Energy efficiency for the poor: Varying subsidies and procedures in a refrigerator replacement program*

Bettina Chlond[1,2], Timo Goeschl[1,2], and Martin Kesternich[2,3]

[1]Alfred Weber Institute for Economics, Heidelberg University
[2]ZEW – Leibniz Centre for European Economic Research
[3]Institute of Economics, University of Kassel

May 6, 2022

Abstract

We study the effects of varying the subsidy level and program procedures on the performance of a national refrigerator replacement program targeted at improving energy efficiency in low-income households in Germany. Successful participation for eligible households, screened via a home energy audit, has two stages: an enrolment stage that leads to a replacement voucher and the voucher redemption stage following replacement. Exploiting two exogenous temporal discontinuities in voucher value (subsidy), enrolment and redemption procedures, we examine the replacement decisions of 77,000 eligible households. Increasing the subsidy of 50 percent raises the rate of replacement among eligible household by 9 to 16 percentage points. Changing procedures – from automatic to elective enrolment and from a three-month renewable to a two-month non-renewable voucher – raises the success rate by 4 to 10 percentage points. We conclude that low-cost changes in procedures that target the behavioral responses of low-income households represent unexploited economies in program design.

Keywords: Energy efficiency, low-income households, durable replacement, energy poverty, technology adoption. **JEL classification:** C25, D15, H23, O33, Q20

*This research has been funded by the Fritz Thyssen Foundation. The authors thank Peter Heindl for his support in the initiation of this cooperation. Correspondence: bettina.chlond@awi.uni-heidelberg.de.

1 Introduction

The impact of high energy prices on low-income households is a major concern among policy-makers. Inelastic demand for energy and restricted access to consumer credit imply that low-income households are disproportionately affected unless effective assistance programs help these households reduce their energy bills through further behavioral adjustments and through accelerated adoption of energy-efficient technologies in consumer durables.

The setting of our paper is a nation-wide Refrigerator Replacement Program (RRP) targeting low-income households that subsidizes the modernization of household refrigeration appliances and has been operational in Germany since 2009. There, refrigerators are the consumer durable that accounts for the largest share (about 25%) of household electricity consumption [BDEW, 2022] [1]. The RRP is embedded in a larger initiative called "Electricity Savings Check" (SSC) being funded by German Federal Ministry for the Environment. Between 2009 and 2020, the 150 local branches of the SSC actively recruited more than 360,000 low-income households through a variety of channels and conducted energy audits in their homes to help them reduce energy and water consumption. Appliance inventory data collected as part of the SSC home energy audit are used to screen for eligibility for the RRP. Three criteria determine eligibility: Being a recipient of at least one of several federal income support schemes; the age of the refrigerator ($>$ 10 years); and expected annual savings from refrigerator replacement of at least 200kWh.[2] In the first twelve years of the RRP's existence, the screening identified 77,305 eligible households. These households are then actively targeted for enrolment into the RRP in a follow-up visit by the team of SSC advisors. Enrolled households receive a voucher that is redeemed in cash upon successful refrigerator replacement by the household. On average, 26% of eligible households take up the voucher-based subsidy to replace energy-inefficient refrigerators.[3] With an electricity prices of €0.289 in

[1] This contrasts with the US experience where air conditioners are the most energy-intensive home durable accounting for 12% of total home energy expenditures in 2015 [EIA, 2015]. In Germany, AC units are yet relatively rare.

[2] This translate into minimum savings of approximately €70-99 per year. For more details on the RRP, see section 2.

[3] To put this into perspective, take-up of financial incentives among low-income U.S. households for energy efficiency improvements such as building weatherization is minimal, even when the gains of doing so are high [Fowlie et al., 2015, 2018, Hancevic and Sandoval, 2022]. Comparable evidence on appliance replacement programs is only available for episodic campaigns directed at

2020 and average annual savings of 342 kWh, successful refrigerator replacement has led to annual savings in electricity bills of €99.

This paper contributes to a growing literature in behavioral public economics that examines how program design affects program performance, in particular among vulnerable groups in society. "Program design" here refers to the totality of features of a policy, from budget-relevant economic incentives to purely situational aspects [Bertrand et al., 2004]. The literature on behavioral public policy has demonstrated that seemingly inconsequential design variations can lead to sizeable changes in program performance, highlighting the importance of psychological co-determinants of program success. For example, Bertrand et al. [2006] discuss how variations in procedural hassle are likely to affect take-up of food stamp programs among eligible households, a key performance metric of anti-poverty programs. Hastings and Weinstein [2008] present evidence from a natural experiment to show that varying information provision to low-income households improves academic outcomes for disadvantaged children, the key performance metric of so-called "choice plans" in education. In a field experiment in partnership with the U.S. Internal Revenue Service, Bhargava and Manoli [2015] demonstrate how variations in tax mailings to low-income households can improve take-up of means-tested cash transfers, the key performance metric of the Earned Income Tax Credit. Deshpande and Li [2019] show that closing Social Security Administration field offices significantly changes the volume and composition of disability benefits recipients and applicants. Against this background, our present study both examines the impact of varying subsidies and of varying procedures on the probability that an eligible household successfully replaces their refrigerator. This probability, referred to as the "replacement rate", is the key performance metric of the RRP, not least because of the considerable cost of each home energy audit to the program.

On account of our setting, we are able to examine a number of aspects that can enrich the literature on program design for low-income households. First, our study exploits the

the general population: A 36-month campaign in Mexico between 2009 and 2012 led to 1.9 million appliance replacements, among them 1.7 million refrigerators. This corresponds to a take-up rate of around 17 percent [Davis et al., 2014]. A much shorter similar campaign in the US (duration between 1 and 91 weeks, 26 weeks on average) led to 1.8 million appliance replacements, among them 631,561 refrigerators, under much less stringent eligibility criteria [Houde and Aldy, 2017]. There is no estimated take-up rate reported.

fact that after being scaled up to its present size in 2013, the RRP experienced two quasi-exogenous shocks that changed different dimensions of the program design unexpectedly and at short notice. These shocks mean that we observe the RRP in three distinct regimes, one until December 2017, a second from January 2018 to March 2019, and a third from April 2019 onwards. Proceeding conservatively and making use of the rolling nature of the program, we make the case that much of the change in replacement rates across the three regimes can be attributed to changes in program design. This attribution relies on a tailored regression discontinuity (RDD) framework that takes into account transitional periods in the RRP as well as seasonality effects and location-specific factors. The RRP therefore constitutes a particularly rich, but also statistically favorable setting in which to explore the comparative performance of design variations on program performance.

Second, our study exploits the fact that the design dimensions changed by each of the policy shocks were almost orthogonal. One shock changed the level of the cash subsidy that households receive upon replacing their appliance: Since the start of the RRP, the cash value of the voucher had always been set at €150, accounting for 35 to 45 percent of the purchase price of the new refrigerator. From April 1, 2019, the subsidy fell to €100, a reduction by one third. The paper can therefore speak to the effects of large relative changes in financial incentives on program performance among low-income households.[4]. The other shock changed program procedures: Since the start of the RRP, enrolment had always been automatic. Every eligible households was enrolled and received the voucher by default. From January 2018, enrolment became elective: Eligible households had to actively enrol by requesting the voucher after the second visit from their local branch. At the same time, the voucher terms changed. Terms had always been flexible: The voucher was valid for three months at a time and repeatedly renewable. From January 2018, voucher terms became rigid: The voucher was valid for two months and not renewable. The paper can therefore speak to the effects of procedural changes on program performance among low-income households, furthering our understanding of how "psychological frictions" [Bhargava and Manoli, 2015] and "hassle" [Bertrand et al., 2006] affect program uptake. This includes new evi-

[4] To our knowledge, empirical evidence on such effects is surprisingly scarce, with the exception of the effects of social benefits on labor supply [Ellwood, 2000]

dence on the effect of deadlines on program performance [Bertrand et al., 2010, Shu and Gneezy, 2010, Altmann et al., 2021], for which – to our knowledge – no specific evidence for low-income households has been available so far. Importantly, we are able to benchmark the effect of these procedural changes against the variation in cash subsidies, providing an intuitive metric of comparison. Third, this paper extends the range of economic decisions taken by low-income household beyond the typical consumption or income support programs towards their investment decisions. In particular, it adds value to the scarce literature on investment decisions in energy efficiency among low income households [Fowlie et al., 2015, 2018], by presenting the – to our knowledge – first evidence that can explicitly speak to the impact of program design variations on investments in energy efficient appliances. The investment decision at the heart of the RRP constitutes a particularly challenging problem that all owners of energy-intensive consumer durables who pay their own electricity bills have to solve [Rapson, 2014, Wang and Matsumoto, 2021]: Due to wear and tear in use, consumer durables become less energy-efficient over time while increasingly energy-efficient devices become available and affordable on the market due to technological progress. Both dynamics play out against a background of short- and long-term changes in electricity prices, further complicating the decision. Compared to high-income groups, low-income households have most to gain from getting the replacement timing right because a larger share of their income is exposed to the cost of energy. At the same time, they are at particular risk of mis-timing: The cognitive challenges of optimal replacement timing accentuate lower financial sophistication, leading to errors in decision-making [Calvet et al., 2009]. Low-income households are also forced, as a result of being poor, to devote a greater share of their cognitive resources to psychologically salient short-term problems [Shah et al., 2012, Mani et al., 2013]: This makes it likelier that households overlook longer-term problem and miss optimal replacement points in consumer durables. For German low-income households, mis-timing is particularly costly: At €0.36 per kWh Germany has some of the highest retail prices for electricity in the world,[5] and German low-income households tend to face higher retail prices for electricity

[5] Consumer electricity prices have doubled since 2002. In March 2022, wholesale prices peaked at a new all-time high. In consequence, some providers started to charge prices of more than 70 cents per kWh.

than the average household.[6] Despite their exposure, low-income households invest less in energy-efficiency consumer durables [Ameli and Brandt, 2015, Schleich, 2019] and are less responsive to public energy-efficiency programs than the average household [Allcott, 2011, Gillingham and Tsvetanov, 2018]. A further aggravating factor for optimal replacement is the annual billing cycle by German electricity suppliers: Households learn about their electricity consumption only with significant delay and with little hope of being able to attribute the annual total to specific appliances, such as refrigerators, or consumption episodes, such as hot weather periods. Taken together, these particular challenges make appliance replacement decisions particularly interesting among the class of economic decisions for which public policies provide support to low-income households.

On the basis of 12 years of RRP data on home energy audits, program enrolment, and voucher redemption in three distinct program regimes, we have three main results on how subsidy and procedural variations in the RRP affected replacement rates among eligible low-income households. First, we find that a 50% percent higher subsidy is associated with a likelihood of refrigerator replacement that is 9 to 16 percentage points higher. We believe this is the first evidence on the "subsidy elasticity" of consumer durables replacement among low-income households, and the elasticity is substantial. The estimates underscore both the presence of an effect of economic incentives on RRP performance and its significant scale: Program performance is demonstrably a question of subsidy levels. The data also provide insights into the nature of the subsidy elasticity, which is important to researchers and program administrators. That nature is that the elasticity operates only at the enrolment stage, but not at the redemption stage: Higher-value vouchers make more households enrol, but higher-value vouchers are not redeemed more frequently.

Second, we find that the procedural changes in the RRP cause replacement rates to rise by 4 to 10 percentage points. The direction of this effect is as interesting as its magnitude and dynamics. At the enrolment stage, the share of enrolled households drops from 100% under automatic to just under 40% under elective enrolment. Through the lens of the behavioral economics of assistance programs, the size of this decrease could be attributed to the change

[6] Using household data for Germany, Frondel et al. [2019] show that the energy price elasticity for low-income households is low. Andor et al. [2021] report that low-income households are on average less efficient in their electricity use per square meter than wealthy households.

in the default [Thaler and Sunstein, 2021] and to the procedural “hassle” imposed on eligible households [Bertrand et al., 2006]. At the same time, electively enrolled households exhibit – under the two-month deadline – vigorous program take-up. Compared to automatically enrolled households, a greater share of enrolled households replaces their refrigerator, and they replace more quickly following the second visit. Selection effects trivially explain some of the intensive-margin difference, but are insufficient for explaining why cumulative replacement rates after the procedural change dominate those before the change for every point in time following the second home visit. Through the lens of behavioral economics, several factors could be at play, such as deliberate ‘opt-ins’ facilitating effective goal-setting towards replacement [Locke and Latham, 1990] and the presence of a rigid deadline helping households to overcome time management problems [Bertrand et al., 2006].[7] Jointly, they lead to an intensive-margin effect that more than compensates for the changes in the enrolment mechanism.

Our third result comes from comparing the effects of varying subsidies and varying procedures and using them to conduct back-of-the-envelope calculations of the merits of alternative program design. Such a comparison yields that procedural changes that added little to no cost to the program[8] generated an improvement in replacement rates that was equivalent to an estimated subsidy increase of €15 to €56 per replacing household. This estimate represents for social assistance programs the first estimate of the monetary equivalent of changes in program procedures, complementing similar estimates for loan marketing targeting the general population [Bertrand et al., 2010]. These estimates support our conservative assessment that implementing these procedural changes from 2013 rather than 2018 would have realized 1,900 additional replacements by low-income households at the same budgetary cost, leading to additional savings in electricity bills of €187,800.

We proceed as follows: In the following section 2 we provide the necessary background on the Refrigerator Replacement Program. In section 3, we explain the data on which the analysis is based. Section 4 lays out the empirical challenges and the empirical strategy. In

7 The impact of deadlines is far from clear: Bertrand et al. [2010] find a negative effect of deadlines on loan take-up among general-population households in South Africa. Shu and Gneezy [2010] and Altmann et al. [2021], on the other hand, find positive effects.

8 The impact on costs could plausibly even be negative due to reductions in administrative work load.

section 5, we present the main effects of the variation in the subsidy levels and the procedures on the success rate of the RRP. We then discuss the underlying mechanisms in Section 6. In section 7 we estimate the effects of the alternative, untried regime. Section 8 concludes.

2 The Refrigerator Replacement Program

Since 2009, the Refrigerator Replacement Program (RRP; German: *Kühlgeräte - Tauschprogramm*) has been offering cash vouchers to households on federal income support[9] in order to encourage replacing their old and inefficient refrigeration devices with modern, highly efficient models. The Program is embedded within a wider initiative, the "Electricity-Saving Check-up" (SSC, German: *Stromspar-Check*) that provides general support to low-income households for reducing their energy and water consumption by conducting home energy audits. These 'SSC households' constitute the pool from which the RPP draws its population. RRP and SSC are implemented jointly by the German Caritas Association, one of the largest social welfare organizations in the country, and the Association of Energy and Climate Protection Agencies (eaD). Caritas and eaD operate around 150 local branches throughout the country. Annual funding of around €10-15 million is provided by the German Federal Ministry for the Environment on the basis of program grants with a funding cycle of three years, subject to successful (re-)application by the implementing agencies. The RRP started on January 1, 2009 and was scaled up to its current size with the start of its second funding cycle of the SSC ("SSC plus") in April 2013. Starting April 1, 2022, the SSC will be in its fifth funding cycle, which lasts until 2025 ("SSC close-by", German: *Stromspar-Check in Ihrer Nähe*) (see Figure 1 for an overview of the funding cycles).

Recruitment of qualified households into the SSC's home energy audits takes place through a variety of channels. SCC and RRP are actively promoted in many employment and social assistance agencies across the country through printed and audiovisual material. They are

[9] To qualify, the household needs to receive at least one type of federal income support such as unemployment benefits ("Arbeitslosengeld II"), housing allowances ("Wohngeld", "Sozialhilfe"), low pensions ("Grundsicherung"), child supplements ("Kinderzuschlag") or benefits for asylum seekers ("Leistungen nach Asylbewerberleistungsgesetz"), or the household's income must be below the income limit for attachment. Around xx% of German households qualify on this basis (data for 2019?).

also present with pop-up booths in shopping streets and malls, with active staffers providing individualized education about the program. Some local branches of the social assistance agency mandate the participation of households with excessively high energy bills. The SCC also maintains a website where information is available about the RRP in eleven languages. Additionally, recruitment takes place directly through one of the local branches. The program has no systematic understanding of how its different channels contribute to overall recruitment, but since 2009, more than 360,000 households have participated in the SSC initiative and undergone, free of charge, a home energy audit by staff employed by one of the local branches.

The typical home energy audit of the SSC consists of two visits to the household by a two-person team within a period of around three weeks. During the first visit, the "energy advisors" make an inventory of all electric devices and their usage in the household, assess the electricity consumption of refrigerators and freezers, and educate the household on electricity-saving behavior. The inventory and electricity consumption assessment are used to screen for eligibility of the household for the RRP. The screening leads to differences in the second visit. Both eligible and non-eligible households receive approximately €70 worth of energy-saving kit such as LED light bulbs, switchable socket strips, TV standby cut-off switches, timers and water flow regulators. These items are directly installed by the two advisors. Non-eligible households then exit the SSC initiative.*A fraction of households are re-contacted one year later to assess savings.* For eligible households, the second visit contains an additional component in which they are specifically targeted for enrolment in the RRP through educational material and promotion.[10]

The rationale for enrolling households in the RRP is the large contribution, roughly 25 percent [BDEW, 2019], that refrigerators make to the electricity consumption of the average German household.[11] Differences in refrigerator efficiency can therefore impact significantly on domestic electricity bills. To be eligible for enrolment, the low-income household has to own a refrigerator older than 10 years and be expected to save at least 200 kWh annually from

[10] Only households that completed the first visit of the home energy audit can become eligible for the RRP.

[11] We use "refrigerator" to refer to both refrigerators, freezers, and combination units within the program.

a replacement with the most energy-efficient class of devices on the market.[12] The expected savings are communicated to the household in writing during the second visit. Under the terms of the RRP, enrolled households can redeem their voucher for cash only after meeting a number of criteria. They need to present the purchase receipt; document that the purchased device is of energy efficiency class A+++; document that it is not larger than the original refrigerator; and provide proof that the original refrigerator has entered the recycling chain. Households have to handle all steps of the refrigerator replacement on their own, including identifying and selecting a model that fulfils the requirements, pre-financing the purchase, and organizing the logistics of delivering the new and of disposing of the old refrigerator.

The RRP is the only federal voucher scheme for replacing refrigerators in low-income households. At the same time, complementary programs exist in at least four of the sixteen states (*Länder*) and in a number of municipalities.[13] This coexistence of programs is one feature of the policy landscape that requires an appropriate empirical strategy. Another feature of the policy landscape are expected and unexpected program changes at the federal level. Expected changes in the RRP occurred at the end and beginning of each of funding cycle: Vouchers are cycle-specific and do not carry over from one funding cycle to the next. As one cycle ends, staff at local branches increase their efforts to encourage enrolled households to redeem their vouchers during the final months of the program. At the same time, enrolment activities cease in the final two to three months before being ramped up again at the beginning of the new cycle.

There were also two unexpected changes in the RRP, one on January 1, 2018 and one on April 1, 2019. The first change, within the third funding period of the SSC, simultaneously affected specific procedures of the program, nameley the enrolment mode of the RRP and

[12] The savings expectations are based on engineering estimates: Based on the inventory data from the first visit, SSC staff use a custom database to calculate expected savings based on a comparison between the current device and a reference device of equivalent size and features that fulfills the A+++ standard, the most efficient class of devices on the EU scale in force between 2009 to 2021. Since March 2021, a revised EU scale has been in force that puts devices previously rated as A+++ in the classes B and C. Transitional arrangements are in place both in the retail sector and in the RRP.

[13] At the level of the federal states, Berlin offers a complementary subsidy of €50 since December 2020, Saxony-Anhalt of €75 since May 2020, and Hamburg of €100 since September 2010. North Rhine-Westphalia complements the federal subsidy with an additional €50 per person (up to €200 per household and up to the purchasing price less €50) since July 2016.

the terms of the voucher. The enrolment mode switched from automatic enrolment until the end of 2017 to elective enrolment from 2018 onwards. Under automatic enrolment, all eligible households received the RRP voucher on the second visit. Under elective enrolment, eligible households have been receiving on the second visit an *invitation to claim* a voucher from the local branch before purchasing a new refrigerator. Enrolment hence requires households to take an active step. In addition, the terms of the voucher changed at the same time: Until the end of 2017, the voucher handed out to all eligible households was valid for three months and renewable for additional periods of three months upon request. From 2018 onwards, the voucher has been valid for two months, without the option to renew. The reason for the change from a flexible three-month renewable to a rigid two-month non-renewable terms in January 2018 was the discovery in late 2017 that a combination of an automatic enrolment mode and an implicit right for voucher renewal had left the RRP open to possible oversubscription and a resulting budget shortfall as the funding cycle approached its end in March 2019. As a result of this discovery, the implementing agencies resolved, at short notice, to alter the enrolment mode and voucher terms as an 'emergency brake'.

The second unexpected change, when turning from the third to the fourth funding cycle on April 1, 2019, affected the value of the voucher. Since the start of the RRP in 2009, vouchers had always been worth €150 to a redeeming household. The implementing agencies' 2018 application for the fourth funding cycle starting 2019 foresaw the same voucher value. Instead, the Federal Ministry's funding approval at the end of 2018 cut its support to €100 per replaced refrigerator, the first such change in the history of the RRP. This

Taken together, the four funding cycles so far constitute a twelve-year history of experience with appliance replacement through a cash voucher scheme.

3 Data

Our data includes more than 360,000 households who participated in an SSC audit between January 2009 and December 2020 (repeated cross-section). Of these, about 77,000 households were eligible for a subsidized refrigerator replacement, the sample of interest for our analysis. About 20,000 households actually replaced their refrigerator. The share of eligible households that successfully participated in the replacement program is therefore around 26% (see Table 1: Program variables). This statistic is important: It implies that for three out of four low-income households owning an old and inefficient refrigerator, the efforts of the RRP do not lead to subsidized replacement. At the level of the household, this means a continuation of paying high electricity bills. At the program level, it means that for one successful replacement, the RRP has to bear the costs of screening and enrolling four households. It also means bearing the costs of issuing and administrating thousands of vouchers that go unused.

For each eligible household, the dataset contains demographic information, such as the number of persons in the household, the type of federal income support received, living space and the state and ZIP code of residence. Documentation from the audit includes the date of the first and second visit, the local branch that administered the audit, the auditors who conducted the first and the second visit, the annual electricity consumption of the household and the price paid per kWh. For the refrigerator replacement program, status of eligibility, enrolment (i.e. voucher request) and voucher redemption after refrigerator replacement is available. Moreover, the dataset contains information on the old refrigerators in the household, such as age, measured kWh consumption and volume. Finally, the data contains information on the newly purchased refrigerator, including the purchasing price, volume and kWh consumption as specified by the manufacturer.

Table 1 presents descriptive statistics for household and old refrigerator characteristics. On average, households eligible for subsidized refrigerator replacement consist of 2.8 household members which live on 69 square meter.[14] Their refrigerators and freezers have an average age of 17.3 years, a capacity of 239 liters and consume around 480 kWh annually. For comparison, a state-of-the-art large A+++ combined refrigerator-freezer consumes around 200 kWh annually. The difference of 280 kWh per year, equivalent to around €84, illustrates the energy efficiency gap present in eligible households.

Of the eligible households, 35% live together in families with at least one child in the household; more than a third of these families have more than two kids. 29% in the sample

[14] An average German household consists of 2.03 members (Mikrozensus – Haushalte und Familien 2020/Statistisches Bundesamt) and lives on 93 square meters (Mikrozensus-Zusatzerhebung 2018/Statistisches Bundesamt)

Table 1: Descriptive statistics

	Observations	Mean	Median	Std. Dev.	Min	Max
RRP variables						
Total No. of eligbile households	77,305					
– Automatic enrolment (2009 - 2017)	49,182	0.99	1	0.04	0	1
– Elective enrolment (since 2018)	28,123	0.40	0	0.49	0	1
Voucher redemption	77,305	0.26	0	0.44	0	1
Subsidy rate (2009 - 2018)	xx,xxx	0.xx	0.xx	0.xx	0.xx	0.xx
Subsidy rate (2019 - 2020)	xx,xxx	0.xx	0.xx	0.xx	0.xx	0.xx
Household variables						
Number of inhabitants	77,305	2.79	2	1.74	1	15
Electricity price per kWh	77,270	0.28	0.28	0.02	0.03	0.90
Living space in m^2	77,305	69.37	65	24.65	10	300
Annual electricity consumption in kWh	71,513	3,021.18	2,571	1,846.97	0	54,329.15
Old refrigerator variables						
Annual consumption in kWh	29,679	479.62	430	6.57	1	5,840
Age in years	77,299	17.31	16	4.76	1	45
Volume in liters	77,299	239.27	238	76.87	37	733
Estimated savings from replacement in kWh	77,305	336.07	286	166.93	0	5,736

are single households, with about a third retired. 14% are single parent households with one or more children and 6% are retired couples. The remaining 16% in the sample have another household composition. Close to all eligible households are on some type of federal income support. 75% receive unemployment benefits and 12% get a basic income.[15] 5% receive a housing allowance[16] and 4% profit from other public benefits. 3% of households in the sample are not on federal income support. The state with the highest population share within the country is prominently represented in the sample: 38% of households live in North Rhine-Westphalia. Another third of households live in the states Baden-Wuerttemberg, Hesse, Lower Saxony and Berlin, which are among the eight states with the highest population share in Germany.

As the overall aim of such energy efficiency programs is to reduce the financial burden of low-income households resulting from high energy bills, we start our analysis by investigating the average savings of households which successfully managed to replace their refrigerator. In our data, households that have replaced their old refrigerator between 2009-2020 save on average 342 kWh annually.[17] In the same period, the electricity price paid by the households audited in the program has strongly increased from an average of €0.205 in 2009 to €0.289 in 2020 (see Figure 27 in the Appendix) mirroring a general increase in electricity prices in Germany during this period. That is, annual savings realized through the replacement have increased over time from about €70 in 2009 to €99 in 2020. In January 2022, the average price per kWh paid in Germany has further increased to €0.362 [BDEW, 2022] resulting in expected annual savings of €123. At an average purchase price of €478 less the program grant of €100, the investment amortizes after about three years. On a monthly basis, savings are at about €10 which is more than a quarter of the lump sum in the basic income rate for single households intended for electricity and maintenance. As electricity

[15] Retired households with a pension below the minimal income and households with a reduced earning capacity are entitled to basic income. Unemployment benefits and basic income contain a fixed amount for electricity costs which depends on the number of persons in the households. For instance, in 2022 unemployment benefit "ALGII" grants €36.42 for monthly electricity costs for a single household. ALGII also includes a monthly grant of €1.89 to save as investment into a new refrigerator. Some job centers offer interest-free loans to finance durable replacements.

[16] Households with sufficiently low incomes qualify for a partial or total grant of their rent costs.

[17] Old refrigerators consume on average 479 kWh and purchased new refrigerators 138 kWh. In the time span we observe saving rates remain rather constant (see Figure 15 in the Appendix). New efficient refrigerators grow in size over the sample period (see Figure 26 in the Appendix).

prices are projected to keep rising, the significance of the potential savings further increases for low-income households and at the same time stresses the importance of such energy efficiency programs.

Figure 1: Program timeline

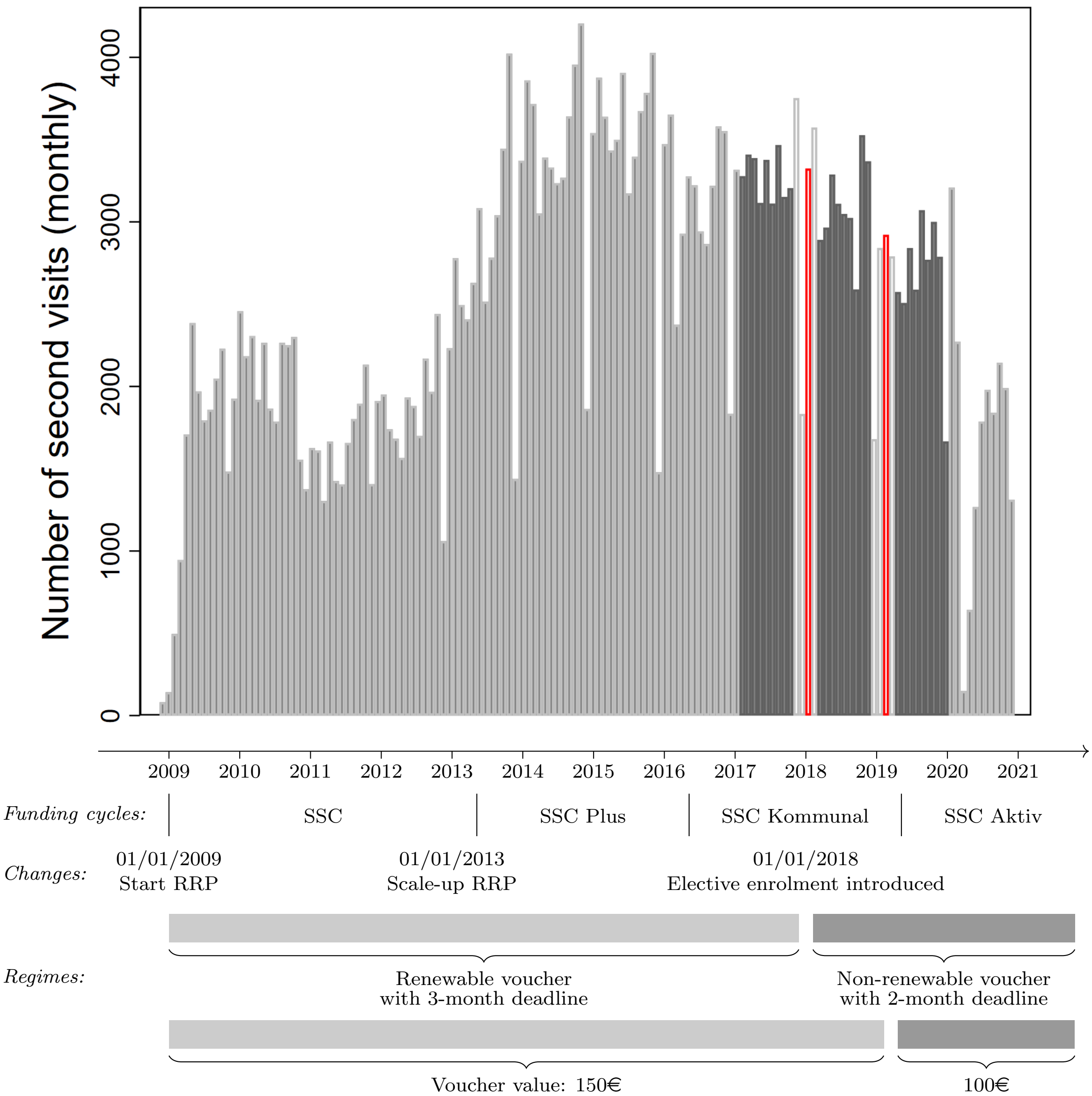


Figure 1 shows the distribution of audits in the program over time, each bar mapping one month from January 2009 to December 2020. The number of monthly audits increases up to 2015 and remains on the high level until it slightly decreases from 2018 on. The dip to zero during the second quarter of 2020 displays the repercussions of the first SARS-CoV-2 lockdown. During the rest of 2020, the number of monthly audits does not yet rebound back to the level of the pre-lockdown months. In our main specification, for the RDD design we use data from February 2017 to January 2020 (see dark-colored bars in Figure 1). Some cyclical fluctuations are visible over the course of each year. The seasonal pattern is particularly pronounced in December, due to the end-of-year and Christmas break at the SSC branches; the month marks the monthly minimum with about a thousand audits less than in the other months each year.

We complement the dataset by a weighted index of cooling appliance prices. We collect data on price indices for refrigerators, freezers, and refrigerator-freezers in Germany (base year 2015) from the Federal Statistical Office (*Verbraucherpreisindex für Deutschland 2021*) and we weight each index according to the share of each RRP category in all newly purchased durables within the program.[18]

4 Empirical strategy

To estimate the effect of varying subsidies and procedures on refrigerator replacement rates, we exploit the temporal variation in the enrolment mode and voucher terms (the procedural change) and in the voucher value (the subsidy change) in a Regression Discontinuity Design (RDD) in time. These program changes mean that over the sample period, we observe eligible low-income households making replacement decisions in three distinct regimes: A regime with automatic enrolment, flexible terms, and a subsidy of €150 up to December 2017, a regime with elective enrolment, rigid terms and a subsidy of €150 up to February 2019[19], and – finally – a regime with elective enrolment, rigid terms, and a subsidy of €100.

RDD analysis relies on a minimum of two important assumptions about our empirical setting. The first is the absence of selection effects: Households need to have been quasi-randomly assigned to the three regimes of the program. The second is comparability: Re-

[18] Refrigerator-freezers make up 77% of all purchased appliances, refrigerators make up 18%, and freezers account for only 5%.

[19] The fourth funding cycle with the new €100 voucher value started on April 1, 2019. Between February 1 and March 31 the RRP paused and no vouchers were issued. Households who underwent a home energy audit during this interim period could request a voucher no sooner than April 1. Therefore, we set the day for the regime change on February 1, 2019 in our analysis.

placement decisions that households took at different points in time and space need to be comparable. On the first, we have three reasons for assuming no evidence for selection bias in the way that households were into the three regimes. One reason is institutional: Both regime changes were unexpected and deviated from the RRP's implementation plan both in terms of substance and timing. Local branches, let alone households, were not given advance information about the discovery of a potential funding shortfall in 2017 or the cut in federal subsidy at the end of 2018. To test formally for evidence that households strategically selected out of or into regimes around regime changes, we test for bunching and discontinuities in household observables around the cutoff points. These tests reveal no visual clues for bunching around the thresholds (see Figures 15 and 16 in the Appendix), and a McCrary test confirms this hypothesis (see Tables 6 and 7 in the Appendix). We also do not find any discontinuities in household observables (see Figures **??** to **??** in the Appendix). Another reason is the dynamic nature of the program: New households become continuously eligible for enrolment into the program as their appliances age while the transparent recruitment process and eligibility criteria remain constant over time. If households responded strategically to the regime, the characteristics of households found eligible would be expected to differ across regimes. Instead, we find that the characteristics of RRP-eligible households, including the features of the refrigerator slated for replacement, do not vary detectably over time (see Figures 9 to 14 in the Appendix). This supports the view that there is no evidence for a clear selection effects and that observations can be treated as independent.

The comparability assumption underlying the RDD analysis is under threat because the conditions within the program under which households take the replacement decision vary over time and space. To be able to compare replacement decision, we account for a range of temporal and spatial factors that likely affect households' investment decision. Such factors comprise changes in the economic environment outside of the RRP such as the persistent increase in German electricity prices during the sample period, the long-run decrease in refrigerator prices during the same period, but also cyclical effects such as seasonal variations in refrigerator prices, reduced household liquidity after the Christmas holidays, and annual adjustments for inflation in federal income support rates at the beginning of each year. Around the federal program, we also account for the presence of complementary programs at the state and municipal level that coexist with the RRP. Finally, temporal and spatial factors also comprise changes inside the program: One example are differences between local branches in program practices and differences in audit quality between advisors, even at the same branch. Transition periods between funding cycles and around unexpected program changes similarly need to be accounted for. The relevance of such transition periods is visible in the data. For example, both right around January 2018 and February 2019, when changes are implemented, the share of audited households that are subsequently enrolled into the RRP drops (see Figure 19 in the Appendix), which can be traced back to some eligible households being denied enrolment. At the same time, the share of redeeming households among eligible households inches higher. Both observations suggest selection to be biased towards households with a high propensity to replace their refrigerator in the transition period.[20]

We ensure comparability through three strategies that help us to jointly account for the dynamic economic environment that the program is embedded in. First, we employ a Donut RDD as proposed by Barreca et al. [2011]. When applying a RDD design to a setting which is prone to irregularities in the observations closely around the policy change, observations in this period should be excluded from the sample on each side of the threshold, creating a "Donut hole".[21] Our preferred Donut RDD excludes two months of observations on each side of the cutoff (program change). This choice controls for the detected bias in the selection towards households with a high propensity to redeem the voucher during the transition periods.

Second, we apply an Augmented Local Linear design to control for seasonality and location effects, thereby increasing the power of estimation. In a two-step approach, we first regress the outcome of interest on time and location indicators using the full sample (2009-2020). We then use the residuals obtained from this first step as outcome in the second step, the RDD estimation in a bandwidth of 6 to 11 months around the policy change [Hausman and Rapson, 2018].[22] We apply Augmented Local Linear using a series of temporal and spa-

[20] This is despite the fact that selection into treatment is not biased as bunching and discontinuity tests indicate.

[21] Other applications of Donut RDD are Ost et al. [2018], Kim and Koh [2020], Gillingham and Huang [2021].

[22] Augmented Local Linear has been used in other RDD analyses [Li et al., 2020, Tan and Mao, 2021, Gillingham and Huang, 2021].

tial indicators. We control for different practices at the local branches and changes at each branch over time by including branch and branch-by-year fixed effects, for audits by different advisors by including fixed effects for the advisors who conducted the first and second visit of the audit separately, and for complementary programs by states, municipalities and energy providers by including state and branch fixed effects. To control for seasonal variation in liquidity we include month fixed effects, for adjustments in the income support rates we include year-by-income support type fixed effects. To control more granularly for differences in the socio-economic environment we also include ZIP code fixed effects, and to control for other time trends we include month-by-year and year fixed effects.

Third, we add further explanatory variables to our econometric model to correct for potential differences in the household groups before and after each regime change.[23] We add the price paid per kWh at the household level to adjust for differences in electricity prices, and we add further controls which could influence the replacement decision of households, such as the number of persons in the household, the type of income support received, living space, total electricity consumption, the age and size of the old refrigerator, and the calculated savings after replacement. We also add a refrigerator price index as control for changes in refrigerator purchasing prices over time.

Our identification strategy uses between-household variation to estimate the effect of the changes in the subsidy level and the procedures on the replacement decision. We estimate the basic RDD equation as follows, separately for the subsidy and procedural variations:

$$Outcome_{it} = \beta_0 + \beta_1 Policy_t + \beta_2 DayCount_t + \beta_3 X_i + \varepsilon_{it} \tag{1}$$

Policy indicates the current regime as a binary treatment variable: 1 for €150/0 for €100 (non-renewable voucher in both regimes) or, alternatively, 0 for the renewable voucher/1 for the non-renewable voucher with a strict deadline (€150 voucher value in both regimes). *DayCount* is the running variable counting the number of days from the policy change. X is

[23] Table 8 provides a comparison in means for relevant covariates before and after each policy change (voucher value and procedures). Small imbalances in some of the variables are due to changes of variables exogenous to the program, e.g. the refrigerator price index which varies seasonally, and the average electricity price per kWh which increases over the sample period in Germany, in turn increasing the estimate for savings after replacement.

the vector of controls. The subscripts t and i denote time in days and individual households.[24]. We estimate the equation for three outcomes of interest:

1. The replacement rate: the share of households that redeem the voucher out of all eligible households. The variable of interest is the binary decision to replace the refrigerator, estimated on the sample of eligible households.

2. The enrolment rate: the share of households that enrol in the program out of all eligible households. The variable of interest is the binary decision to enrol, estimated on the sample of eligible households. We only observe this outcome for the period as of 2018.

3. The redemption rate: the share of households that redeem the voucher out of all enrolled households. The variable of interest is the binary decision to redeem the voucher after replacement, estimated on the sample of enrolled households. We only observe this outcome for the period as of 2018.

In our main specification, we apply the Donut RDD and use the Augmented Local Linear to adjust the outcome as described above. We estimate equation (1) as linear probability model in a bandwidth of six to eleven months around each regime change.[25] We bootstrap standard errors, using 50 repetitions. We run robustness checks that estimate the treatment effects for subsamples of only households in North Rhine-Westphalia that receive a large amount of additional funding from the state government on top of the federal subsidy, and only non-NRW households. Additionally, we vary the size of the Donut or omit the Donut design, skip the Augmented Local Linear, and estimate a binary probability model instead of a linear probability model.

[24] We choose the most basic RDD specification without allowing for a more flexible functional form as this is in line with both the empirical appearance of the data and economic reasoning, and as is practice in many empirical studies [Gelman and Imbens, 2019, Pei et al., 2021]

[25] We choose the minimum bandwidth at $+/-$ 6 months as precision of the estimates is low with a bandwidth below 6 months. The maximum bandwidth of $+/-$ 11 months is determined by data constraints: a longer bandwidth choice for both treatment effects would include observations located inside the transition period of the respective other design change and would bias the estimations.

5 Main Results

5.1 Subsidy variations

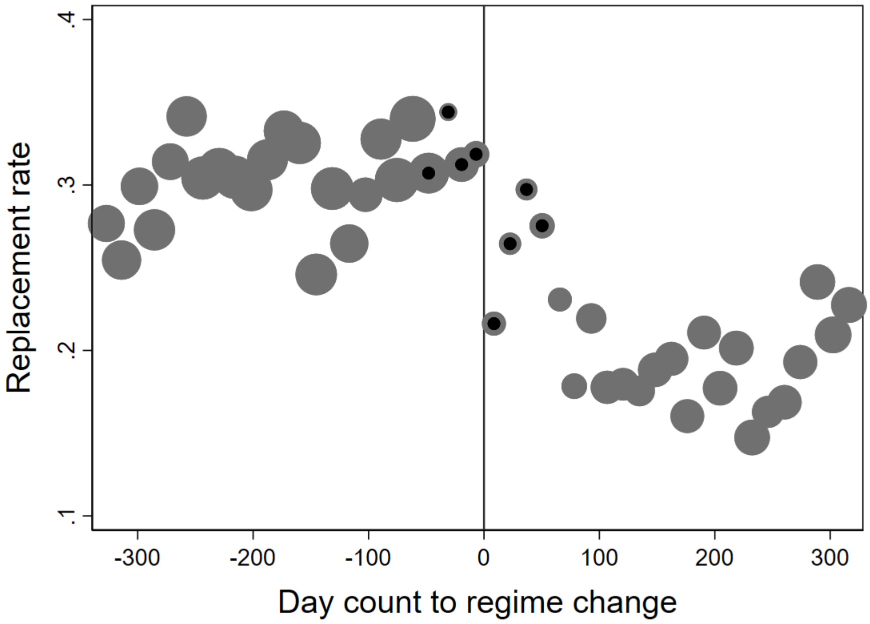
We first investigate to what extent replacement decisions among eligible households respond to a €50 variation in the voucher-based subsidy. This variation is large relative to the voucher values of €100 and €150, respectively. It also leads to sizeable variations in the subsidy share as a percentage of the retail price of new refrigerators (see Table 1), from around 40% of the price before to around 20% after the change (CROSS-CHECK). The effect size of the €50 variation, predicted to be significant, positive, and economically meaningful, also provides an intuitive benchmark for gauging the effects of procedural variations in the following section.

Figure 2 shows the replacement rate around the subsidy change from €150 to €100. Day 0 is February 1, 2019. Negative day counts cover the period when the voucher value is €150, positive day counts the period when the voucher value is €100. Each bubble captures the average replacement rate within a 14 day interval, with larger bubbles signifying more observations. Observations marked with black dots lie in the transition period and are excluded by the Donut design. By inspection, replacement rates respond to subsidy levels as expected. They vary around 0.3 for negative day counts: About one in three eligible households elects to enrol and redeems the €150 voucher. For positive day counts, replacement rates vary around 0.2: About one in five households elects to enrol and redeems the €100 voucher. This suggests that the reduction in the subsidy is associated with a 10 percentage point reduction in the share of eligible households replacing their refrigerator.

Table 2 provides our estimation results for a bandwidth between six and eleven months, with and without controls. All models indicate the treatment indicator of subsidy variation (= 1 for the voucher of €150, 0 for €100) to significantly differ from zero ($p<0.001$),[26] confirming the visual impression of Figure 2: Households react to prices, leading to a lower replacement rate after the reduction of the voucher value to €100. In our preferred specifications (Columns II and IV) that account for the Donut design, the Augmented Local Linear approach and further control variables we estimate the replacement rate to be 8.7 to 15.8 percentage points

[26] Appendix Table 9 provides robustness check results.

Figure 2: Subsidy variation: Discontinuity graph



higher for a voucher that has a €50 higher value.[27] In other words, a 33% percent lower subsidy level is associated with a likelihood of appliance replacement that is 9 to 16 percentage points lower.

5.2 Procedural variations

Against the background of the effect of a €50 variation in subsidy levels, we now turn to the effect of simultaneous procedural changes from automatic to elective enrolment and from flexible to rigid voucher terms.

Figure 3 shows the replacement rate around the procedural change. Day 0 is January 1, 2018. Negative day counts cover the period when enrolment was automatic and voucher terms flexible, positive day counts the period when enrolment was elective and voucher terms rigid. As before, each bubble captures the average replacement rate within a 14 day interval, with larger bubbles signifying more observations. Observations marked with black dots lie in the transition period and are excluded by the Donut design. By inspection, noise is strong inside the transition period. Outside, the average replacement rate lies around 0.25 before the

[27] Figure 20 in the Appendix shows how the treatment effect changes as function of the bandwidth.

Table 2: Estimated effects of subsidy changes on the replacement rate

	I	II	III	IV
Subsidy Variation	0.096*** (0.016)	0.087*** (0.022)	0.154*** (0.009)	0.158*** (0.010)
Day count	yes	yes	yes	yes
Controls		yes		yes
Bandwidth in months	6	6	11	11
No. observations	7,222	6,739	16,434	15,401

Notes: Controls include the number of inhabitants, living space, the type of federal income support received, kWh price paid, yearly electricity consumption, old appliance age, appliance price index and estimated replacement savings. The Donut design excludes 2 months around the regime change. The Augmented Local Linear approach uses month, month-by-year, year, state, branch, branch-by-year, ZIP code and auditor controls. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Estimated on the sample of eligible households in a bandwidth around February 1, 2019.

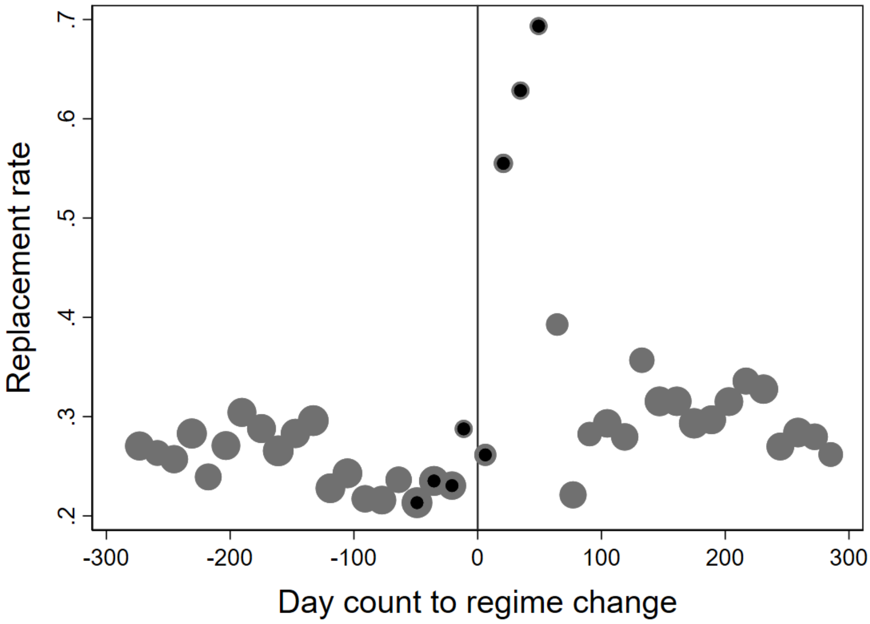
transition: About a quarter of automatically enrolled eligible households redeem the €150 voucher upon replacing their refrigerator. After the transition, the replacement rate rises to around 0.3: Around a third of eligible household elect to enrol in the RRP and successfully redeem the €150 voucher with rigid terms.

Table 3 provides our estimation results. The specifications are analogous to the estimation of the subsidy effect. We estimate a significant positive coefficient in all four specifications, which confirms the visual impression.[28] Based on our preferred specifications II and IV, we estimate the replacement rate to be 3.9 to 9.7 percentage points higher under elective enrolment with rigid terms compared to automatic enrolment with flexible terms.[29] The direction and size of the effect of the procedural variations merit attention, in particular in light of their small, possibly negative costs to the program. Comparing these effects of varying procedures to those of a variation in a subsidy in a back-of-the envelope calculations stresses the merits of alternative program design. The procedural variations within the RRP

[28] Appendix Table 12 provides robustness check results.

[29] Figure 31 in the Appendix shows how the treatment effect changes as function of the bandwidth.

Figure 3: Procedural variations: Discontinuity graph



appears to stimulate the adoption of energy-efficient appliance among low-income households and to deliver one half to two-thirds of an increase that would require a €50 increase in the subsidy.

6 Mechanisms

6.1 Subsidy variations: Enrolment and redemption effects

The procedures in place when the subsidy is changed from €150 to €100 are elective enrolment and rigid voucher terms. Since RRP records register whether a household enrolled and whether the enrolled households redeemed the voucher, we are able to examine the effect of varying the subsidy on refrigerator replacement more closely by decomposing it into two distinct effects, one at the enrolment stage and one at the redemption stage.

Figure 4 shows a discontinuity graph similar to Figure 3 for the enrolment stage. The key difference is the enrolment rate as the outcome variable, i.e. the share of households that enrol in the program out of all eligible households. By inspection, enrolment rates are around 0.4 before the subsidy change and the transition period (black dots): Around 40%

Table 3: Estimated effects of procedural changes on the replacement rate

	I	II	III	IV
Procedural change	0.095*** (0.016)	0.097*** (0.018)	0.044*** (0.009)	0.039*** (0.007)
Day count	yes	yes	yes	yes
Controls		yes		yes
Bandwidth in months	6	6	11	11
No. observations	9,102	8,539	20,532	19,174

Notes: Controls include the number of inhabitants, living space, the type of federal income support received, kWh price paid, yearly electricity consumption, old appliance age, appliance price index and estimated replacement savings. The Donut design excludes 2 months around the regime change. The Augmented Local Linear approach uses month, month-by-year, year, state, branch, branch-by-year, ZIP code and auditor controls. * p<0.05, ** p<0.01, *** p<0.001. Estimated on the sample of eligible households in a bandwidth around January 1, 2018.

of eligible households elect to enrol in the RRP for a subsidy of €150. After the change in the subsidy and the transition period, the enrolment rate settles around 0.3: Roughly 30% of eligible households elect to enrol for a subsidy of €100. During the transition period, enrolment rates are elevated.[30]

Table 4 provides estimation results, using the same specifications as for the replacement rate in section 5. All specifications show a positive significant coefficient (p<0.001), mirroring the results of our descriptive analysis.[31] In our preferred specifications II and IV, we estimate the enrolment rate to be 24.8 to 24.9 percentage points higher for a €50 higher voucher value.[32] That is, for a higher subsidy, we observe significantly more households electing to enrol in the program.

[30] An important factor in the elevated levels are irregularities in the issuance of the invitation letters to households during the transition period: Despite fulfilling the eligibility criteria, there is evidence of invitation letters being withheld (see the eligibility ratio in Figure 19 in the Appendix). This has the effect of decreasing the denominator of the enrolment rate, driving up the enrolment rate.

[31] Appendix Table 10 provides robustness check results.

[32] Figure 21 in the Appendix shows how the treatment effect changes as function of the bandwidth.

Figure 4: Subsidy variations, enrolment stage: Discontinuity graph

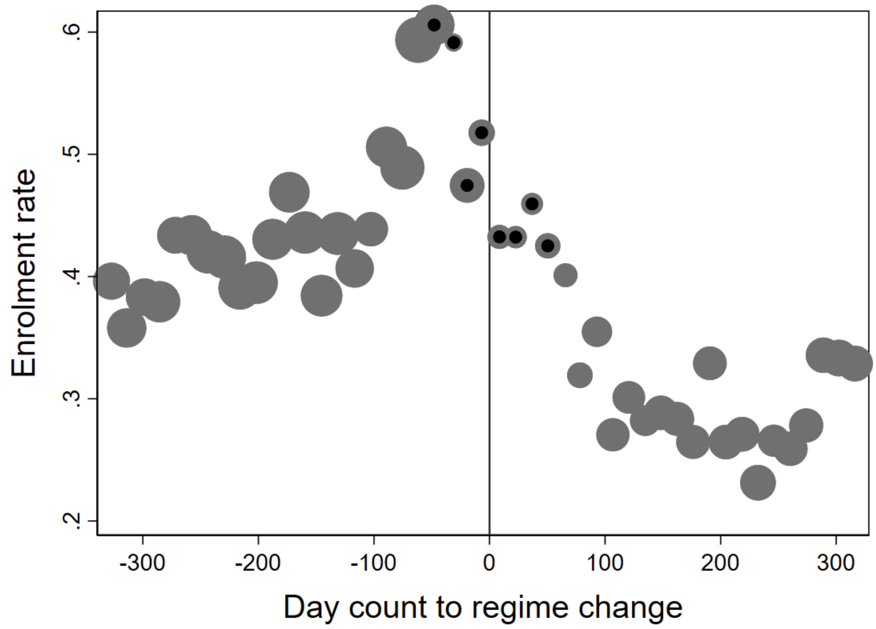


Table 4: Estimated effects of subsidy variation of €50 on the enrolment rate

	I	II	III	IV
Voucher €150	0.210*** (0.018)	0.248*** (0.025)	0.225*** (0.011)	0.249*** (0.010)
Day count	yes	yes	yes	yes
Controls		yes		yes
Bandwidth in months	6	6	11	11
No. observations	7,222	6,739	16,434	15,401

Notes: Controls include the number of inhabitants, living space, the type of federal income support received, kWh price paid, yearly electricity consumption, old appliance age, appliance price index and estimated replacement savings. The Donut design excludes 2 months around the regime change. The Augmented Local Linear approach uses month, month-by-year, year, state, branch, branch-by-year, ZIP code and auditor controls. * p<0.05, ** p<0.01, *** p<0.001. Estimated on the sample of eligible households in a bandwidth February 1, 2019.

The redemption stage of the replacement process is captured in the discontinuity graph of Figure 5. The key difference to the previous analysis is the redemption rate as the dependent variable, i.e. the share of enrolled households that redeem the voucher. Redemption rates are characterized by considerable variation, both before, around (black dots), and after the change in voucher value. By inspection, they lie in the range between 0.5 and 0.8 up to 300 days before the change and 0.5 to 0.66 up to 100 days before: One half to two thirds of enrolled household redeem their voucher for €150 in cash after replacing their refrigerator. After the change, the redemption rates are between 0.50 and 0.75: One half to three quarters of enrolled household redeem their €100 voucher. As a result, there is no clear effect visible at the redemption stage.

Figure 5: Subsidy variation, redemption stage: Discontinuity graph

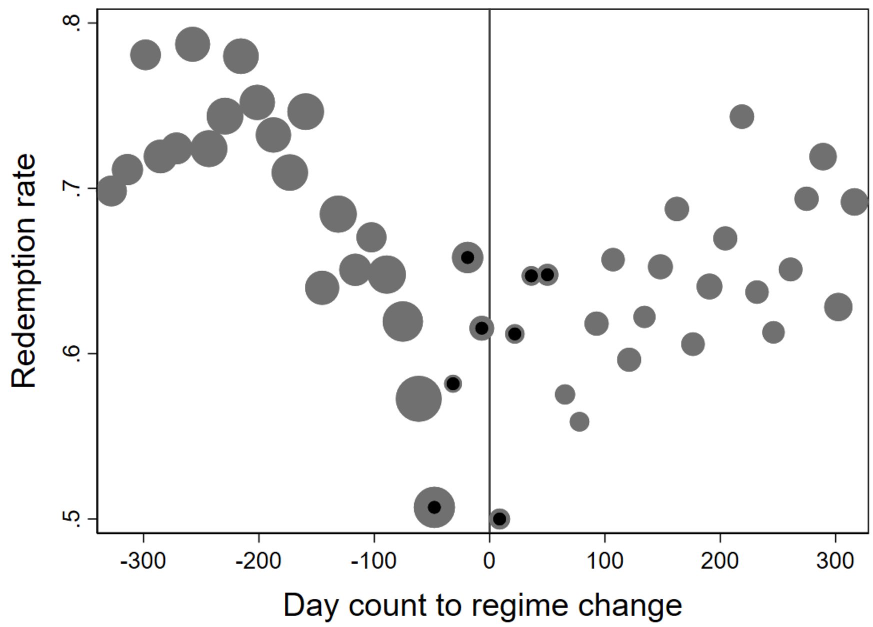


Table 5 reports the formal estimation results, using the same specifications as in the previous models. Only the treatment coefficients in I and II are significant.[33] In our preferred specifications II and IV, we estimate the redemption rate to decrease by 3.1 to 16.5 percentage points for a €50 higher voucher value. [34] Intuitively, it could be expected that households holding a voucher worth €150 rather than €100 are more likely to replace successfully their

33 Appendix Table 11 provides robustness check results.
34 Figure 22 in the Appendix shows how the treatment effect changes as function of the bandwidth.

refrigerator and redeem the voucher. Statistically, however, the evidence is weak.

Table 5: Estimated effects of a €50 subsidy variation on the redemption rate

	I	II	III	IV
Subsidy variation	-0.093** (0.034)	-0.165*** (0.038)	0.008 (0.022)	-0.031 (0.022)
Day count	yes	yes	yes	yes
Controls		yes		yes
Bandwidth in months	6	6	11	11
No. observations	2,774	2,617	5,955	5,616

Notes: Controls include the number of inhabitants, living space, the type of federal income support received, kWh price paid, yearly electricity consumption, old appliance age, appliance price index and estimated replacement savings. The Donut design excludes 2 months around the regime change. The Augmented Local Linear approach uses month, month-by-year, year, state, branch, branch-by-year, ZIP code and auditor controls. * p<0.05, ** p<0.01, *** p<0.001. Estimated on the sample of households that have requested a voucher in a bandwidth around February 1, 2019.

Combining these insights, our data suggests that out of the two candidate mechanisms, one at the enrolment and one at the redemption stage, only one is operational. This means that the effect of varying the subsidy estimated in section 5 is predominantly a recruitment effect at the enrolment stage. Reducing the value of the voucher from €150 to €100 results in a lower propensity among eligible households to elect enrolment by requesting the voucher. Once households hold the voucher, its cash value no longer reliably influences the chance that the household will actually replace the refrigerator. This finding demonstrates that the economic incentive did not succeed at every margin of decision-making.

6.2 Procedural variation: Behavioral effects

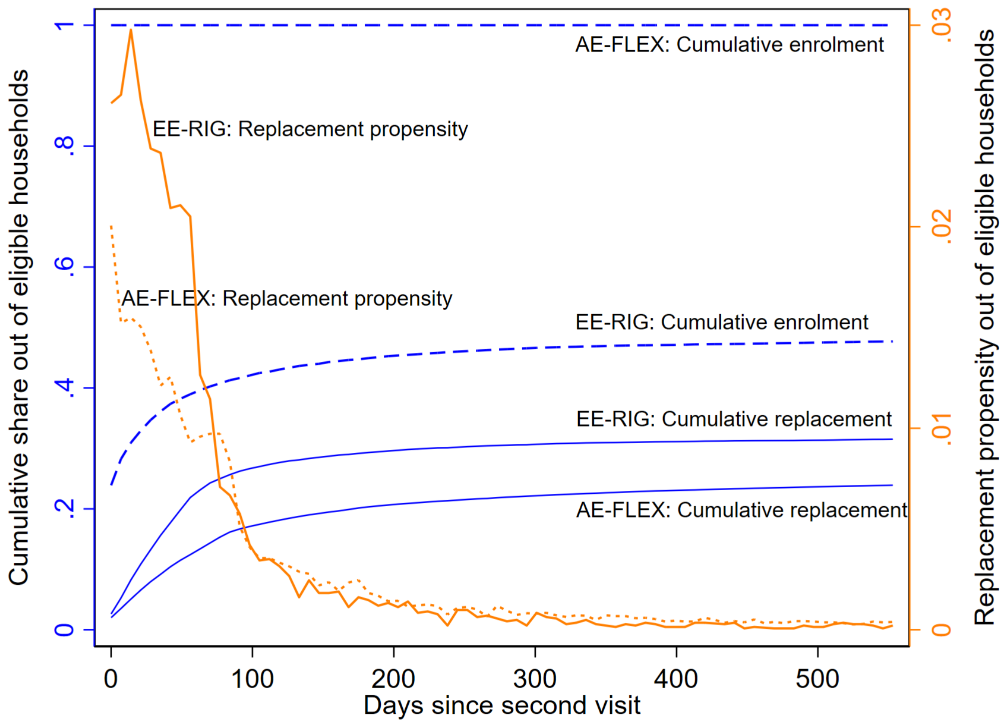
To understand more about the mechanisms behind the effect of procedural variation on the success rate of the RRP, we take a closer look at how the behavioral patterns before and after the procedural changes compare.

Figure 6 shows, as a function of days passed since the second home visit, three temporal patterns, two cumulative (in blue, left scale), one intensive (in yellow, right scale), under two regimes, automatic enrolment and flexible terms (AE-FLEX) and elective enrolment and rigid terms (EE-RIG). The first cumulative dynamic is the share of enrolled households among all eligible households, the second the cumulative replacement rate among all eligible households. The intensities over time are the replacement propensities among eligible households.

Enrolment before the change is automatic (AE-FLEX). As a result, cumulative enrolment of eligible households (blue, right scale) mechanically jumps to 100% on the day of the second visit. After the change, enrolment is elective (EE-RIG). Cumulative enrolment starts at around 20% of eligible households that enrol on the day of the second visit and grows at a slowing rate to top out at 44%. 90% of elective enrolment occurs within 90 days following the second visit. The differences in enrolment patterns mean that under elective enrolment, more than half of eligible households never request the voucher that they would have automatically received under the previous scheme. This removes thousands of households for whom replacement has been determined to be economically advantageous from the pool of potentially replacing households. The sizeable drop in cumulative enrolment can plausible be traced to 'hassle' costs of overcoming psychological frictions, time and effort costs when enrolment is elective. Despite their small size relative to the gains from replacement, such costs have been shown to effectively deter households from enrolling in social assistance programs. [Bertrand et al., 2006, Bhargava and Manoli, 2015]. At the same time, the drop in cumulative enrollment provides important information to the manager of the program, in particular if vouchers are costly to issue and require managers to set aside funds.

The key performance metric of the RRP is not the enrolment, but the replacement rate. As expected, these rates start at zero for both regimes and grow more slowly than enrolment. Despite the lower cumulative enrolment, the cumulative replacement rate reaches 29% of eligible households when enrolment is elective and voucher terms are rigid (EE-RIG). This is consistently higher than under automatic enrolment and flexible terms. There, 22% of eligible households ultimately replace their refrigerator, most within the 90-day validity period of their first voucher. The reasons for the difference in performance between the two procedural regimes are not obvious. While selection effects could trivially explain why cumulative

Figure 6: Cumulative share and replacement propensity as function of days since second visit



replacement under EE-RIG is <u>not lower</u> than under AE-FLEX, additional mechanisms must be at play in order to explain why it is higher.

To dig deeper, we examine the temporal patterns of replacement propensity between the two regimes. Under AE-FLEX, about 2% of eligible households replace immediately after the second visit. This points to households having advance notice of their eligibility and awaiting voucher receipt on the second visit for final implementation. Replacement intensity then falls off, before increasing again to 1% as the first voucher approaches the end of its 90-day validity. After that, the decline is fairly rapid, but some replacement activity still takes place long after the second visit. Progressively smaller peaks of replacement activity are detectable after 180 and 270 days, when the second and third voucher expire. Under EE-RIG, replacement intensity starts at a considerably higher level, indicating more preparedness among households ready to enrol than under AE-FLEX, and first increases, peaking at about 3% roughly a month after the second visit. It then falls off, with a shoulder at around 60 days. This could indicate the expiry of those vouchers that were requested immediately on or following the second visit. After 80 days, replacement intensity under EE-RIG falls below that of AE-FLEX and does not recover.

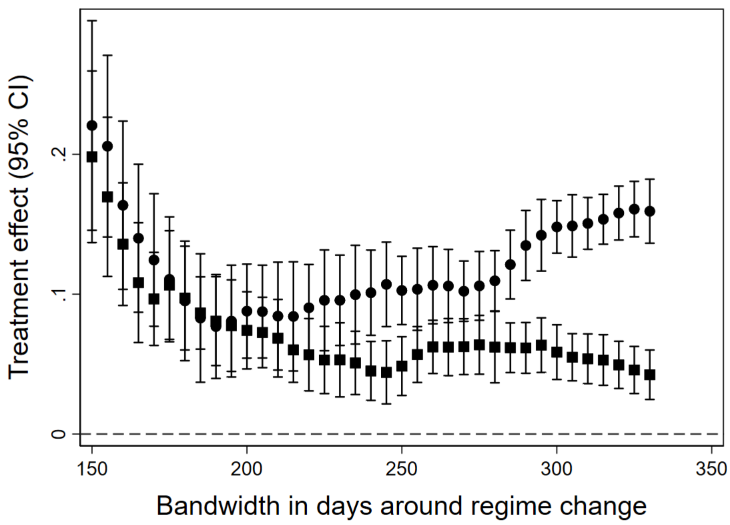
Comparing these patterns, it becomes clear that the differences in cumulative replacement rates stem from phenomena that arise at and right after the second visit. The typical electively enrolled households replace more vigorously and complete their planned replacement faster than their automatically enrolled counterpart. One candidate explanation advanced by psychologists relates such behavior to the extensive and intensive margins of goal setting [Locke and Latham, 1990] implicit in the voucher terms. Rigid terms commit the enrolling household receiving the voucher to meeting a two-month replacement goal. Such terms have been referred to as a ‘pseudo ‘self-set’ goal’ [Burdina et al., 2017] because the terms are set by an outside agency, but voluntarily adopted by a subset of households wishing to receive the subsidy. They have little impact on the median household, but affect the tail end of the distribution. At the extensive margin, such goals lead to a demotivation effect: Individuals who consider the goals set by the outside agency as unattainable do not adopt the goal [Burdina et al., 2017]. In the RRP, the changes to rigid terms could therefore demotivate those eligible households who consider themselves unable to undertake – within two months – the not insignificant efforts required from themselves to complete all the steps of the RRP. At the intensive margin, there is an counteracting motivation effect: Challenging, but attainable goals lead to a higher likelihood of task completion [Harding and Hsiaw, 2014, Burdina et al., 2017]. Related to this argument, voucher terms can also sharpen the implementation intention to support the realization of goal intentions by specifying ”when, where, and how goal-directed responses should be initiated” [Achtziger et al., 2008](p.381). This in turn does not only facilitate the starting process but also prevents households to stray from the intended path. In the RRP, some households who would not have completed the replacement within 60 days under the flexible regime could therefore adopt the goal and be more motivated to redeem the voucher within its term limits. This positive effect on the implementation decision can therefore explain the sharper increase in cumulative replacement rates in EE-RIG compared to AE-FLEX within the first 60 days. In addition, we observe a deadline effect in EE-RIG: Approaching the 60 days under the rigid regime leads again to an increase in the redemption probability (see Figure 24).

This highlights the potential to use behaviorally informed procedural changes such goal setting in the future in an effort to target more narrowly the motivation effect detected here.

7 Policy assessment

In this section we discuss the economic magnitude of our estimates in a stylized back-of-the envelope calculation. To do so, we discuss the benefits of both the change in the subsidy levels and the procedures against the cost of implementation. As shown in Figure 7, a direct comparison of the treatment effects reveals the subsidy variation to be up to four times but at least as effective as the procedural variation to stimulate additional refrigerator replacement. The relative difference depends on the choice of bandwidth.

Figure 7: Comparison of subsidy and procedural variation as function of bandwidth



Notes: This figure shows the estimate of the subsidy (dot) and the procedural effect (square) for different bandwidth choices.

The direct comparison in size has to be evaluated against the background that the procedural variation is considerably less costly to implement. A change in the voucher redemption value from €100 to €150 increases the average replacement rate by 9 to 16 percentage points. The higher voucher value increases efficiency of each conducted home visit since the probability of replacement increases, so that the net benefit per home visit rises.[35] But implementation of the change in the subsidy increases the cost per replacement by 50%. In addition, as shown in the results section, a higher voucher values predominantly affects the enrolment stage: More households request a voucher but the share of successfully redeemed vouchers remains unchanged which in turn leads to higher administrative costs. The efficiency of the subsidy variation would be higher if it affected households at the redemption stage.

In comparison, the procedural changes boost the replacement rate by "only" 4 to 10 percentage points. In contrast to the subsidy variation it comes with close to zero additional costs or even reduces the administrative cost of the program as less vouchers have to be kept track of and kept on the balance sheet. Using the estimates of the procedural change (4 to 10 percentage points) and the subsidy variation (9 to 16 percentage points) to conduct a back-of-the envelope calculation yields that procedural changes generate an improvement in replacements that is equivalent to an estimated subsidy increase of €12 to €56 per replacing household.

In addition to these comparisons, we can use our estimated for a counterfactual program scenario. We ask to what extent energy efficiency programs for low-income households using voucher-based subsidies, which – in contrast to the SSC program – have not yet introduced elective enrollment and rigid terms, could have increased the number of replacements. As shown, our point estimates suggest at least 400 additional refrigerators being replaced for every 10,000 invitations to actively claim a voucher with a rigid term. Applied to our observation period (2013 – 2017) and assuming a constant treatment effect over time (4 to 10 percentage points) [95% CI: 3.2 pp; 11.5 pp][36] we calculate that elective enrollment and rigid terms to have led to at least 1,900 (= 0.04 x 48,615) [95% CI: 1,556; 5,591] additional refrigerator replacements. At an average electricity prices of €0.289 in 2020 and average annual savings of 342 kWh this would have led to additional savings in electricity bills of €187,800.

[35] We cannot precisely quantify the replacement program's cost (recording refrigerator information and reporting it in the database, generating the infoletter, handing it to the household and explaining the details of the replacement process) as the share of the cost of home visits.

[36] Here and in the following, we provide the lower and upper bound of the 95% confidence interval of the low and high estimate respectively.

8 Conclusion

A growing literature in behavioral public policy has lately been demonstrating how program design affects program performance, in particular for policies targeting low-income households. Our paper adds to this literature by studying – in the form of investments in energy-efficient appliances – a type of household decision that differs from the labor supply or consumption decisions typically examined. This study benefits from empirically favorable circumstances: Not only were the changes in program design quasi-exogenous, they also varied both subsidy levels and procedures separately. As a result, our paper cannot just speak to the impact of each variation on program performance individually, but also how variations along these two dimensions compare. To our knowledge, such evidence has so far not been available in the literature.

Over its lifetime so far, the national Refrigerator Replacement Program in Germany has amassed twelve years (2009 - 2020) of data from over 77,000 eligible low-income households making a replacement decision about the most energy-intensive home appliance through a voucher-based subsidy program. In combination with the changes in subsidies and procedures, these data offer a rare glimpse into the 'black box' of consumer durable replacement decisions among the poor under three different program design regimes. As a result, We have three main findings. One is the first evidence on the subsidy elasticity of replacement decisions: A 50% higher subsidy increases the likelihood of refrigerator replacement by 9 to 16 percentage points. We can attribute this effect to changes exclusively at the extensive margin of the program: More households enrol in the program when the subsidy is higher, but the same share of enrolled household replaces their fridge. The second is the evidence on how replacement rates are affected by procedural changes. These rates are 4 to 10 percentage points higher under elective enrolment and rigid terms than under automatic enrolment and flexible terms. This overall change consists of a drop at the enrolment stage from 100% to 40% of eligible households combined with an increase in redemption rates strong enough such that cumulative replacement rates after the changes outperform those before for every point in time following the second home visit. Additional observational evidence points to households entering the RRP more prepared and accelerating replacement as the fixed deadline approaches. Such patterns are consistent with a behavioral interpretation that the procedural changes facilitated goal setting by households and helped overcome time management problems.

Our third main finding is that, comparing the subsidy and the procedural variation, the accidental changes in how to enrol households and what voucher terms to set were equivalent – in terms of replacement rates – to raising the subsidy by between €15 and €56. These numbers give a intuitive metric to the potential of procedural changes to affect program performance. They are also at the basis of our conservative estimate of an additional 1,900 refrigerators that could have been replaced if the new procedures had been in place from 2013 onwards. We believe that this finding in particular should be of interest to researchers investigating how best to deliver energy efficiency improvements to low-income households.

In our mind, the novel evidence on the comparative impact of procedural changes on program performance has implications for future research for three reasons. One is that our results make it more likely that (re-)evaluations of existing programs will also uncover effects of procedural changes on program performance. Many small changes in procedures happen for reasons other than deliberate program optimization. The RRP is a case in point. There, unexpected budgetary considerations of the program sponsor and sudden realization of problematic implications of current procedures for budgeting were the main drivers. Such changes may be easily treated as an empirical nuisance in ex post evaluations of programs or simply be overlooked as seemingly irrelevant. A wider effort to identify procedural changes and to estimate their effects on program success is likely to contribute to a greater understanding of how and why procedures matter for program success.

The second reason is that our evidence highlights the potential of the economics of program design benefiting from progress towards theoretically and empirically informed procedural changes. Changes that are accidental or driven by expediency should over time give way to deliberate changes. These deliberate changes will be progressively informed by evidence that was generated through purposeful experimentation. This evidence should be complemented by careful studies of how changes in procedures affect program costs. For example, in the RRP there was a perception that having fewer voucher in circulation simplified administrative procedures, reduced workload fluctuation, and required less budget to be set aside to cover possible late redemption. If correct, these changes therefore came at negative cost. The joint presence of accidental procedural changes delivering both unanticipated performance improvements and unanticipated cost savings leads us to believe that the economics of program design retain the potential to make significant contributions to behavioral public policy.

The third reason why the evidence presented here can inform future research is that it highlights an unexplored dimension of program design. This dimension is how to optimally integrate economic incentives and procedures for program design. When the subsidy and the procedural variations were introduced in the RRP, design optimization was not part of the agenda. On the basis of results in the marketing literature, however, the conjecture that combining economic and procedural elements in a single program re-design could help boost program performance further appears promising but will need to await future empirical opportunities in order to be tested.

References

Anja Achtziger, Peter M. Gollwitzer, and Paschal Sheeran. Implementation intentions and shielding goal striving firm unwanted thoughts and feelings. Personality and Social Psychology Bulletin, 34:381–393, 2008.

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

Steffen Altmann, Christian Traxler, and Philipp Weinschenk. Deadlines and memory limitations. Management Science, 2021.

Nadia Ameli and Nicola Brandt. Determinants of households' investment in energy efficiency and renewables: Evidence from the OECD survey on household environmental behaviour and attitudes. Environmental Research Letters, 10(4), 2015.

Mark A. Andor, David H. Bernstein, and Stephan Sommer. Determining the energy efficiency gap of residential electricity consumption. Empirical Economics, 60:2897–2923, 2021.

Alan I Barreca, Melanie Guldi, Jason M Lindo, and Glen R Waddell. Saving babies? revisiting the effect of very low birth weight classification. The Quarterly Journal of Economics, 126 (4):2117–2123, 2011.

BDEW. Bundesverband der Energie- und Wasserwirtschaft e.V.: Energiemarkt Deutschland. Technical report, 2019.

BDEW. Bundesverband der Energie- und Wasserwirtschaft e.V.: Strompreisanalyse Januar 2022. `https://www.bdew.de/service/daten-und-grafiken/bdew-strompreisanalyse/`, 2022. Accessed: 2022-02-28.

Marianne Bertrand, Sendhil Mullainathan, and Eldar Shafir. A behavioral-economics view of poverty. American Economic Review, 94(2):419–423, 2004.

Marianne Bertrand, Sendhil Mullainathan, and Eldar Shafir. Behavioral economics and marketing in aid of decision making among the poor. Journal of Public Policy & Marketing, 25(1):8–23, 2006.

Marianne Bertrand, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman. What's advertising content worth? evidence from a consumer credit marketing field experiment. The quarterly journal of economics, 125(1):263–306, 2010.

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

Mariya Burdina, R Scott Hiller, and Neil E Metz. Goal attainability and performance: Evidence from boston marathon qualifying standards. Journal of Economic Psychology, 58:77–88, 2017.

Laurent E. Calvet, John Y. Campbell, and Paolo Sodini. Measuring the financial sophistication of households. American Economic Review, 99(2):393–98, 2009.

L.W. Davis, A. Fuchs, and P. Gertler. Cash for coolers: evaluating a large-scale appliance replacement program in mexico. American Economic Journal: Economic Policy, 6(4): 207–238, 2014.

Manasi Deshpande and Yue Li. Who is screened out? application costs and the targeting of disability programs. American Economic Journal: Economic Policy, 11(4):213–48, 2019.

EIA. U.s. energy information administration: https://www.eia.gov/todayinenergy/detail.php?id=36692. Technical report, 2015.

David T Ellwood. Anti-poverty policy for families in the next century: From welfare to work–and worries. Journal of Economic Perspectives, 14(1):187–198, 2000.

Meredith Fowlie, Michael Greenstone, and Catherine Wolfram. Are the non-monetary costs of energy efficiency investments large? Understanding low take-up of a free energy efficiency program. American Economic Review, 105(5):201–04, 2015.

Meredith Fowlie, Michael Greenstone, and Catherine Wolfram. Do energy efficiency investments deliver? Evidence from the Weatherization Assistance Program. The Quarterly Journal of Economics, 133.3:1557–1644, 2018.

Manuel Frondel, Gerhard Kussel, and Stephan Sommer. Heterogeneity in the price response of residential electricity demand: A dynamic approach for Germany. Resource and Energy Economics, 57:119–134, 2019.

Andrew Gelman and Guido Imbens. Why high-order polynomials should not be used in regression discontinuity designs. Journal of Business & Economic Statistics, 37(3):447–456, 2019.

Kenneth Gillingham and Pei Huang. Air pollution, health, and racial disparities: Evidence from ports. 2021.

Kenneth Gillingham and Tsvetan Tsvetanov. Nudging energy efficiency audits: Evidence from a field experiment. Journal of Environmental Economics and Management, 90:303–316, 2018.

Pedro I Hancevic and Hector H Sandoval. Low-income energy efficiency programs and energy consumption. Journal of Environmental Economics and Management, page 102656, 2022.

Matthew Harding and Alice Hsiaw. Goal setting and energy conservation. Journal of Economic Behavior & Organization, 107:209–227, 2014.

Justine S Hastings and Jeffrey M Weinstein. Information, school choice, and academic achievement: Evidence from two experiments. The Quarterly journal of economics, 123 (4):1373–1414, 2008.

Catherine Hausman and David S Rapson. Regression discontinuity in time: Considerations for empirical applications. Annual Review of Resource Economics, 10:533–552, 2018.

Sebastién Houde and Joseph Aldy. Consumers' response to state energy efficient appliance rebate programs. American Economic Journal: Economic Policy, 9(4):227–255, 2017.

Seonghoon Kim and Kanghyock Koh. Does early access to pension wealth improve health? Economic Inquiry, 58(4):1783–1794, 2020.

Pei Li, Yi Lu, and Jin Wang. The effects of fuel standards on air pollution: Evidence from China. Journal of Development Economics, 146, 2020.

Edwin A Locke and Gary P Latham. A theory of goal setting & task performance. Prentice-Hall, Inc, 1990.

Anandi Mani, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. Poverty impedes cognitive function. Science, 314(6149):976–980, 2013.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics, 142.2:698–714, 2008.

Ben Ost, Weixiang Pan, and Douglas Webber. The returns to college persistence for marginal students: Regression discontinuity evidence from university dismissal policies. Journal of Labor Economics, 36(3):779–805, 2018.

Zhuan Pei, David S Lee, David Card, and Andrea Weber. Local polynomial order in regression discontinuity designs. Journal of Business & Economic Statistics, pages 1–9, 2021.

David Rapson. Durable goods and long-run electricity demand: Evidence from air conditioner purchase behavior. Journal of Environmental Economics and Management, 68(1):141–160, 2014.

Joachim Schleich. Energy efficient technology adoption in low-income in the European Union - what is the evidence? Energy Policy, 125:196–206, 2019.

Anuj K. Shah, Sendhil Mullainathan, and Eldar Shafir. Some consequences of having too little. Science, 338(6107):682–685, 2012.

Suzanne B Shu and Ayelet Gneezy. Procrastination of enjoyable experiences. Journal of

Marketing Research, 47(5):933–944, 2010.

Yutong Tan and Xianqiang Mao. Assessment of the policy effectiveness of central inspections of environmental protection on improving air quality in China. Journal of Cleaner Production, 288, 2021.

Richard H Thaler and Cass R Sunstein. Nudge. Yale University Press, 2021.

Jiaxing Wang and Shigeru Matsumoto. An economic model of home appliance replacement: Application to refrigerator replacement among Japanese households. Environmental Economics and Policy Studies, pages 1–20, 2021.

9 Appendix

9.1 Tables

Table 6: Results for the McCrary test: cutoff on January 1, 2018

Voucher €150	-1,392.44 (1,818.82)	-379.58 (597.64)	-1,003.81 (771.83)	-285.94 (171.30)	-153.20 (198.25)	-399.03 (266.33)	-319.24 (194.38)
Bin size	50	25	25	25	10	10	10
Bandwidth in days	150	150	100	50	150	100	50

Notes: We conduct the McCrary Test [McCrary, 2008] for different bin sizes and bandwidths around the cutoff where the voucher value changes.

Table 7: Results for the McCrary test: cutoff on February 1, 2019

Voucher €150	901.097 (1,369.933)	527.033 (341.001)	236.981 (353.834)	463.678 (54.090)	183.923 (94.996)	60.653 (107.208)	136.028* (54.090)
Bin size	50	25	25	25	10	10	10
Bandwidth in days	150	150	100	50	150	100	50

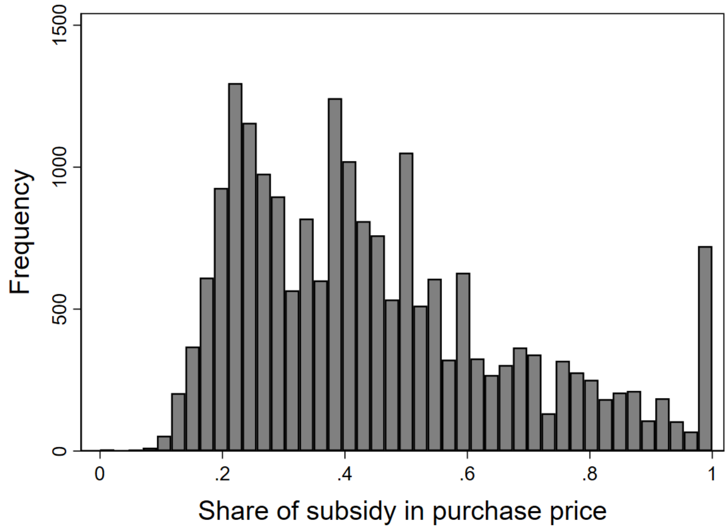
Notes: We conduct the McCrary Test [McCrary, 2008] for different bin sizes and bandwidths around the cutoff where the voucher deadline changes and elective enrolment is introduced.

Table 8: Mean comparison of covariates before and after regime changes

	Procedural variations		
	Mean before Jan 2017 - Dec 2017	Mean after Jan 2018 - Dec 2018	Difference
Household variables			
No. inhabitants	2.886	2.967	-0.083***
Living space in m2	69.830	70.557	-0.727*
Electricity price per kWh	0.276	0.276	-0.0001
Annual electricity consumption in kWh	3,053.508	3,069.962	-16.454
Old refrigerator variables			
Age in years	17.624	18.214	-0.590***
Volume in liters	239.618	245.750	-6.132***
Estimated savings from replacement	329.040	341.715	-12.675***
Price index cooling appliances	95.826	95.370	0.456***
	Subsidy variations		
	Mean before Feb 2018 - Jan 2019	Mean after Feb 2019 - Jan 2020	Difference
Household variables			
No. inhabitants	2.975	3.014	-0.039
Living space in m2	70.662	70.662	-0.0004
Electricity price per kWh	0.276	0.279	-0.003***
Annual electricity consumption in kWh	3,068.515	3,030.006	38.509
Old refrigerator variables			
Age in years	18.241	17.209	1.032***
Volume in liters	245.945	254.001	-8.056***
Estimated savings from replacement	341.438	330.333	11.105***
Price index cooling appliances	95.375	96.321	-0.947***

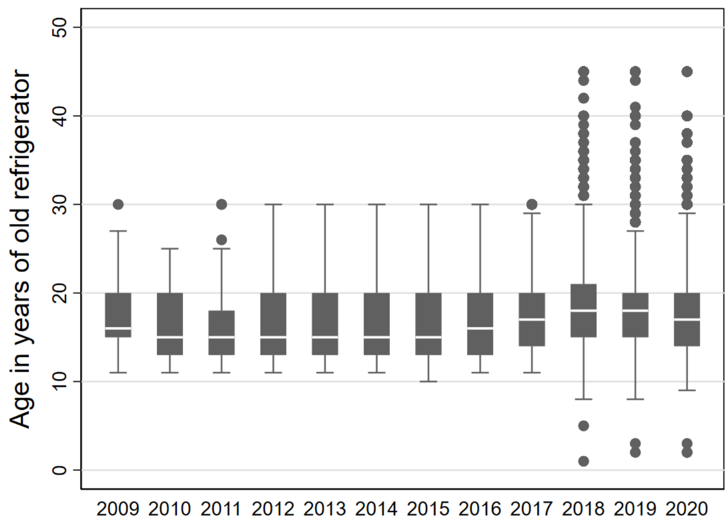
9.2 Figures

Figure 8: Share of subsidy in purchase price of new refrigerator



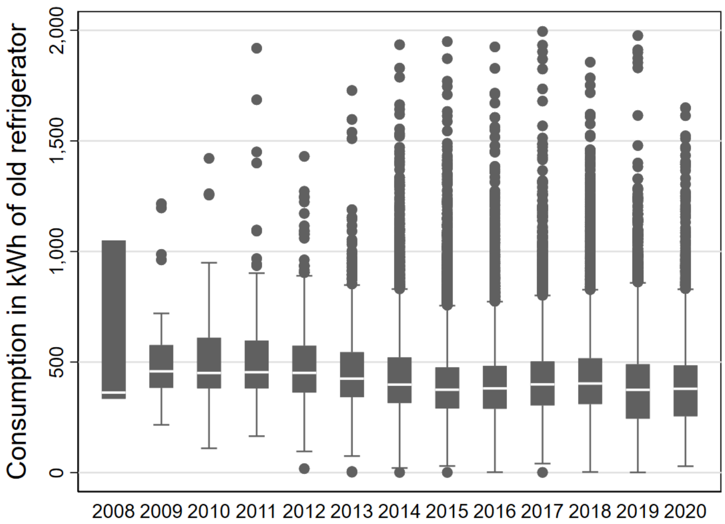
Notes: This figure shows the share that the subsidy on the replacement voucher covers of the total purchase price for the new refrigerator. The subsidies considered here include the federal subsidy of €150 up to 2017 and €100 as of 2018 respectively as well as the complementary programs by several state governments as listed in section 2.

Figure 9: Age structure of old refrigerators over sample period



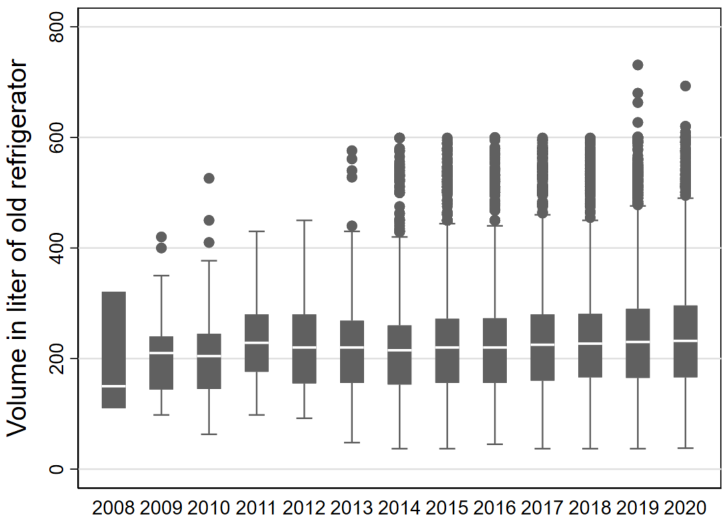
Notes: This figure shows the age distribution of old refrigerators for each year in the sample period. The upper part of the distribution is truncated at 30 years up to 2017, and at 45 years after. The median is relatively stable up to 2018, but drops in 2019. Note that in the years 2018-2020, 13 refrigerators were marked for replacement even though they were younger than 10 years. The figure was created with the sample of for replacement eligible households.

Figure 10: KWh consumption structure of old refrigerators over sample period



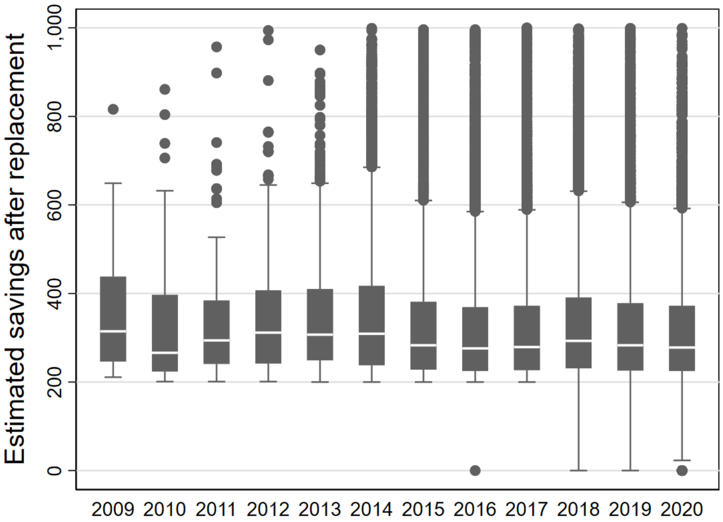
Notes: This figure shows the consumption distribution of old refrigerators in kWh for each year in the sample period. The median decreases relatively steadily over the sample period. Note that a few outliers lie above 2,000 kWh which we omit in the figure due to clarity. The figure was created with the sample of for replacement eligible households.

Figure 11: Volume structure of old refrigerators over sample period



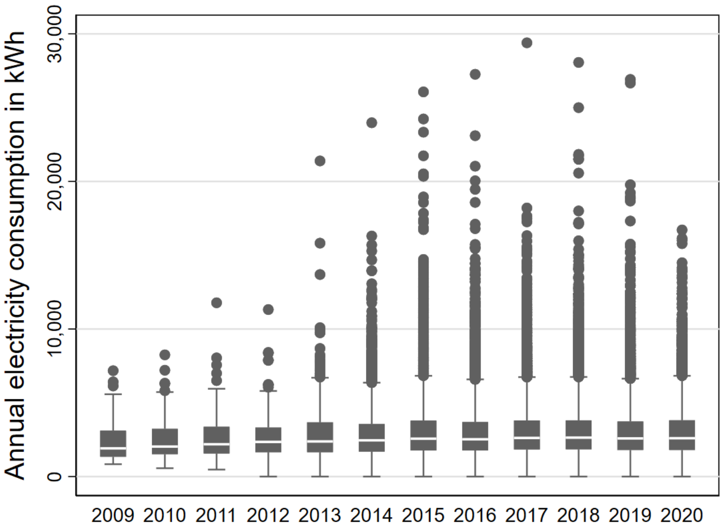
Notes: This figure shows the volume distribution of old refrigerators in liter for each year in the sample period. The median is relatively constant over the later years. The figure was created with the sample of for replacement eligible households.

Figure 12: Estimated savings after replacement over sample period



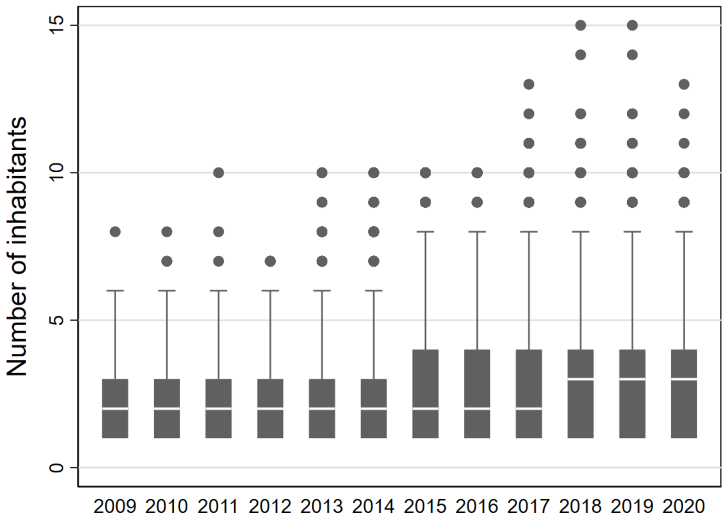
Notes: This figure shows the distribution of estimated savings after replacement for each year in the sample period. The median is relatively constant over the sample period. Note that a few outliers lie above 1,000 kWh which we omit in the figure due to clarity, and that in the years 2016-2020, 19 refrigerators were marked for replacement even though their replacement would have estimatedly saved these household less than 200 kWh annually. The figure was created with the sample of for replacement eligible households.

Figure 13: Annual electricity consumption structure of households over sample period



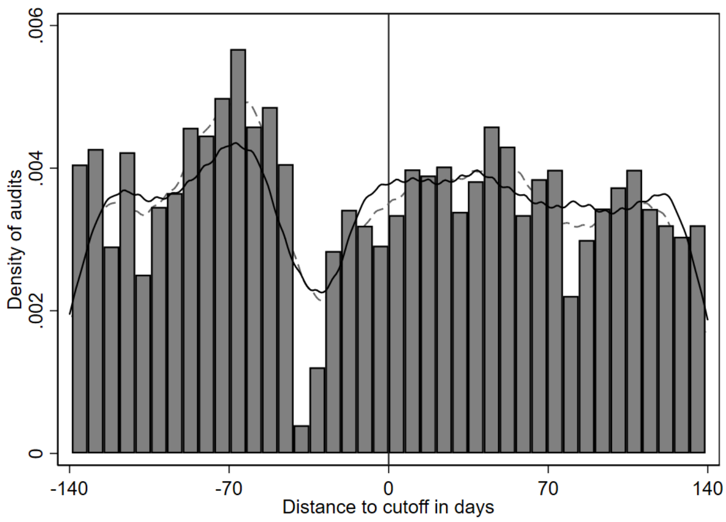
Notes: This figure shows the distribution of the annual electricity consumption of households for each year in the sample period. The median is stable over the entire period. The figure was created with the sample of for replacement eligible households.

Figure 14: Household structure over sample period



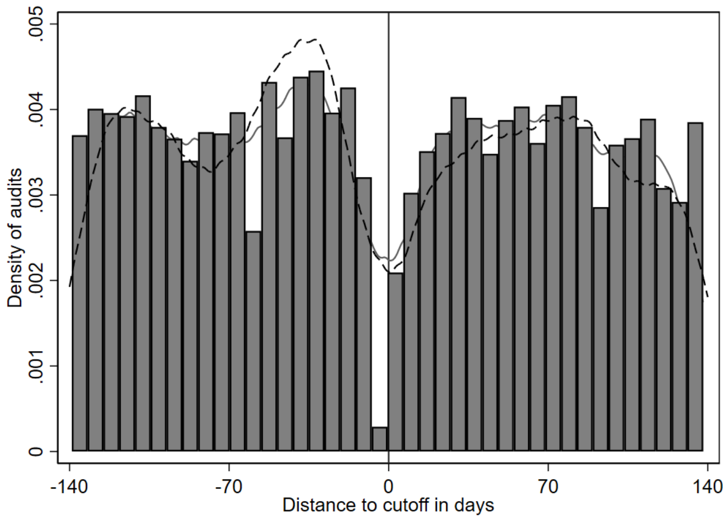
Notes: This figure shows the distribution of the number of inhabitants for each year in the sample period. The upper part of the distribution is truncated at 10 inhabitants up to 2016, and at 15 in the years after. The median is stable over the period 2009-2017, but increases to a higher level in 2018-2020. The figure was created with the sample of for replacement eligible households.

Figure 15: Density around the subsidy variations



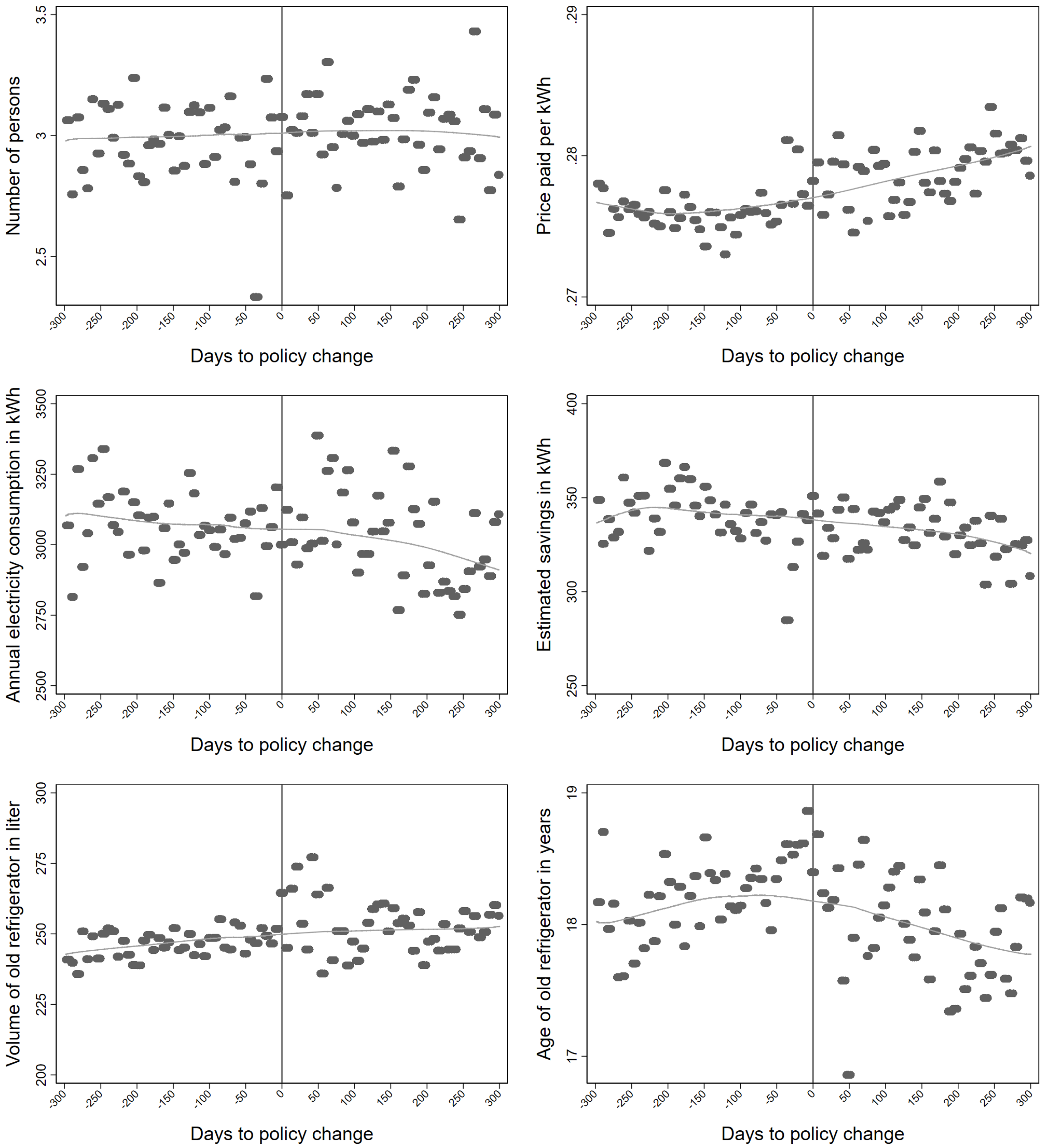
Notes: This figure shows the density and Kernel density (dashed line) of audits (second home visits) in a bandwidth of 20 weeks around the regime change. No bunching is apparent on either side of the threshold (we would expect bunching to occur on the left side if households wanted to get one of the last vouchers with €150 value). A bump appears 6 to 5 weeks before the regime change which coincides with the Christmas and end-of-year break when most local branches close for one or two weeks. To demonstrate that this pattern is usual we also provide the Kernel density of audits in the year before during the same season (solid line). Both Kernel densities are almost perfectly aligned.

Figure 16: Density around the procedural variations



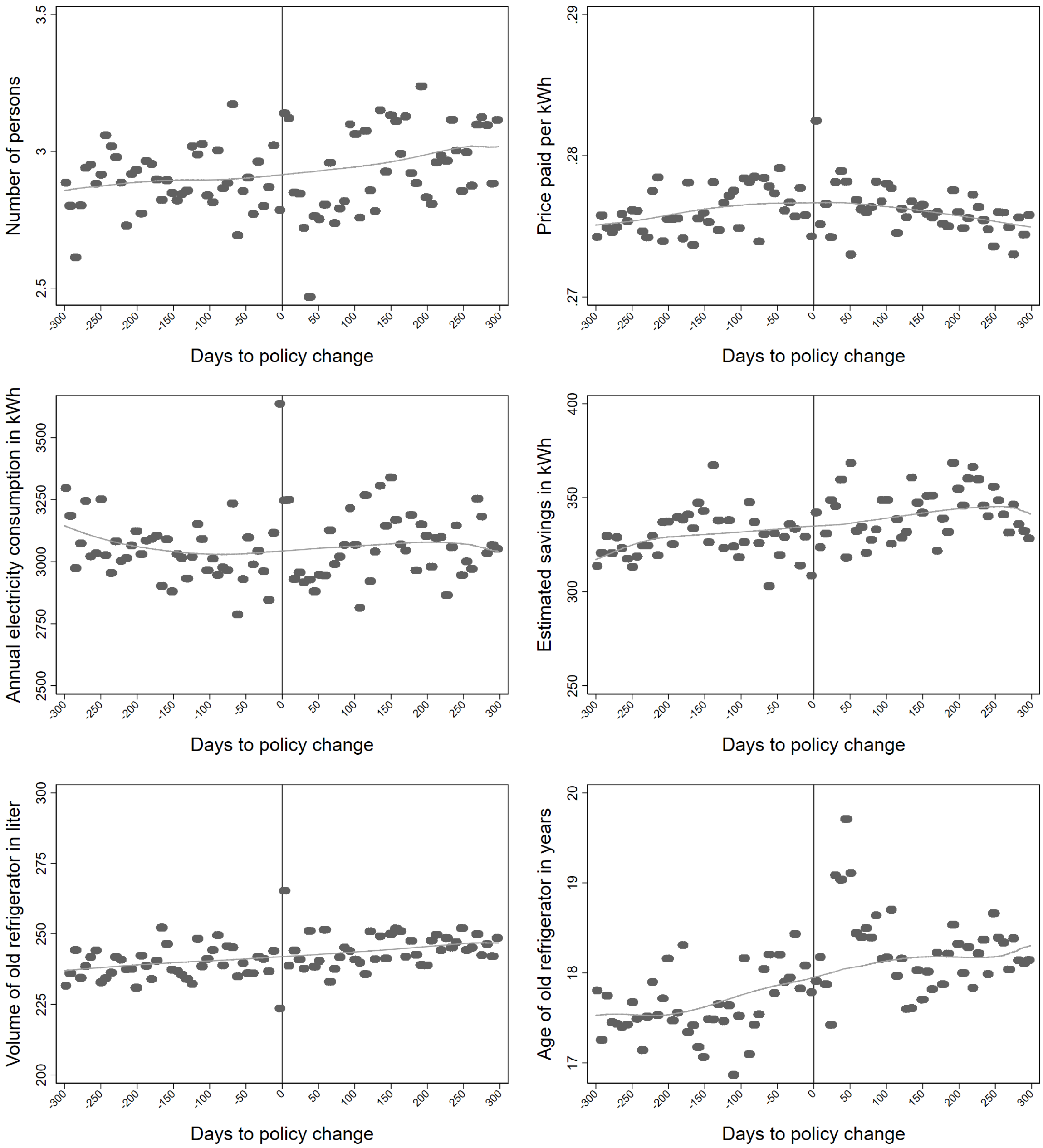
Notes: This figure shows the density and Kernel density (solid line) of audits (second home visits) in a bandwidth of 20 weeks around the regime change. The pattern is symmetric on either side of the threshold (we do not expect bunching to occur; the more attractive extendable vouchers had a definite deadline set on the day the regime changed to vouchers with a strict deadline so that no additional incentive was present on either side of the threshold). A bump appears directly before the regime change which coincides with the Christmas and end-of-year break when most local branches close for one or two weeks. To demonstrate that this pattern is usual we also provide the Kernel density of audits in the year after during the same season (dashed line). Both Kernel densities are almost perfectly aligned.

Figure 17: Discontinuity check for observables around the pecuniary inducement change



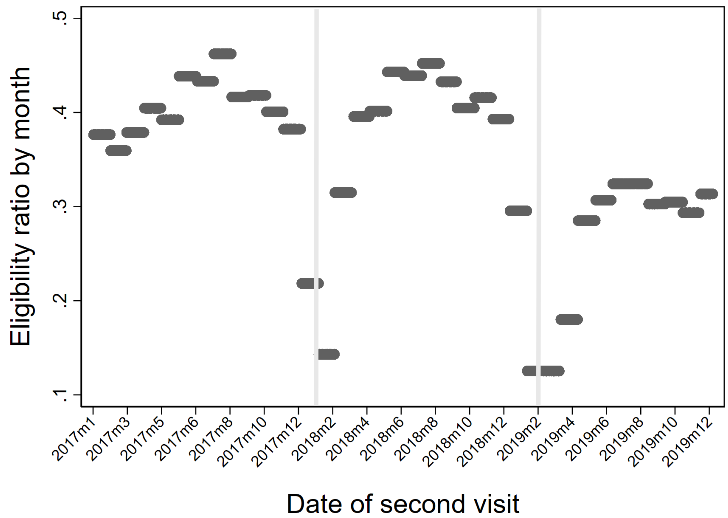
Notes: The figures show weekly averages of household observables in a bandwidth of 300 days around the change in the voucher value, and a locally weighted regression through the individual data points. There is no evidence for a systematic discontinuity at the point where policy design changes.

Figure 18: Discontinuity check for observables around the procedural variations



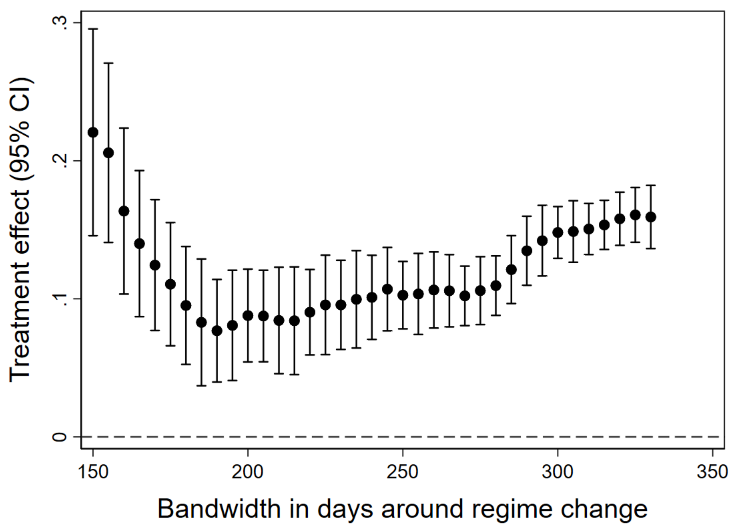
Notes: The figures show weekly averages of household observables in a bandwidth of 300 days around the change in the voucher redemption conditions (validity and extension), and a locally weighted regression through the individual data points. There is no evidence for a systematic discontinuity at the point where policy design changes.

Figure 19: Eligibility ratio around both design changes



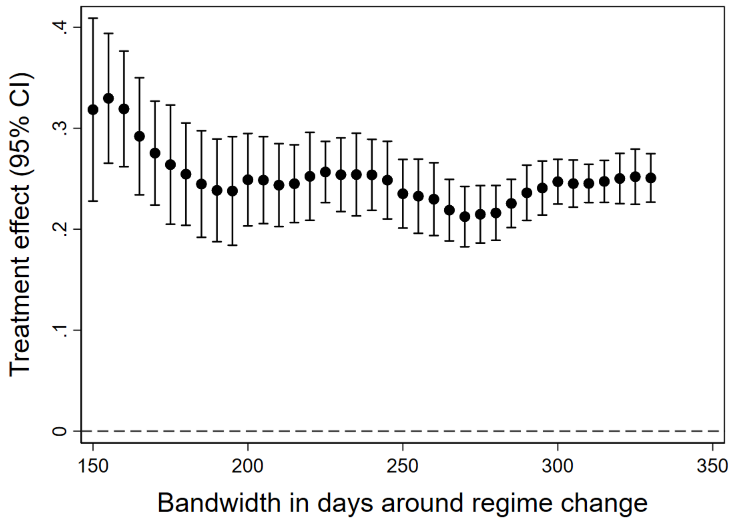
Notes: This figure shows the monthly ratio of households who are found eligible for replacement and receive an information letter out of all audited households. Around both regime changes (introduction of strict voucher deadline and reduction of voucher value), the eligibility ratio drops considerably. In the data, we see that this is not due to fewer households whose refrigerators fulfill the criteria for replacement (older than 10 years, annual savings of at least 200 kWh). Instead we observe that not all households who fulfill the criteria receive an information letter or voucher which enables them to join the program. This pattern may origin in irregularities in the program process due to the introduction of the infoletter at the first regime change and due to the end and start of a new funding phase at the second regime change.

Figure 20: Pecuniary effect (replacement rate) as function of bandwidth



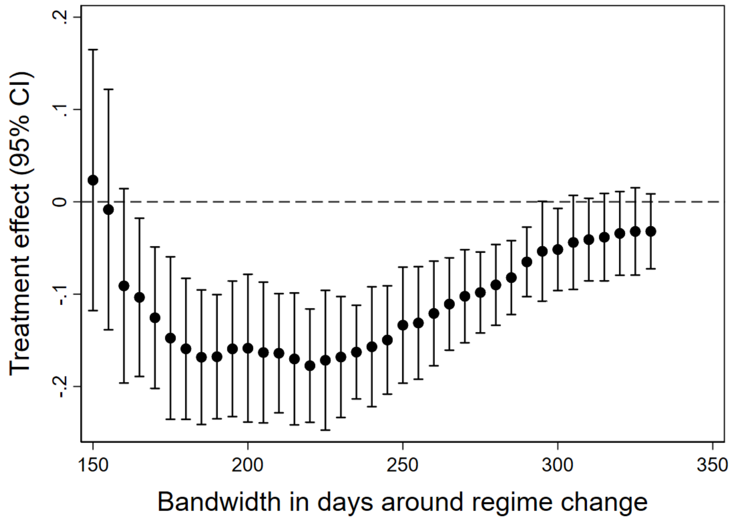
Notes: This figure shows the pecuniary effect for different bandwidth choices.

Figure 21: Pecuniary effect (enrolment rate) as function of bandwidth



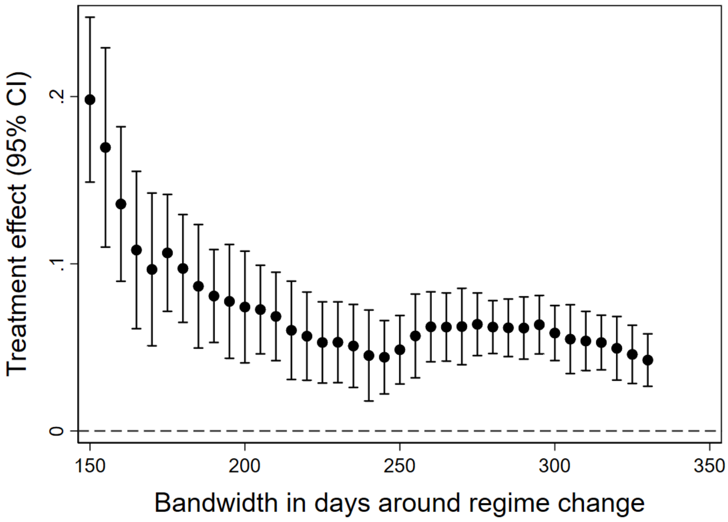
Notes: This figure shows the pecuniary effect for different bandwidth choices.

Figure 22: Pecuniary effect (redemption rate) as function of bandwidth



Notes: This figure shows the pecuniary effect for different bandwidth choices.

Figure 23: Effect of procedural variations (replacement rate) as function of bandwidth

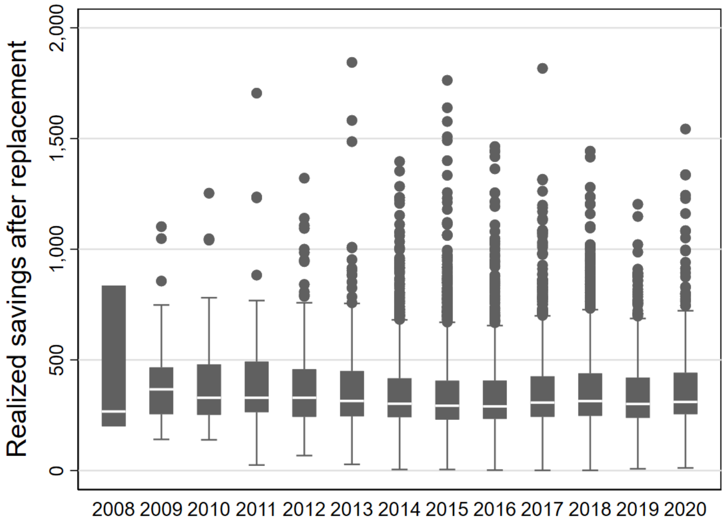


Notes: This figure shows the pecuniary effect for different bandwidth choices.

Figure 24: Redemption propensity out of enrolled households in EE-RIG

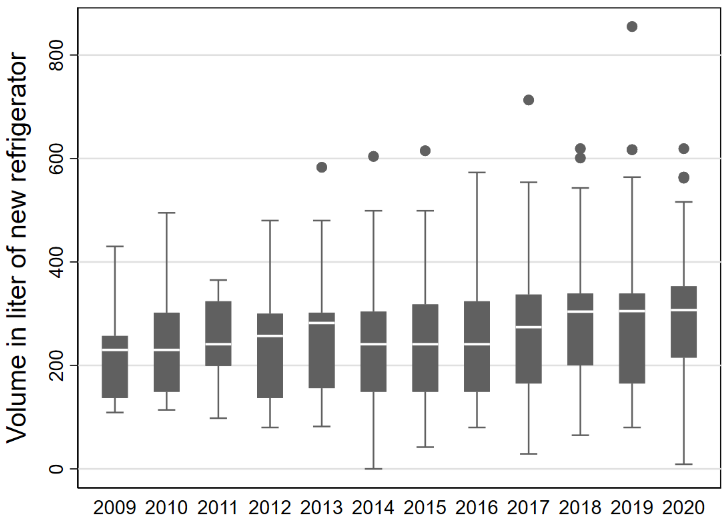
Notes: This figure shows the propensity of enrolled households to redeem the voucher as function of the days passed since the voucher was generated.

Figure 25: Realized savings after replacement over sample period



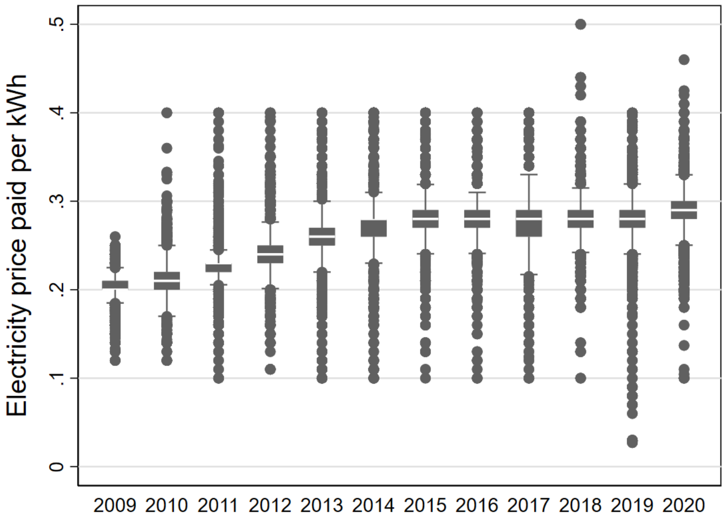
Notes: This figure shows the distribution of realized savings after replacement (different from estimated savings *before* replacement) for each year in the sample period. The median is relatively stable over the entire period. The figure was created with the sample of households that replaced their refrigerator.

Figure 26: Volume of new refrigerator over sample period



Notes: This figure shows the distribution of the volume in liters of the new efficient refrigerator that households purchase as replacement for the old inefficient one for each year in the sample period. The median trends towards a higher volume over time.

Figure 27: Electricity price paid per kWh over sample period



Notes: This figure shows the distribution of the price per kWh paid for electricity. The median increases significantly over time mirroring a general rise in electricity prices in Germany during the entire sample period. The figure was created with the sample of all audited households.

9.3 Robustness checks

Table 9: Robustness checks for the pecuniary effect on the replacement rate

	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
Voucher €150	0.058 (0.031)	0.105*** (0.012)	0.119*** (0.028)	0.212*** (0.010)	0.100*** (0.018)	0.152*** (0.009)	0.108*** (0.011)	0.146*** (0.008)	0.081* (0.039)	0.124*** (0.032)	0.086 (0.076)	0.136** (0.048)	0.267* (0.114)	0.397*** (0.101)	0.318* (0.124)	0.385*** (0.103)
Daycount	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Bandwidth	6	11	6	11	6	11	6	11	6	11	6	11	6	11	6	11
Donut (months)	2	2	2	2	1	1	0	0	2	2	2	2	2	2	2	2
Augmented Local Linear	yes	yes	yes	yes	yes	yes	yes	yes								
Model/Estimation	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	BPM	BPM	BPM	BPM
No. observations	3,562	8,122	3,562	7,279	7,804	16,466	8,645	17,307	7,628	17,478	5,774	14,270	7,628	17,478	7,539	17,416
Full sample					yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Non-NRW only	yes	yes														
NRW only			yes	yes												

Note: Controls include the number of inhabitants, living space, the type of federal income support received, kWh price paid, yearly electricity consumption, old appliance age, appliance price index and estimated replacement savings. The Augmented Local Linear approach uses month, month-by-year, year, state, branch, branch-by-year, ZIP code and auditor controls. XV-XVI include state, branch, ZIP code, and auditor FE. XIX-XX include state and branch FE. The binary probability model (BPM) is estimated as probit. Standard errors are bootstrapped for V-XII and clustered by branch for XIII-XX. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. Estimated on the sample of eligible households around February 1, 2019.

Table 10: Robustness checks for the pecuniary effect on the enrolment rate

	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
Voucher €150	0.235*** (0.039)	0.217*** (0.016)	0.262*** (0.025)	0.281*** (0.015)	0.357*** (0.018)	0.303*** (0.011)	0.258*** (0.017)	0.271*** (0.008)	0.268*** (0.054)	0.244*** (0.038)	0.243** (0.087)	0.217*** (0.048)	0.700*** (0.139)	0.653*** (0.100)	0.761*** (0.142)	0.675*** (0.101)
Daycount	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Bandwidth	6	11	6	11	6	11	6	11	6	11	6	11	6	11	6	11
Donut (months)	2	2	2	2	1	1	0	0	2	2	2	2	2	2	2	2
Augmented Local Linear	yes	yes	yes	yes	yes	yes	yes	yes								
Model/Estimation	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	BPM	BPM	BPM	BPM
No. observations	3,562	8,122	3,177	7,279	7,804	16,466	8,645	17,307	7,628	17,478	5,774	14,270	7,628	17,478	7,511	17,384
Full sample					yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Non-NRW only	yes	yes														
NRW only			yes	yes												

Note: Controls include the number of inhabitants, living space, the type of federal income support received, kWh price paid, yearly electricity consumption, old appliance age, appliance price index and estimated replacement savings. The Augmented Local Linear approach uses month, month-by-year, year, state, branch, branch-by-year, ZIP code and auditor controls. XV-XVI include state, branch, ZIP code, and auditor FE. XIX-XX include state and branch FE. The binary probability model (BPM) is estimated as probit. Standard errors are bootstrapped for V-XII and clustered by branch for XIII-XX. * p<0.05, ** p<0.01, *** p<0.001. Estimated on the sample of eligible households around February 1, 2019.

Table 11: Robustness checks for the pecuniary effect on the redemption rate

	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
Voucher €150	-0.181** (0.054)	-0.093** (0.030)	-0.142** (0.042)	0.008 (0.026)	-0.261*** (0.034)	-0.126*** (0.020)	-0.121*** (0.026)	-0.087*** (0.016)	-0.200* (0.085)	-0.107 (0.058)	-0.039 (0.173)	-0.023 (0.127)	-0.568* (0.233)	-0.311 (0.164)	-0.444 (0.238)	-0.280 (0.161)
Daycount	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Bandwidth	6	11	6	11	6	11	6	11	6	11	6	11	6	11	6	11
Donut (months)	2	2	2	2	1	1	0	0	2	2	2	2	2	2	2	2
Augmented Local Linear	yes	yes	yes	yes	yes	yes	yes	yes								
Model/Estimation	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	BPM	BPM	BPM	BPM
No. observations	1,288	2,724	1,329	2,892	3,212	6,211	3,593	6,592	3,087	6,636	1,660	4,402	3,087	6,636	3,011	6,561
Full sample					yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Non-NRW only	yes	yes														
NRW only			yes	yes												

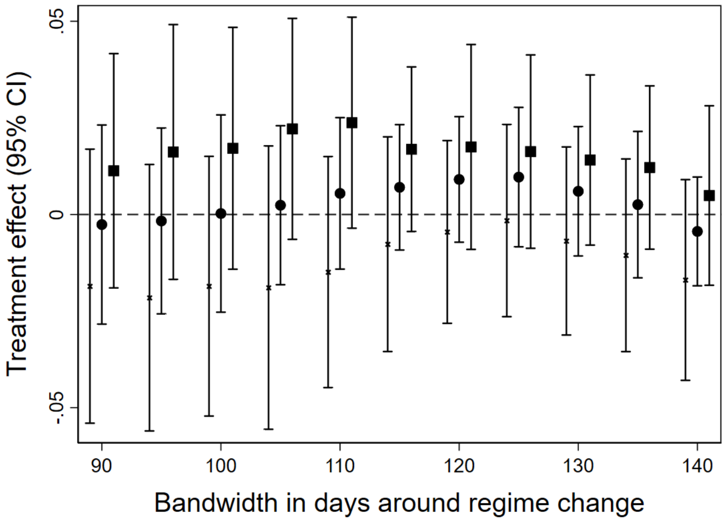
Note: Controls include the number of inhabitants, living space, the type of federal income support received, kWh price paid, yearly electricity consumption, old appliance age, appliance price index and estimated replacement savings. The Augmented Local Linear approach uses month, month-by-year, year, state, branch, branch-by-year, ZIP code and auditor controls. XV-XVI include state, branch, ZIP code, and auditor FE. XIX-XX include state and branch FE. The binary probability model (BPM) is estimated as probit. Standard errors are bootstrapped for V-XII and clustered by branch for XIII-XX. * p<0.05, ** p<0.01, *** p<0.001. Estimated on the sample of households that have requested a voucher around February 1, 2019.

Table 12: Robustness checks for effect of the procedural variations on the replacement rate

	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
Strict voucher deadline	0.083** (0.027)	0.033** (0.011)	0.111*** (0.024)	0.040** (0.012)	0.265*** (0.015)	0.128*** (0.008)	0.225*** (0.010)	0.136*** (0.008)	0.095* (0.039)	0.043* (0.018)	0.126* (0.057)	0.048 (0.025)	0.289* (0.120)	0.130* (0.055)	0.331* (0.122)	0.155** (0.059)
Daycount	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Bandwidth	6	11	6	11	6	11	6	11	6	11	6	11	6	11	6	11
Donut (months)	2	2	2	2	1	1	0	0	2	2	2	2	2	2	2	2
Augmented Local Linear	yes	yes	yes	yes	yes	yes	yes	yes								
Model/Estimation	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	LPM	BPM	BPM	BPM	BPM
No. observations	4,632	10,285	3,907	8,889	10,071	20,706	11,363	21,998	9,824	22,003	7,484	18,012	9,824	22,003	9,726	21,913
Full sample					yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Non-NRW only	yes	yes														
NRW only			yes	yes												

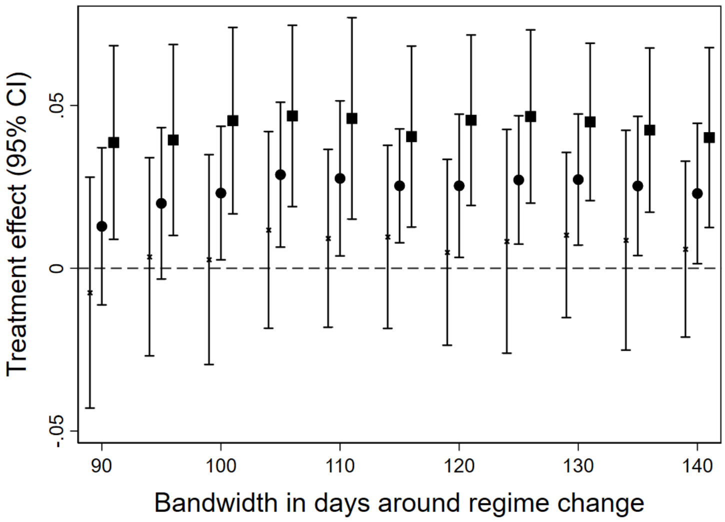
Note: Controls include the number of inhabitants, living space, the type of federal income support received, kWh price paid, yearly electricity consumption, old appliance age, appliance price index and estimated replacement savings. The Augmented Local Linear approach uses month, month-by-year, year, state, branch, branch-by-year, ZIP code and auditor controls. XV-XVI include state, branch, ZIP code, and auditor FE. XIX-XX include state and branch FE. The binary probability model (BPM) is estimated as probit. Standard errors are bootstrapped for V-XII and clustered by branch for XIII-XX. * p<0.05, ** p<0.01, *** p<0.001. Estimated on the sample of eligible households around January 1, 2018.

Figure 28: Treatment effect for placebo inducement on July 15, 2018 as function of bandwidth



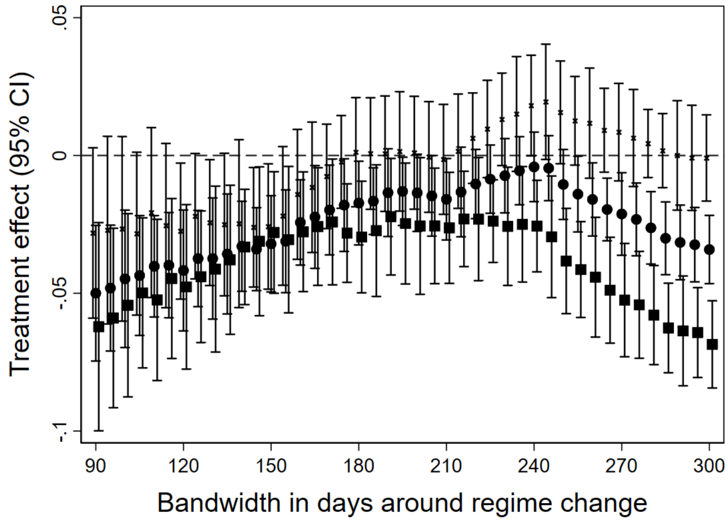
Notes: This figure shows the treatment effect for a placebo inducement on July 15, 2018, halfway between the variations in subsidy levels and procedures for different bandwidth choices. The dot represents the point estimate for the full sample, the square represents the sample of only households in North Rhine-Westphalia, and the x represents the sample of households not from North Rhine-Westphalia. None of the coefficients is significant to the 5% confidence level. The point estimates vary for the different samples as funding in the NRW top-up state program differs from the federal program only and other less comprehensive state programs, and funding in the NRW program is not as stable as in the federal program.

Figure 29: Treatment effect for placebo inducement on August 15, 2019 as function of bandwidth



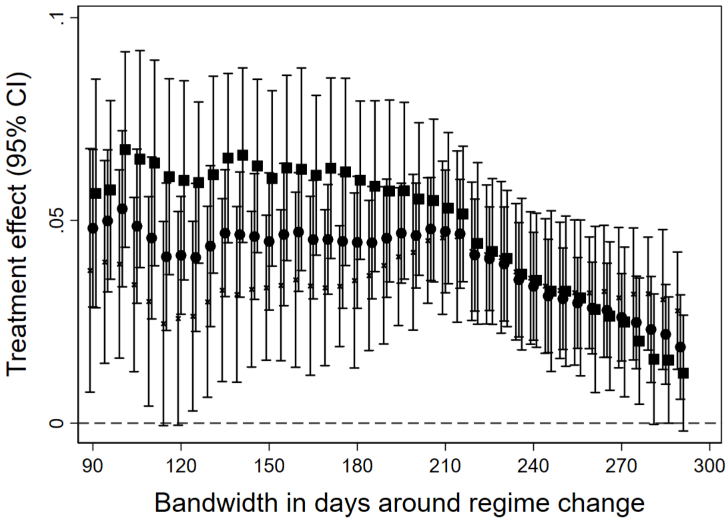
Notes: This figure shows the treatment effect for a placebo inducement on July 15, 2018, halfway between the variations in subsidy levels and procedures for different bandwidth choices. The dot represents the point estimate for the full sample, the square represents the sample of only households in North Rhine-Westphalia, and the x represents the sample of households not from North Rhine-Westphalia. Coefficients for the full sample and the NRW-only sample are significantly different from zero at the 5% level. However, the coefficient on the sample of non-NRW households is insignificant. The point estimates vary for the different samples as funding in the NRW top-up state program differs from the federal program only and other less comprehensive state programs, and funding in the NRW program is not as stable as in the federal program.

Figure 30: Treatment effect for placebo inducement on January 1, 2017 as function of bandwidth



Notes: This figure shows the treatment effect for a placebo inducement on July 15, 2018, halfway between the changes in the voucher value and the procedures for different bandwidth choices. The dot represents the point estimate for the full sample, the square represents the sample of only households in North Rhine-Westphalia, and the x represents the sample of households not from North Rhine-Westphalia. Coefficients for the full sample and the NRW-only sample are significantly different from zero at the 5% level. However, the coefficient on the sample of non-NRW households is insignificant. The point estimates vary for the different samples as funding in the NRW top-up state program differs from the federal program only and other less comprehensive state programs, and funding in the NRW program is not as stable as in the federal program.

Figure 31: Treatment effect for placebo inducement on February 1, 2017 as function of bandwidth



Notes: This figure shows the treatment effect for a placebo inducement on July 15, 2018, halfway between the changes in the voucher values and the procedures for different bandwidth choices. The dot represents the point estimate for the full sample, the square represents the sample of only households in North Rhine-Westphalia, and the x represents the sample of households not from North Rhine-Westphalia. Coefficients for the full sample and the NRW-only sample are significantly different from zero at the 5% level. However, the coefficient on the sample of non-NRW households is insignificant. The point estimates vary for the different samples as funding in the NRW top-up state program differs from the federal program only and other less comprehensive state programs, and funding in the NRW program is not as stable as in the federal program.