2016). Does better information lead to better choices? Evidence from energy-efficiency labels. *Journal of the Association of Environmental and Resource Economists 3*(3), 589–625.

Davis, L. W. and E. Muehlegger (2010). Do americans consume too little natural gas? An empirical test of marginal cost pricing. *The RAND Journal of Economics 41*(4), 791–810.

De Bondt, W. F. and R. Thaler (1985). Does the stock market overreact? *The Journal of Finance 40*(3), 793–805.

De Meester, T., A.-F. Marique, A. De Herde, and S. Reiter (2013). Impacts of occupant behaviours on residential heating consumption for detached houses in a temperate climate in the northern part of europe. *Energy and Buildings 57*, 313–323.

Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature 48*(2), 424–55.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy 65*(2), 135–150.

Du, C., B. Li, Y. Cheng, C. Li, H. Liu, and R. Yao (2018). Influence of human thermal adaptation and its development on human thermal responses to warm environments. *Building and Environment 139*, 134–145.

Duflo, E. and E. Saez (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of Economics 118*(3), 815–842.

Fabra, N. and M. Reguant (2020). A model of search with price discrimination. *European Economic Review 129*, 103571.

Finkelstein, A. (2009). E-ztax: Tax salience and tax rates. *The Quarterly Journal of Economics 124*(3), 969–1010.

Firpo, S., N. M. Fortin, and T. Lemieux (2009). Unconditional quantile regressions. *Econometrica 77*(3), 953–973.

Fowlie, M., C. Wolfram, C. A. Spurlock, A. Todd, P. Baylis, and P. Cappers (2017). Default effects and follow-on behavior: Evidence from an electricity pricing program. Technical report, National Bureau of Economic Research.

Frederiks, E. R., K. Stenner, and E. V. Hobman (2015). Household energy use: Applying behavioural economics to understand consumer decision-making and behaviour. *Renewable and Sustainable Energy Reviews 41*, 1385–1394.

Gianniou, P., C. Reinhart, D. Hsu, A. Heller, and C. Rode (2018). Estimation of temperature setpoints and heat transfer coe ffi cients among residential buildings in Denmark based on smart meter data. *Building and Environment 139*, 125–133.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Grubb, M. D. and M. Osborne (2015). Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review 105*(1), 234–71.

Hahn, J. and J. Hausman (2003). Weak instruments: Diagnosis and cures in empirical econometrics. *American Economic Review 93*(2), 118–125.

Harari, M. (2020). Cities in bad shape: Urban geometry in India. *American Economic Review 110*(8), 2377–2421.

Harold, J., S. Lyons, and J. Cullinan (2018). Heterogeneity and persistence in the effect of demand side management stimuli on residential gas consumption. *Energy Economics 73*, 135–145.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association 81*(396), 945–960.

Holland, S. P. and E. T. Mansur (2006). The short-run effects of time-varying prices in competitive electricity markets. *The Energy Journal 27*(4), 127–155.

Hossain, T. and J. Morgan (2006). ... plus shipping and handling: Revenue (non) equivalence in field experiments on eBay. *The BE Journal of Economic Analysis & Policy 6*(2).

Ito, K. (2014). Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing. *American Economic Review 104*(2), 537–63.

Jeffreys, H. (1961). *Theory of probability.* Clarendon Press, Oxford.

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *The Quarterly Journal of Economics 122*(3), 879–924.

Jessoe, K. and D. Rapson (2014a). Knowledge is (less) power: Experimental evidence from residential energy use. *American Economic Review 104*(4), 1417–38.

Jessoe, K. and D. Rapson (2014b). Knowledge is (Less) power: Experimental evidence from residential energy use. *American Economic Review 104*(4), 1417–1438.

Joskow, P. L. and C. D. Wolfram (2012). Dynamic pricing of electricity. *American Economic Review 102*(3), 381–85.

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

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives 5*(1), 193–206.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–292.

Kim, D. G. and Y. Yoon (2019). A theory of faqs: Public announcements with rational ignorance. *Journal of Economic Behavior & Organization 158*, 560–574.

Kőszegi, B. and F. Matějka (2020). Choice Simplification: A Theory of Mental Budgeting and Naive Diversification. *The Quarterly Journal of Economics 135*(2), 1153–1207.

Laes, E., I. Mayeres, N. Renders, P. Valkering, and S. Verbeke (2018). How do policies help to increase the uptake of carbon reduction measures in the EU residential sector? Evidence from recent studies. *Renewable and Sustainable Energy Reviews 94*(June), 234–250.

Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review 73*(1), 31–43.

Liang, F., R. Paulo, G. Molina, M. A. Clyde, and J. O. Berger (2008). Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association 103*(481), 410–423.

Lopez de Leon, F. L. and R. Rizzi (2014). A test for the rational ignorance hypothesis: Evidence from a natural experiment in Brazil. *American Economic Journal: Economic Policy 6*(4), 380–98.

Ludwig, J., J. R. Kling, and S. Mullainathan (2011). Mechanism experiments and policy evaluations. *Journal of Economic Perspectives 25*(3), 17–38.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(3), 531–542.

Met (2010). Exceptional weather events, Type of event: Cold weather, Date: Winter 2009/2010. *The Irish Meterological Service*.

Nyborg, K. (2011). I don't want to hear about it: Rational ignorance among duty-oriented consumers. *Journal of Economic Behavior & Organization 79*(3), 263–274.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

Owens, S. and L. Driffill (2008). How to change attitudes and behaviours in the context of energy. *Energy Policy 36*(12), 4412–4418.

Reiss, P. C. and M. W. White (2005). Household electricity demand, revisited. *The Review of Economic Studies 72*(3), 853–883.

Sexton, S. (2015). Automatic bill payment and salience effects: Evidence from electricity consumption. *Review of Economics and Statistics 97*(2), 229–241.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review 71*(3), 421–436.

Simon, H. A. (1957). *Models of man; social and rational.* Wiley.

Simon, H. A. (1959). Theories of decision-making in economics and behavioral science. *The American Economic Review 49*(3), 253–283.

Simon, H. A. (1978). Rationality as process and as product of thought. *The American Economic Review 68*(2), 1–16.

Simon, H. A. (1979). Information processing models of cognition. *Annual Review of Psychology 30*(1), 363–396.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization 1*(1), 39–60.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science 4*(3), 199–214.

Verhallen, T. M. and W. Fred van Raaij (1986). How consumers trade off behavioural costs and benefits. *European Journal of Marketing 20*(3/4), 19–34.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p-values. *Psychonomic Bulletin & Review 14*(5), 779–804.

Yohanis, Y. G. (2012). Domestic energy use and householders' energy behaviour. *Energy Policy 41*, 654–665.

5 Appendix I: Tables and Figures

Table 15: *Indicated prices for natural gas*

Euro cents / kWh	**Frequency** September 2010	December 2010	Control group
1	0	1	1
2	0	0	4
3	4	3	5
4	4	2	12
5	2	3	7
6	0	0	7
7	1	0	5
8	1	1	4
9	0	0	2
10	1	1	8
11	0	1	3
12	0	2	8
13	2	1	5
14	2	0	9
15	0	0	6
16	0	1	1
17	2	2	5
18	0	0	3
19	0	0	0
20+	13	11	77
N	32	29	172
actual price	3.3	4.6	3.9

This table returns the prices for natural gas in Euro cent per kWh households indicated in the post-trial survey. Columns 2 and 3 are responses from treated households whose prices varied in September and December 2010. The final column returns the answers from the control group. We aggregated all indicated prices above 20 Euro cent per kWh into one group. We indicate the actual price in the final row. N indicates the total number of responses.

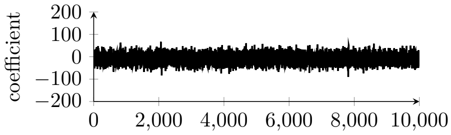
Table 16: *Bayes factor classification (Jeffreys, 1961)*

Bayes factor: BF_{10}	Interpretation
$< 1/100$	Extreme evidence in favor of H_0
$1/100 - 1/30$	Very strong evidence in favor of H_0
$1/30 - 1/10$	Strong evidence in favor of H_0
$1/10 - 1/3$	Substantial evidence in favor of H_0
$1/3 - 1$	Anecdotal evidence in favor of H_0
1	No evidence
$1 - 3$	Anecdotal evidence in favor of H_1
$3 - 10$	Substantial evidence in favor of H_1
$10 - 30$	Strong evidence in favor of H_1
$30 - 100$	Very strong evidence in favor of H_1
> 100	Extreme evidence in favor of H_1

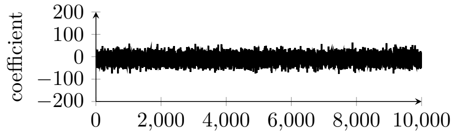
The table was taken verbatim from Wagenmakers (2007), who adapted the labels for the strength of the evidence slightly compared to Jeffreys (1961).

Figure 4: *Trace plots for price treatments*

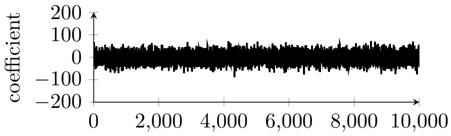
(a) *June 2010*



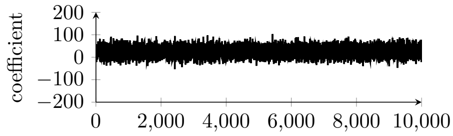
(b) *July 2010*



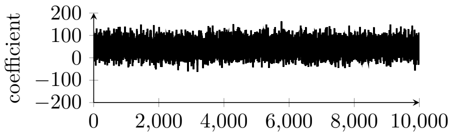
(c) *August 2010*



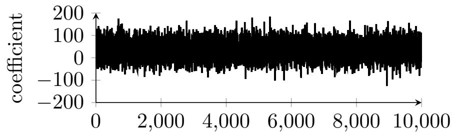
(d) *September 2010*



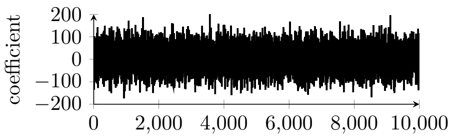
(e) *October 2010*



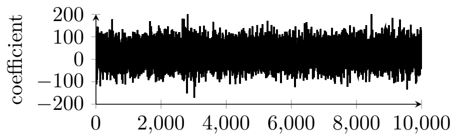
(f) *November 2010*



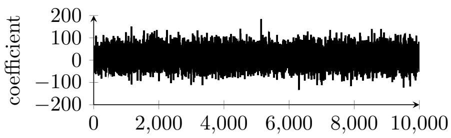
(g) *December 2010*



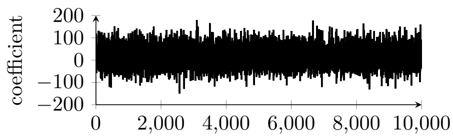
(h) *January 2011*



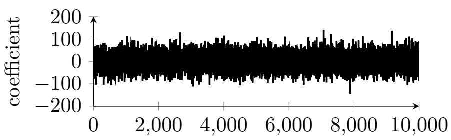
(i) *February 2011*



(j) *March 2011*



(k) *April 2011*



(l) *May 2011*

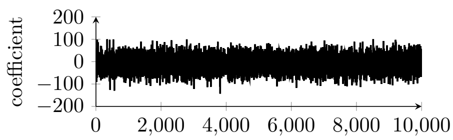


Table 17: *Do household characteristics explain responding to post-trial questions on natural gas prices? (N=426)*

Variable	**Estimates**
1 if gas is used for hot water	0.019 (0.046)
1 if household size: 2 persons	−0.091 (0.058)
1 if household size: 3 persons	−0.003 (0.062)
1 if household size: 4+ persons	0.011 (0.061)
1 if house is semi-detached	−0.004 (0.125)
1 if house is detached	0.005 (0.133)
1 if house is terraced	−0.052 (0.129)
1 if house is bungalow	−0.072 (0.153)
1 if property is owned	0.020 (0.082)
1 if property is owned with mortgage	0.015 (0.081)
1 if house built between 2001 and 2005	0.041 (0.095)
1 if house built between 1981 and 2000	−0.015 (0.091)
1 if house built between 1936 and 1980	−0.010 (0.096)
1 if house built before 1935	0.001 (0.111)
1 if number of bedrooms is 2	−0.129 (0.196)
1 if number of bedrooms is 3	−0.093 (0.194)
1 if number of bedrooms is 4	−0.133 (0.197)
1 if number of bedrooms is 5 +	−0.202 (0.208)
1 if income is between 15001 and 30000 Euro	0.035 (0.091)
1 if income is between 30001 and 50000 Euro	0.089 (0.087)
1 if income is between 50001 and 75000 Euro	0.034 (0.091)
1 if income is over 75000 Euro	0.016 (0.090)
1 if income was not revealed	0.040 (0.085)
1 if household has no alternative heating source	−0.022 (0.036)
1 if share of double glazed windows is up to 0.25	0.113 (0.139)
1 if share of double glazed windows is up to 0.50	0.071 (0.117)
1 if share of double glazed windows is up to 0.75	−0.088 (0.140)
1 if all windows are double glazed	0.074 (0.081)
1 if house has no wall insulation	−0.003 (0.045)
1 if household does not know about wall insulation	−0.027 (0.052)
intercept	0.170 (0.227)

** (*) indicates statistical significance at the 1% (5%) level. Standard errors in parentheses.

Table 18: *Do home improvements differ between treated and control households? (N=346)*

Variable	Estimates
1 if gas is used for hot water	−0.007 (0.077)
1 if household size: 2 persons	0.101 (0.094)
1 if household size: 3 persons	0.106 (0.099)
1 if household size: 4 persons	0.054 (0.097)
1 if house is semi-detached	−0.242 (0.243)
1 if house is detached	−0.324 (0.252)
1 if house terraced	−0.214 (0.248)
1 if house is bungalow	−0.521 (0.287)
1 if property is owned	0.054 (0.141)
1 if property is owned with mortgage	0.088 (0.139)
1 if house built between 2001 and 2005	0.008 (0.161)
1 if house built between 1981 and 2000	0.023 (0.159)
1 if house built between 1936 and 1980	0.101 (0.166)
1 if house built before 1936	0.012 (0.186)
1 if number of bedrooms is 2	0.398 (0.332)
1 if number of bedrooms is 3	0.213 (0.326)
1 if number of bedrooms is 4	0.197 (0.330)
1 if number of bedrooms is 5	0.195 (0.349)
1 if income is between 15001 and 30000 Euro	0.049 (0.152)
1 if income is between 30001 and 50000 Euro	0.160 (0.149)
1 if income is between 50001 and 75000 Euro	0.102 (0.154)
1 if income is over 75000 Euro	0.095 (0.152)
1 if income is not revealed	0.008 (0.145)
1 if household has no alternative heating source	0.060 (0.058)
1 if share of double glazed windows is 0.25	0.117 (0.227)
1 if share of double glazed windows is 0.50	0.175 (0.202)
1 if share of double glazed windows is 0.75	0.002 (0.231)
1 if all windows are double glazed	0.061 (0.137)
1 if house has no wall insulation	−0.008 (0.074)
1 if household does not know about wall insulation	−0.012 (0.084)
1 if household upgraded double glazed windows	0.059 (0.118)
1 if household added insulation	0.136 (0.072)
1 if household added jacket to hot water tank	0.000 (0.084)
1 if household added energy-saving devices	0.069 (0.119)
1 if household added solar panels	−0.542 (0.551)
1 if household weatherized home	−0.105 (0.074)
1 if household serviced boiler	0.034 (0.075)
1 if household added thermostatic controls to heaters	0.022 (0.076)
intercept	0.160 (0.386)

** (*) indicates statistical significance at the 1% (5%) level. Standard errors in parentheses.

Table 19: *Indication of whether individuals understand how much natural gas individual appliances consume*

Category	**Frequency** Control	Treatment	Total
strongly disagree	12	12	24
disagree	12	13	25
neither agree nor disagree	8	15	23
agree	6	5	11
strongly agree	8	12	20

With $p(\chi^2) = 0.753$, there is no evidence to suggest systematic differences between treated and control households.

Table 20: *Differences in pre-trial and trial consumption in kWh*

group	**pre-trial**	**trial**	**Δ (pre-trial - trial)**
control	11,652.0	10,102.3	$1,549.7^{**}$
	(345.0)	(314.1)	(113.5)
treatment	11,587.3	9,956.0	$1,631.3^{**}$
	(347.3)	(296.6)	(145.8)
Δ (control - treatment)	64.7	146.3	-81.6
	(489.6)	(431.8)	(185.2)

This table compares the pre-trial consumption between December 2009 and May 2010 to the trial consumption between December 2010 and May 2011. The final column presents the difference between pre-trial and trial period. The ultimate row presents the difference between control and treatment group. ** (*) indicates statistical significance at the 1% (5%) level.

Table 21: *Do household characteristics explain treatment status?*

Variable	EUS	monthly	IHD	all info	price
observations	653	394	410	665	426
R^2	0.0425	0.0576	0.0766	0.0339	0.0523
1 if gas is used for hot water	0.011	−0.049	0.026	0.044	−0.024
	(0.046)	(0.066)	(0.066)	(0.046)	(0.066)
1 if household size: 2 persons	0.071	−0.094	−0.129	−0.022	0.114
	(0.059)	(0.087)	(0.083)	(0.060)	(0.084)
1 if household size: 3 persons	0.024	0.025	−0.028	0.020	0.125
	(0.068)	(0.097)	(0.092)	(0.066)	(0.089)
1 if household size: 4+ persons	0.017	0.009	−0.027	0.011	0.080
	(0.066)	(0.094)	(0.088)	(0.065)	(0.088)
1 if house is semi-detached	0.070	0.236	0.069	0.091	−0.139
	(0.141)	(0.279)	(0.223)	(0.142)	(0.180)
1 if house is detached	0.124	0.196	−0.008	0.080	−0.195
	(0.147)	(0.284)	(0.228)	(0.149)	(0.191)
1 if house is terraced	0.109	0.245	0.020	0.095	−0.158
	(0.142)	(0.280)	(0.226)	(0.144)	(0.185)
1 if house is bungalow	0.173	0.188	0.044	0.186	−0.387
	(0.161)	(0.293)	(0.248)	(0.164)	(0.221)
1 if type of house was not revealed	0.186			0.087	
	(0.689)			(0.699)	
1 if property is owned	−0.020	0.143	0.101	0.048	0.090
	(0.082)	(0.128)	(0.114)	(0.083)	(0.119)
1 if property is owned with mortgage	0.039	0.085	0.005	0.033	0.106
	(0.080)	(0.125)	(0.110)	(0.081)	(0.116)
1 if property status was not revealed	−0.256	0.840		−0.311	
	(0.342)	(0.544)		(0.347)	
1 if house built between 2001 and 2005	−0.054	0.032	−0.152	−0.248	0.129
	(0.131)	(0.175)	(0.144)	(0.120)	(0.137)
1 if house built between 1981 and 2000	−0.120	0.129	−0.062	−0.224	0.053
	(0.127)	(0.165)	(0.138)	(0.116)	(0.131)
1 if house built between 1936 and 1980	−0.043	−0.019	−0.274	−0.331**	0.148
	(0.128)	(0.167)	(0.142)	(0.119)	(0.138)
1 if house built before 1935	−0.089	−0.177	−0.216	−0.342**	0.079
	(0.139)	(0.194)	(0.165)	(0.129)	(0.160)
1 if number of bedrooms is 2	−0.080	0.074	0.071	0.032	0.088
	(0.202)	(0.335)	(0.297)	(0.208)	(0.283)
1 if number of bedrooms is 3	−0.153	0.073	0.178	0.030	−0.094
	(0.201)	(0.327)	(0.291)	(0.209)	(0.280)
1 if number of bedrooms is 4	−0.079	0.016	0.118	0.048	−0.099
	(0.206)	(0.335)	(0.297)	(0.213)	(0.283)
1 if number of bedrooms is 5 +	0.030	−0.031	0.041	0.101	−0.103
	(0.222)	(0.350)	(0.310)	(0.228)	(0.300)
1 if number of bedrooms was not revealed	−0.406			−0.435	
	(0.515)			(0.524)	
1 if income is between 15001 and 30000 Euro	0.177	−0.078	−0.182	0.061	0.128
	(0.094)	(0.147)	(0.134)	(0.092)	(0.131)
1 if income is between 30001 and 50000 Euro	0.129	−0.112	−0.228	−0.058	0.193
	(0.090)	(0.145)	(0.130)	(0.087)	(0.125)
1 if income is between 50001 and 75000 Euro	0.120	−0.142	−0.195	−0.036	0.170
	(0.095)	(0.151)	(0.134)	(0.091)	(0.131)
1 if income is over 75000 Euro	0.110	−0.104	−0.155	−0.013	0.144
	(0.095)	(0.152)	(0.135)	(0.091)	(0.130)
1 if income was not revealed	0.064	−0.092	−0.075	0.010	0.092
	(0.089)	(0.146)	(0.127)	(0.084)	(0.123)
1 if no alternative heating source	0.000	−0.074	−0.020	−0.039	0.048
	(0.038)	(0.053)	(0.051)	(0.038)	(0.051)
1 if share of double glazed windows is 0.25	−0.046	−0.056	0.202	0.131	0.014
	(0.155)	(0.259)	(0.211)	(0.151)	(0.199)
1 if share of double glazed windows is 0.50	−0.037	−0.103	0.103	0.023	0.132
	(0.116)	(0.195)	(0.173)	(0.119)	(0.168)
1 if share of double glazed windows is 0.75	0.113	0.059	−0.038	0.136	−0.009
	(0.135)	(0.178)	(0.181)	(0.144)	(0.202)
1 if all windows are double glazed	−0.024	0.058	0.060	0.037	0.033
	(0.076)	(0.117)	(0.112)	(0.079)	(0.116)
1 if house has no wall insulation	0.015	0.046	0.014	0.022	0.030
	(0.046)	(0.067)	(0.063)	(0.047)	(0.064)
1 if wall insulation not revealed	0.021	−0.030	−0.021	−0.012	−0.004
	(0.054)	(0.079)	(0.075)	(0.055)	(0.075)
intercept	0.245	0.246	0.609	0.377	0.274
	(0.257)	(0.418)	(0.324)	(0.237)	(0.327)

** (*) indicates statistical significance at the 1% (5%) level. Standard errors in parentheses. Blank cells indicate empty categories.

Table 22: *Do household characteristics explain attrition? (N=1,342)*

Variable	**Estimates** $R^2 = 0.0297$
1 if gas is used for hot water	-0.004 (0.016)
1 if household size: 2 persons	0.003 (0.022)
1 if household size: 3 persons	0.001 (0.023)
1 if household size: 4+ persons	0.001 (0.022)
1 if house is semi-detached	-0.074 (0.065)
1 if house is detached	-0.045 (0.067)
1 if house is terraced	-0.029 (0.067)
1 if house is bungalow	-0.010 (0.074)
1 if answer to housetype was refused	0.028 (0.089)
1 if property is owned	0.011 (0.031)
1 if property is owned with mortgage	0.004 (0.029)
1 if property status was not revealed	0.158 (0.221)
1 if house built between 2001 and 2005	-0.051 (0.048)
1 if house built between 1981 and 2000	-0.067 (0.046)
1 if house built between 1936 and 1980	-0.050 (0.048)
1 if house built before 1935	-0.043 (0.054)
1 if number of bedrooms is 2	-0.042 (0.104)
1 if number of bedrooms is 3	-0.057 (0.105)
1 if number of bedrooms is 4	-0.067 (0.105)
1 if number of bedrooms is 5 +	-0.068 (0.107)
1 if number of bedrooms was not revealed	-0.157 (0.109)
1 if income is between 15001 and 30000 Euro	-0.009 (0.032)
1 if income is between 30001 and 50000 Euro	-0.004 (0.033)
1 if income is between 50001 and 75000 Euro	-0.010 (0.032)
1 if income is over 75000 Euro	-0.019 (0.032)
1 if income was not revealed	0.019 (0.032)
1 if household has no alternative heating source	-0.008 (0.012)
1 if share of double glazed windows is up to 0.25	-0.033 (0.050)
1 if share of double glazed windows is up to 0.50	-0.019 (0.038)
1 if share of double glazed windows is up to 0.75	0.055 (0.057)
1 if all windows are double glazed	0.007 (0.027)
1 if house has no wall insulation	-0.006 (0.016)
1 if household does not know about wall insulation	0.012 (0.019)
intercept	0.215 (0.113)

** (*) indicates statistical significance at the 1% (5%) level.Standard errors in parentheses.

Table 23: *Unconditional quantile treatment effects (N=426)*

Month	**Percentile**		
	25th	50th	75th
June 2010	8.798	6.711	−38.386
	(15.608)	(24.184)	(33.603)
July 2010	−2.062	−9.422	−36.382
	(13.280)	(22.167)	(30.723)
August 2010	−13.045	−1.406	−18.354
	(15.242)	(23.784)	(36.684)
September 2010	17.499	69.520	48.366
	(23.874)	(26.082)	(37.161)
October 2010	70.851	56.444	53.481
	(49.318)	(48.051)	(62.971)
November 2010	61.836	50.017	145.677
	(80.140)	(76.765)	(90.578)
December 2010	60.917	36.589	98.403
	(119.698)	(112.643)	(140.007)
January 2021	19.349	25.301	116.775
	(98.974)	(92.161)	(118.451)
February 2021	57.776	−41.664	75.468
	(70.508)	(68.426)	(83.024)
March 2021	−13.598	−25.423	135.200
	(65.925)	(67.111)	(86.307)
April 2021	13.081	−20.655	29.964
	(38.025)	(43.809)	(61.460)
May 2021	42.319	27.629	−22.564
	(37.814)	(41.332)	(63.380)

** (*) indicates statistical significance at the 1% (5%) level. Standard errors in parentheses.

Working Papers of the Center of Economic Research at ETH Zurich

(PDF-files of the Working Papers can be downloaded at www.cer.ethz.ch/research/working-papers.html).

21/348 N. Ritter, J. A. Bingler
Do homo sapiens know their prices? Insights on dysfunctional price mechanisms from a large fi

eld experiment

20/347 C. Daminato, M. Filippini, F. Haufler
Personalized Digital Information and Tax-favoured Retirement Savings: Quasi-experimental Evidence from Administrative Data

20/346 V. Britz, H. Gersbach
Open Rule Legislative Bargaining

20/345 A. Brausmann, E. Grieg
Resource Discoveries and the Political Survival of Dictators

20/344 A. Jo
The Elasticity of Substitution between Clean and Dirty Energy with Technological Bias

20/343 I. van den Bijgaart, D. Cerruti
The effect of information on market activity; evidence from vehicle recalls

20/342 H. Gersbach, R. Wattenhofer
A Minting Mold for the eFranc: A Policy Paper

20/341 L. Bretschger
Getting the Costs of Environmental Protection Right

20/340 J. A. Bingler, C. Colesanti Senni
Taming the Green Swan: How to improve climate-related fi

nancial risk assessments

20/339 M. Arvaniti, T. Sjögren
Temptation in Consumption and Optimal Redistributive Taxation

20/338 M. Filippini, S. Srinivasan
Voluntary adoption of environmental standards and limited attention: Evidence from the food and beverage industry in Vietnam

20/337 F. Böser, C. Colesanti Senni
Emission-based Interest Rates and the Transition to a Low-carbon Economy

20/336 L. Bretschger, E. Grieg, P. J.J. Welfens, T. Xiong
Corona Fatality Development and the Environment: Empirical Evidence for OECD Countries

20/335 M. Arvaniti, W. Habla
The Political Economy of Negotiating International Carbon Markets

20/334 N. Boogen, C. Daminato, M. Filippini, A. Obrist
Can Information about Energy Costs Affect Consumers Choices? Evidence from a Field Experiment

20/333 M. Filippini, N. Kumar, S. Srinivasan
Nudging the Adoption of Fuel-Efficient Vehicles: Evidence from a Stated Choice Experiment in Nepal

20/332 L. Bretschger, E. Grieg
Exiting the fossil world: The effects of fuel taxation in the UK

20/331 H. Gersbach, E. Komarov
Research Bubbles

20/330 E. V. Dioikitopoulos, C. Karydas
Sustainability traps: patience and innovation

19/329 M. Arvaniti, C. K. Krishnamurthy, A. Crepin
Time-consistent resource management with regime shifts

19/328 L. Bretschger, K. Pittel
Twenty Key Questions in Environmental and Resource Economics

19/327 C. Karydas, A. Xepapadeas
Climate change financial risks: pricing and portfolio allocation

19/326 M. Filippini, S. Srinivasan
Investments in Worker Health and Labor Productivity: Evidence from Vietnam

19/325 H. Gersbach
Democratizing Tech Giants! A Roadmap

19/324 A. Brausmann, M. Flubacher and F. Lechthaler
Valuing meteorological services in resource-constrained settings: Application to smallholder farmers in the Peruvian Altiplano

19/323 C. Devaux and J. Nicolai
Designing an EU Ship Recycling Licence: A Roadmap

19/322 H. Gersbach
Flexible Majority Rules for Cryptocurrency Issuance

19/321 K. Gillingham, S. Houde and A. van Benthem
Consumer Myopia in Vehicle Purchases: Evidence from a Natural Experiment

19/320 L. Bretschger
Malthus in the Light of Climate Change

19/319 J. Ing and J. Nicolai
Dirty versus Clean Firms' Relocation under International Trade and Imperfect Competition

19/318 J. Ing and J. Nicolai
North-South diffusion of climate-mitigation technologies: The crowding-out effect on relocation

19/317 J. Abrell, M. Kosch and S. Rausch
How Effective Was the UK Carbon Tax? - A Machine Learning Approach to Policy Evaluation

19/316 D. Cerruti, C. Daminato and M. Filippini
The Impact of Policy Awareness: Evidence from Vehicle Choices Response to Fiscal Incentives

19/315 M. Filippini, N. Kumar and S. Srinivasan
Energy-related financial literacy and bounded rationality in appliance replacement attitudes: Evidence from Nepal

19/314 S. Houde and E. Myers
Heterogeneous (Mis-) Perceptions of Energy Costs: Implications for Measurement and Policy Design

19/313 S. Houde and E. Myers
Are Consumers Attentive to Local Energy Costs? Evidence from the Appliance Market

19/312 N. Kumar
A model-based clustering approach for analyzing energy-related financial literacy and its determinants

19/311 C. Karydas and A. Xepapadeas
Pricing climate change risks: CAPM with rare disasters and stochastic probabilities

19/310 J.Abrell,S. Rausch and C. Streitberger
Buffering Volatility: Storage Investments and Technology-Specific Renewable Energy Support

19/309 V. Britz
Negotiating with frictions

19/308 H. Gersbach and S. Papageorgiou
On Banking Regulation and Lobbying