Essays on the public economics of decarbonization: Low-income households and industry

Dissertation zur Erlangung des akademischen Grades
Doctor rerum politicarum

Vorgelegt an der Fakultät für Wirtschafts- und Sozialwissenschaften
der Ruprecht-Karls-Universität Heidelberg

von Bettina Larissa Chlond

Heidelberg, April 2025

Acknowledgement

The support from numerous people over the past five years has been invaluable while working on this dissertation. Without my co-authors, supervisors, mentors, and colleagues, this work would not have been possible.

I thank my supervisor, Timo Goeschl, for all his guidance and support throughout. His (Zoom) door was always open and he always provided me with the advice needed on research and on navigating the academic world.

I am also grateful to Martin Kesternich, who has been my mentor since my very first day at ZEW. Being part of his group at ZEW, I benefited greatly from his encouragement and trust in my abilities.

Moreover, I thank Anca Balietti, Sébastien Houde and Kathia Serrano Velarde for agreeing to read and evaluate this dissertation and providing me with valuable feedback on the papers.

I would also like to thank our partners at Stromspar-Check, in particular, Christian Müller at Caritas Frankfurt, for the trusted and fruitful collaboration which is continuing. I am grateful to Kathrine von Graevenitz for all her advice and pep talks during coffee breaks and walks around the Mannheim palace. I thank Erica Myers for her extensive feedback on my job market paper and the generous amount of time she dedicated to supervising me during my research stays at the University of Calgary.

My colleagues at the Environmental and Climate Economics department at ZEW and at the Research Center for Environmental Economics at Heidelberg University made my work days enjoyable and provided me with invaluable comments on my research. I am especially thankful to my co-authors Robert Germeshausen and Madeline Werthschulte. Moreover, I would like to thank Lara Bartels, my office mate at ZEW, for all her encouragement and positivity she used to bring to our office. I am grateful to Anna Straubinger for her support, kind words and honest feedback.

Finally, I am grateful to my parents whose support made my academic development possible. I thank my family and friends for their support, distraction from thesis writing, and, most importantly, reminding me of the non-economist perspective on the world.

Table of Contents

Chapter 1

Introduction

1.1 Motivation

Climate change is "the defining issue of our time" (United Nations, 2025) and one of mankind's greatest challenges. If global human-made greenhouse gas emissions exceed the IPCC carbon budget, global warming beyond 2°C above pre-industrial levels will be inevitable, leading to widespread adverse effects on nature, people and the economy (IPCC, 2025). However, despite extensive knowledge about human-induced atmospheric processes and the availability of viable mitigation technologies in most sectors, emissions continue to rise. Economists contribute crucially to addressing this paradox and tackling challenges related to climate change. First, they offer explanations for why individual and institutional action often does not reflect the urgency of the climate crisis, despite widespread awareness. Second, economists provide evidence on suitable policy instruments and evaluate and improve existing instruments that can accelerate the decarbonization of economies. Concepts from environmental and behavioral economics help explain why the challenges that climate change poses remain insufficiently addressed. A key idea from environmental economics is externalities, which offers an explanation for market failures, such as over-pollution and the depletion of carbon budgets (Marshall, 1890; Pigou, 1920). Market actors base decisions mainly on their individual costs and benefits. In unregulated markets individual persons' and countries' incentives are not always aligned with the required actions that would maximize (global) societal welfare. However, many economic activities have an impact on other individuals besides the initial consumer. Failing to internalize these externalities can lead to market failure that results in over- or under-consumption. An example for externalities is car ownership and usage. Individuals besides the consumer in the nearer and farther environment are negatively affected via wear and tear of public roads, local pollutants, congestion, and the risk of accidents. Additionally, CO2 emissions from car production and usage contribute to climate change, affecting people globally. Without regulation that internalizes these externalities, for instance, via pricing, many consumers will not consider the impact of their activities on others, and car usage will exceed the social optimum.

The theory of bounded rationality from behavioral economics can be used to understand behavioral barriers to decarbonization (Simon, 1955). Individuals make choices under cognitive limitations and limited information, deviating from the model of homo oeconomicus in neoclassical theory. Instead of rationally weighing all options, individuals use heuristics as simplified decision rules. Biases arise if heuristics introduce systematic and predictable errors into decision making (Tversky and Kahneman, 1974). One particularly relevant bias for decarbonization is hyperbolic discounting, also known as present bias. The term describes the phenomenon that individuals irrationally disproportionately discount the future. While the discount rate between any two periods in the future remains constant, the rate for discounting between present and future is higher. This results in time-inconsistent preferences and behaviors such as procrastination (Laibson, 1997). In the context of climate change mitigation, hyperbolic discounting is relevant in decisions to invest in energy-efficient technologies. For instance, a present-biased homeowner considering to invest in home insulation to increase the energy efficiency and lower the heating cost of her dwelling, will focus on the high initial investment cost, while excessively discounting future savings in heating cost. Other behavioral biases that potentially affect investment in efficient technologies are inattention to energy cost and loss aversion (Palmer and Walls, 2015; Gerarden et al., 2017). At the economy-wide level, these biases can be a relevant factor in slowing down the energy transition and lead to a need for policy interventions that account of biases.

As a contribution to global efforts to mitigate climate change, the EU has pledged to reduce greenhouse gas emissions by 55% by 2030 against 1990 levels and achieve net-zero emissions by 2050 (European Commission, 2023). One key sector to be decarbonized is the residential sector. In 2024, activities of households accounted for 18 percent of total greenhouse gas emissions in the EU (Eurostat, 2024). To accelerate decarbonization in this sector, a focus must be on closing the energy efficiency gap, the difference between the optimal and realized level of energy efficiency (Jaffe and Stavins, 1994). Investments in energy-efficient and renewable energy technologies offer long-term emissions reductions. In contrast, behavioral adjustments in consumption decisions tend to be short-term and, therefore, have a more limited impact on emissions (Brandon et al., 2017). Policy instruments to accelerate the energy transition will have to encourage a large volume of decentralized investments. To reduce energy demand, households need to improve the energy efficiency of their dwellings. To decarbonize energy use, homeowners can install PV solar panels and heat pumps.

The political debate about which instruments are best suited to achieve substantial emissions reductions in households has intensified regarding the objectives of cost-effectiveness and equity. Scientific contributions from the field of public economics aim to create an understanding of how instruments can be used to address market distortions, such as externalities and behavioral bias. This research informs the optimal design of instruments to correct economic decisions of actors given certain policy objectives. The public economics of decarbonization in the household sector studies the optimal design of policy instruments to reduce greenhouse gas emissions and accelerate the energy transition. In practice, this literature investigates which instruments can effectively encourage sustainable investment decisions, for instance, how to incentivize households to replace their appliance, car or heating system with energy-efficient alternatives.

Most policy instruments studied in the public economics of decarbonization feature two key mechanisms that affect individual behavior: monetary and behavioral incentives. Monetary incentives are central to many policy instruments, such as taxes and subsidies. Pigouvian taxes aim to address externalities to the environment by pricing pollution and emissions at their marginal social cost. By increasing prices, these taxes incentivize individuals to take into account the cost of their behavior to social welfare, thereby internalizing externalities. Research on the monetary aspects of policy instruments can be organized around consumption and investment subsidies and taxes. Important criteria in the evaluation of monetary instruments are their cost-effectiveness and distributional impacts. Examples for policies targeting consumption include carbon taxes and emission trading schemes (see Döbbeling-Hildebrandt et al., 2024, for a review of the literature). Instruments targeting investment are direct lump-sum payments at purchase (e.g., Fowlie et al., 2018) and indirect financial incentives, such as interest-free loans or tax credits (e.g., Hassett and Metcalf, 1995; Eryzhenskiy et al., 2023).

Behavioral incentives in policy design can be categorized in two types: the intentional purposeful utilization of behavioral effects for policy design and unintended incidental incorporation of behavioral features into policy instruments. Intentional behavioral incentives come in the form of behavioral interventions ("nudges") and information policies. A well-known example of nudging is the use of social comparisons in "Home Energy Reports", which gives consumers feedback on their energy consumption relative to their neighbors (Allcott, 2011). Information instruments include the use of labelling, for example, concerning energy and food choices, to provide easily understandable and simplified information on a product (e.g., Houde, 2018b). Unintended behavioral effects inherent in the design of public policies come in different forms. The design of public assistance programs can have unintended effects on program outcomes, such as informational complexity deterring low-income households from claiming cash transfers (e.g., Bhargava and Manoli, 2015). Other examples include the design of tax policy, such as the asymmetric impacts of a positive or negative balance of withholding taxes on filing returns due to loss aversion (e.g., Engström et al., 2015) or the moral attribute inherent in 'sin' taxes and 'virtue' subsidies (e.g., Andersson, 2019).

An important question in the public economics of decarbonization is which instruments effectively increase energy efficiency in low-income households. In recent years, European policymakers have prioritized addressing the disproportionate burden from high energy cost on low-income households. This issue has been exacerbated by rising energy prices in the last years, driven mainly by the energy crisis caused by the Russian invasion of Ukraine, and increasing carbon prices. Low-income households are especially vulnerable to high and rising energy cost for at least three reasons. First, they spend a larger share of their income on energy expenditures than wealthier households, which increases their exposure to rising prices. Second, their energy use tends to be less efficient (Andor et al., 2021). And third, low-income households tend to be less able to respond to energy price increases (Frondel et al., 2019). In 2024, 11 percent of Europeans were unable to adequately heat their homes (European Commission, 2024). Recognizing energy poverty as a serious policy concern, policymakers are seeking effective instruments to support low-income households in becoming more efficient in their energy use. The EU Energy Efficiency Directive calls for measures that reduce energy consumption and energy bills as central strategies for addressing energy poverty (European Union, 2023). According to the directive, suitable instruments to this objective include access to grants and subsidies, publicly supported energy consumption assessments and directed schemes that provide better information, and technical and administrative assistance.

Despite policymakers' demand for suitable instruments, few economic studies examine how effective policies tailored to low-income households should be designed. Both monetary and behavioral barriers could contribute to lower energy efficiency in these households, suggesting that policy design should incorporate both types of incentives. Pigouvian instruments that use monetary incentives could be relevant to encourage a higher level of investment for efficient technologies. Low-income households often face budget and credit constraints, making the high upfront cost of efficient technologies difficult to afford. In these cases, monetary incentives could help to finance the investment cost. Behavioral incentives may also be relevant in the design of policies, as energy operating cost that accumulates in many small payments over the lifetime of a durable is not salient at the time of purchase. Additionally, individuals whose decisions are subject to behavioral biases may not rationally value future energy savings from energy-efficient investments. Behavioral factors may play a more prominent role in investment decisions by low-income households than by wealthier consumers due to two reasons. First, financial literacy, including energy-related knowledge, tends to be lower with lower income (Calvet et al., 2009; Lusardi and Mitchell, 2014; Blasch et al., 2021; Brent and Ward, 2018). Second, experiencing financial stress can capture cognitive capacity, diverting low-income individuals' focus away from long-term investment decisions (Shah et al., 2012; Haushofer and Fehr, 2014; Ong et al., 2019).

The literature on energy efficiency in low-income households in industrialized economies includes studies on the level of efficiency, energy price elasticities and assistance programs.[1] Most studies conducted in the general population include income as a covariate, allowing for heterogeneity analyses, comparing results for higher- versus lower-income subjects. It is found that the level of energy efficiency in low-income households is generally lower than that in high-income households (Andor et al., 2021). Low-income households replace energy-using durables less frequently so that their appliances and heating systems are older and less efficient (Ameli and Brandt, 2015; Schleich, 2019). This may explain why low-income households adjust electricity consumption less in response to changes in electricity prices (Frondel et al., 2019; Alberini et al., 2011). However, they are found to be more price-elastic in their consumption of natural gas (Hahn and Metcalfe, 2021; Rubin and Auffhammer, 2024), suggesting a capacity for behavioral adjustments. Literature explicitly focused on energy efficiency in low-income households remains scarce. Existing studies mainly investigate energy efficiency assistance programs and their effectiveness in reducing participants' energy bills. Most of the programs studied provide home energy audits that assess the thermal efficiency of dwellings and the implementation of retrofit works. Typically, such programs cover a significant portion or the full investment cost for retrofit works which pass a cost-benefit test. The main eligibility criterion is typically the income level of applicants. Fowlie et al. (2015, 2018) conduct a large-scale field experiment in Michigan's Weatherization Assistance Program, which offers free-of-charge retrofitting to low-income households living in poorly insulated or uninsulated dwellings. Pre-intervention participation rates in the program are very low at 2 percent. Using an encouragement design, the researchers increase the participation rate by 4 percentage points. High non-monetary cost in the form of barriers resulting from the necessary efforts of application procedures and hassle from the disruption of retrofit works, may explain the low take-up rates (Fowlie et al., 2015). Program investments reduce energy consumption in participating households by 10 to 20 percent. However, as model-projected energy savings strongly overestimate the realized savings, reductions in energy consumption are not cost-effective. On average, investments in the program are neither privately nor socially beneficial, yielding a rate of return around -2.3% (Fowlie et al., 2018). Christensen et al. (2023) use data from the Weatherization Assistance Program in Illinois to investigate factors to explain the large gap between ex-ante estimated and ex-post realized energy savings. Bias in the predictions of engineering

[1] The literature cited here exclusively focuses on the Global North. As Dubois and Sinea (2023) note, (energy-)poor households in the Global South differ in their characteristics and the challenges they face.

models, particularly for wall insulation, and heterogeneity in workmanship conducting the retrofit works each account for more than 40 percent of the gap. The rebound effect, increases in consumption due to reduced cost of energy services, accounts for 6 percent. A systematic review of energy efficiency retrofit studies finds programs serving low-income households to achieve higher reported energy savings and lower cost per unit of energy saved than programs targeting the general population (Giandomenico et al., 2022). On average, retrofits reduce energy consumption by 13.2 percent, more than four times the reduction compared to programs for higher-income households. As compared to energy efficiency assistance programs supporting investment, California's CARE program provides consumption-based subsidies for electricity and natural gas to low-income households. Hahn and Metcalfe (2021) conduct a welfare analysis of the program and find it to reduce societal welfare via increases in energy use and emissions as well as a large consumer surplus loss for higher-income consumers who are charged higher prices. They find the optimal level of the subsidy to be at $0.

The limited evidence base likely derives from the significant methodological challenges in studying low-income households. Low- and especially lowest-income households are generally "hard to reach" by researchers (Hurley, 2007; Ambrose et al., 2019). Reasons for the difficulties in recruiting vulnerable groups for research include limited access to information and time constraints which frequently intersects with a low income (Shaghaghi et al., 2011; Chester, 2016), and feelings of shame due to language barriers (Lees, 2014). Low levels of literacy further complicate participation (Chester, 2016). Hence, low-income households are underrepresented in many general population studies. Allcott (2011) provides a practical example from the energy efficiency literature documenting higher drop-out rates among low-income households in a Home Energy Report study. The issue extends to studies on programs that are specifically tailored to low-income households. Fowlie et al. (2018) document very low participation rates in Michigan's Weatherization Assistance Program, despite the implementation of a high-effort encouragement design and offering free-of-charge audits.

1.2 Contribution and summary

This thesis contributes to the public economics of decarbonization. A particular focus is on energy efficiency in lowest-income households. Across four chapters, the thesis makes a significant contribution to extending the knowledge base on monetary and behavioral barriers to investment as well as the optimal design of assistance programs to promote energy-efficient investment in lowest-income households. Contributing novel evidence on various aspects of investment decisions for energy-using durables of these households, the research adds to the scarce literature on energy efficiency in low-income households. Drawing on a unique dataset that spans more than 10 years of home energy audits allows an in-depth analysis of Germany's largest energy efficiency assistance program. The data provides an intimate look into how a particularly vulnerable group of households makes investment decisions. The detailed nature of the data allows for a close examination of the decision processes of participating households from the lower 7% of the income distribution. This group is usually difficult to access for researchers and underrepresented in most studies, but simultaneously highly policy-relevant. Analysis of the data allows to derive novel conclusions about the optimal design of policy instruments to address barriers to energy efficiency and to narrow the energy efficiency gap in the lower tail of the income distribution. The object of the study is the German Stromspar-Check program (English: Electricity Saving-Check) that has been running since 2009. The focus of the analysis is on the refrigerator replacement scheme in the program that subsidizes the purchase of a new efficient appliance for households that own old inefficient appliances via a voucher. The empirical analyses are based on administrative data from the program database, covering audit records from more than 400,000 audits, a field experiment conducted within the replacement scheme and a survey administered among the energy advisors of the program.

Chapters 1 to 3 deal with different aspects of the extensive margin of the investment decision – the binary decision to invest, providing answers to the question which factors matter for encouraging energy-efficient investment in lowest-income households. Chapter 1 studies the relevance of monetary and procedural barriers to investment. The analysis exploits two exogenous changes in the design of the assistance program. A reduction of the subsidy on the replacement voucher by 50 Euro altered the monetary incentives for investment. A change in the program procedures introduced elective enrolment – as compared to automatic enrolment of eligible households before – and implemented a strict deadline for the voucher, potentially triggering behavioral adjustments to the decision process. The detail of the data allows to further investigate the mechanisms behind both, observing at which step in decision making the administrative procedures and monetary incentive work. As both program changes were implemented on short notice and imposed by external forces, they are suitable as quasi-experiments. The analysis applies a Regression Discontinuity Design, finding that the reduction of the subsidy for replacement reduces the investment rate, on average at 26 percent, by 5 percentage points while the change in procedures increases appliance replacement by 4 percentage points. The monetary incentive affects the first-stage decision of households to enroll in the program, while it does not affect their decision to redeem the voucher. Descriptive analysis of the timing of voucher request and appliance replacement suggests that behavioral effects inherent in the procedural change are driven by goal-setting and the strict deadline. Comparing the effect sizes estimated from both changes yields a monetary equivalent of the change in procedures worth at least 35 Euro in terms of the subsidy value. Altogether, the chapter provides insights on the relevance of monetary and behavioral incentives for investment in lowest-income households and suggests that outcomes of assistance programs may be cost-effectively improved using behavioral instruments.

Chapter 2 builds on the insights from the first chapter, further investigating the scope behavioral elements can have for program success. The analysis is based on a field experiment conducted within the refrigerator replacement scheme. The experiment randomly assigns a set of behavioral interventions ("nudges") to households at the participating program sites. The first set of nudges targets the information letter that invites households to request a replacement voucher. These interventions include a gain and a loss frame of (lost) energy savings from (no) replacement and calculation and presentation of energy savings based on household composition, appealing to peer experience. The second set is made up of two sorts of reminders: a written SMS or letter and a visual reminder in the form of a refrigerator tag. The selection of implemented nudges is based on insights from the behavioral economics literature on the effectiveness of nudges in studies of the general population. The analysis of more than 1,800 household decisions demonstrates that low-income households respond differently to nudging as compared to the general population. Investment rates under the gain frame are higher than under the loss frame, a result contrary to what prospect theory predicts. The effect is mainly driven by the lowest-income households in the sample receiving unemployment benefits. Peer experience and reminders do not significantly increase investment rates of households, and some even backfire. The chapter contributes to the nudging literature, providing a test of the transportability of nudges from a general population, where these interventions have been proven effective, to lowest-income households. The insights are relevant for program designers of public assistance programs directed at vulnerable groups who consider using nudging.

Chapter 3 explores another behavioral aspect – the importance of outreach and information, focusing on the role of energy advisors in households' investment decisions. To address this question, a survey was conducted among energy advisors in the Stromspar-Check program, collecting data on their demographic characteristics, economic preferences and attitudes toward the program, and their own appliance replacement. These survey responses were linked to the administrative data from the program database, creating a novel dataset that tracks 113 advisors with detailed profiles of personal characteristics and preferences and their interactions with more than 7,000 households they visited. The analysis investigates whether advisors influence households' decisions to replace appliances, the factors that drive this influence, and whether demographic similarities between advisors and advisees shape investment decisions through peer effects.

The findings indicate that advisors play a significant role in households' replacement decisions. Their performance when it comes to advising households, offering support and convincing them of the benefits of appliance replacement varies. Several factors contribute to this variation: advisors' own investment in an efficient appliance, their economic preferences and attitudes towards the program all impact households' investment decisions. Interestingly, while socio-demographic characteristics do not directly influence households' decisions, they matter via peer effects in audits where advisors and households share common characteristics. This chapter provides insights on the relevance of advisors for the successful implementation of assistance programs. Like the previous chapter, these insights inform program designers of assistance programs on potential factors of improvement.

Chapter 4 investigates the intensive margin of the investment decision, specifically whether households consider lifetime energy cost of the new appliance they purchase, and how monetary and behavioral factors affect their decision. Data on electricity prices and the purchases of households that replaced appliances allow to determine whether these households choose the optimal level of energy efficiency. The analysis compares the strength of demand responses of households to changes in appliance purchase prices and changes in energy cost driven by electricity price increases, respectively. If households rationally consider energy cost in their investment decision, their response to changes in both cost components should be equivalent. Using conservative assumptions regarding the discount rate and lifetime of the appliance, households are found to not rationally consider energy cost in their purchases. The importance of budget constraints as reason for the low consideration of energy cost is evaluated using exogenous changes in the program subsidy for replacement. The relevance of inattention to energy efficiency – a behavioral barrier – is examined using a revision of the EU Energy Label that increased salience in the differences of the level of efficiency between appliance models. The results suggest that low consideration of energy cost is driven by inattention to energy efficiency, while budget constraints are not a barrier in this context, demonstrating that behavioral barriers are also present at the intensive margin of investment for lowest-income households. However, information instruments, such as energy labels, can help address this barrier. With heterogeneity in the consideration of energy cost along the income distribution, policy implications also arise for the effectiveness and distributional impacts of policy instruments. If households with lower incomes do not consider energy cost, carbon taxes, being part of energy cost, will not affect their investment decisions, and if, contrary to higher-income households, these households do not purchase at the optimal energy efficiency level, their burden from carbon taxes will be unequally higher.

While Chapters 1 to 4 focus on the economics of decarbonization in the residential sector studying private households, Chapter 5 shifts the perspective to the industrial sector, analyzing the effects of mergers and acquisitions on the environmental performance of European industrial firms. This chapter connects to the core chapters 1 to 4, also investigating investment decisions, though on a much larger scale. Here, investment relates to technology adoption of acquiring and acquired entities as well as acquisitions of entire facilities and firms. The scope of analysis extends beyond energy efficiency, encompassing a broader range of environmental outcomes. In addition to examining the effect of acquisitions on greenhouse gas emissions connected to energy generation, the study considers various local pollutants with externalities. Within this context, improving energy efficiency is just one of several possible strategies for reducing emissions. The transfer of technologies between acquiring and acquired firms is often named as a benefit from mergers and acquisitions. This chapter tests this hypothesis for the case of environmental technologies. A merge between the European Pollutant and Transfer Registry (E-PRTR) and the Orbis Database allows observing pollutant emissions alongside economic and financial indicators and ownership relations of large industrial firms in the manufacturing sector in the EU-15 economies, Hungary and Norway in the period 2007 to 2016. The sample consists of around 12,000 facilities associated with around 7,000 firms that change ownership once during the observation period. An event-study analysis with staggered treatment adoption (Sun and Abraham, 2021) exploits variation in the timing of acquisitions of these entities; the treatment timing is shown to be independent of observed firm characteristics. The results indicate that facilities and firms do not change their emissions and emissions intensity as long as they remain in operation after acquisition. Entities that shut down after acquisition strongly reduce their emissions via reductions in output. However, these reductions are observed to start already before acquisition so that they cannot be attributed to the ownership changes. The analysis does not find evidence for transfers in environmental technologies between targets and acquiring parent companies either. Hence, the results do not provide evidence for the hypothesized benefits from ownership changes. The chapter contributes first evidence on the effect of ownership changes on firms' environmental performance in Europe. Spanning 17 countries and a decade, the analysis allows to draw broader conclusions less dependent on country-specific peculiarities than earlier studies. The findings are informative for competition authorities and imply that regulators should consider environmental components when evaluating the costs and benefits of M&A events. The research in this chapter links to the public economics of decarbonization in the industrial sector. Acquiring more knowledge about firms' behavior is important for regulators to develop suitable instruments to steer firms' investment decisions regarding environmental technologies.

Chapter 2

Money versus procedures – Evidence from an energy efficiency assistance program

Coauthor(s): Timo Goeschl and Martin Kesternich

Abstract: In many countries, governments have put in place targeted programs intended to support energy efficiency investments by low-income households, but have encountered low take-up even when subsidies are high. Using evidence from a large energy efficiency assistance program, we demonstrate that seemingly small procedural changes can substantially improve take-up and that these changes have effects comparable to significantly raising subsidies. Observing 77,305 durable goods purchase decisions in a refrigerator replacement program, our RD design exploits two quasi-exogenous temporal discontinuities in voucher value and procedures. Despite seeming disadvantageous, the procedural changes actually raise replacement rates among the target demographic of low-income households, an effect roughly equivalent to raising voucher values by €35. These results suggest that even under fixed budgets, the performance of energy efficiency assistance programs can be improved through empirically guided procedural design.

Publications: This chapter has been published in the *Journal of Environmental Economics and Management*. A previous version of this chapter is available as ZEW Discussion Paper No. 22-020.

Acknowledgments: We acknowledge financial support by the Fritz Thyssen Foundation. We thank the SSC staff for the partnership and Peter Heindl for his support in initiating this cooperation.

2.1 Introduction

In recent years, electricity retail prices for residential customers have been increasing steeply across much of the developed world (U.S. Energy Information Administration, 2023; Eurostat, 2023), bringing the adverse distributional effects of high prices on low-income households to the attention of researchers (Fabra and Reguant, 2024; Frondel et al., 2019) and public entities (Levinson and Silva, 2022; Sirin et al., 2023). Policy-makers have responded by designing programs intended to attenuate the adverse distributional effects of rising energy prices by helping low-income households become more energy-efficient.[1] At a conceptual level, these programs appear simple: They combine technical advice and practical help to households with economic incentives, such as subsidies. At a more detailed level, however, they reveal at times astonishing procedural complexity for all involved parties. For example, the 2014 evaluation report of Weatherization Assistance Program (WAP) dedicates an entire section to 'program complexity' as experienced by households, agencies, auditors, and weatherization crews (Tonn et al., 2014).[2] Onerous design features are an intuitive explanation for low take-up of privately beneficial assistance programs among low-income households (Fowlie et al., 2018). It would seem obvious that thinking carefully about how to design a program may be as important as setting the right economic incentives. Yet, the contribution of program design[3] to program performance is rarely substantiated, let alone quantified, in policy discussions.[4] To fill this gap, we exploit an empirical opportunity presented by one of the world's largest energy efficiency assistance programs to show that procedural design is impactful: Even seemingly small and arguably cost-saving procedural changes can substantially substitute for monetary inducements.

The empirical opportunity for studying and monetizing the effects of procedural change on household behavior arises in the context of a nation-wide energy efficiency assistance program targeting low-income households in Germany. Since 2009, the Refrigerator Replacement Program (RRP) has subsidized the modernization of household refrigeration appliances. There, refrigerators are the consumer durable that accounts for the largest

[1]Examples are the Weatherization Assistance Program (WAP) by the U.S. Department of Energy, Services Locaux d'Intervention pour la Maîtrise de l'Énergie (SLIME) of the French Ministry for Ecological Transition, and the Electricity Savings Check (SSC) by the German Ministry for the Environment.

[2]As an illustration, the report points out that the program's knowledge base comprises more than 100 work categories and more than 800 gradable actions.

[3]"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).

[4]This is despite important examples in other policy areas, from procedural hassles in food stamp programs (Bertrand et al., 2006), information provision in school choice (Hastings and Weinstein, 2008) and variations in tax mailings (Bhargava and Manoli, 2015) to the local presence of Social Security field offices (Deshpande and Li, 2019) and electronic food vouchers (Banerjee et al., 2022).

share (about 25 percent) of household electricity consumption (BDEW, 2019).[5] The RRP is embedded in a larger initiative called "Energy-Saving-Check" (SSC), funded by the German 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 audit are used to screen for eligibility for the RRP.[6] Three criteria determine eligibility: a welfare recipient status, refrigerator age; and expected enegy savings from refrigerator replacement. In its first twelve years, the SSC identified 77,305 eligible households. Eligible households are 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. Average program take-up among eligible households is between 25 and 30 percent for an average payback period of 3.5 years.[7]

In this setting, we study 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 indicator of the RRP, not least because of the considerable cost of each home energy audit to the program. Three aspects of this setting help enrich the literature on program design for low-income households. First, the RRP experienced two quasi-exogenous shocks that changed different dimensions of the program design unexpectedly and at short notice. The shocks mean that we observe the RRP in three distinct regimes over time. Proceeding conservatively and making use of the rolling nature of the program, our empirical strategy shows that much of the change in replacement rates across the three regimes can be attributed to the program design changes.

Second, the design dimensions changed by each of the policy shocks were essentially orthogonal. One shock changed the level of the cash subsidy from € 150 to € 100.[8] The paper can therefore speak to the effects of large relative changes in financial incentives

[5] In the US, air conditioners are the most energy-intensive home durable, with 12 percent of total home energy expenditures in 2015 (EIA, 2015). In Germany, AC units remain rare.

[6] In Germany, the vast majority of low-income households own their refrigerators. In our sample, only 2.6% of households do not, making them ineligible for the RRP.

[7] In 2020 for example, average electricity prices were € 0.289 and annual savings 342 kWh per replaced refrigerator, implying annual electricity bill reductions of € 99. The take-up rate compares favorably to that induced by financial incentives of WAP, which is minimal, even for shorter payback (Fowlie et al., 2015, 2018; Hancevic and Sandoval, 2022). Evidence on appliance replacement programs is only available for episodic campaigns directed at the general population: A 36-month campaign in Mexico between 2009 and 2012 achieved 17 percent take-up (Davis et al., 2014), increasing by 34 percent following an increase in the subsidy of $80 (Boomhower and Davis, 2014). For a similar, but shorter U.S. campaign (26 weeks on average), take-up rates are difficult to compute (Houde and Aldy, 2017).

[8] This corresponds to a drop from 37 percent of the appliance price to 24 percent, on average.

on program performance among low-income households.[9] The other shock changed program procedures from automatic enrolment (every eligible household received the voucher by default) to elective enrolment (vouchers had to be requested) and voucher terms from flexible (three-month validity, unlimited renewability) to rigid (two-month validity, non-renewable). The paper can therefore speak to the effects of procedural changes on program performance among low-income households, including the role of "psychological frictions" (Bhargava and Manoli, 2015), "hassle" (Bertrand et al., 2006), and deadlines (Bertrand et al., 2010; Shu and Gneezy, 2010; Altmann et al., 2021). Importantly, we are able to benchmark the effect of these procedural changes against the variation in cash subsidies, providing an intuitive but novel metric of comparison.

Third, low-income households take investment decisions in our setting, rather than consumption or labor supply decisions typically studied. This decision constitutes a particularly challenging problem for all owners of energy-intensive consumer durables who pay their own electricity bills (Rapson, 2014; Wang and Matsumoto, 2021): Due to wear and tear in use, the 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 literacy, specifically to energy-related questions, leading to errors in decision-making (Calvet et al., 2009; Brent and Ward, 2018; Blasch et al., 2021). 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 problems and miss optimal replacement points, a particularly costly mistake for German low-income households due to high electricity prices,[10] lower investments in energy-efficiency consumer durables (Ameli and Brandt, 2015; Schleich, 2019), and an annual billing cycle.[11]

On the basis of twelve years of RRP data on home energy audits, program enrolment,

[9]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).

[10]At €0.37 per kWh in July 2022 Germany has some of the highest retail prices for electricity in the world. Consumer electricity prices have more than doubled since 2002. In July 2022, wholesale prices peaked at a new all-time high. In consequence, some providers started to charge prices of close to €1.00 per kWh. In addition, German low-income households tend to face higher retail prices for electricity than the average German household.

[11]As a result, 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.

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 5 to 7 percentage points higher. This "subsidy elasticity" underscores that program performance is demonstrably a question of subsidy levels. Program administrators will want to take note 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 from automatic enrolment with flexible voucher terms to elective enrolment with rigid voucher terms in the RRP cause replacement rates to rise by 4 to 15 percentage points. The direction of this effect is as interesting as its composition, magnitude, and dynamics. At the enrolment stage, the share of enrolled households drops from 100 percent under automatic to just under 40 percent under elective enrolment. Through the lens of the behavioral economics of assistance programs, the size of this decrease is consistent with a change in the default (Thaler and Sunstein, 2021) and with procedural "hassle" being imposed on eligible households (Bertrand et al., 2006). At the same time, electively enrolled households exhibit – under the rigid two-month deadline – vigorous program take-up at the redemption stage. 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. They cannot explain, however, why cumulative replacement rates among eligible households 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, this evidence is consistent with deliberate 'opt-ins' facilitating effective "goal-setting" (Locke and Latham, 1990) towards replacement and with rigid deadlines helping households to overcome time management problems (Bertrand et al., 2006).[12] Jointly, they lead to an intensive-margin effect that more than compensates for the changes in the enrolment mechanism.

Third, we conduct back-of-the-envelope calculations of the merits of alternative program design. Comparing effect sizes,[13] we find that the procedural changes improved replacement rates equivalent to an estimated subsidy increase of about €50 per re-

[12] *Prima facie,* the effect 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.

[13] Examples of such comparisons for the general population are interest rate equivalents to changes in deadlines and in advertising content in loan marketing (Bertrand et al., 2010) or monetary equivalents to product information in the purchase of energy-efficient light bulbs (Allcott and Taubinsky, 2015).

placing household while adding little to no cost to the program.[14] As a conservative illustration, bringing these procedural changes forward to 2013 rather than 2018 would have realized 2,000 additional replacements by low-income households at the same budgetary cost, leading to average additional savings of €99 in annual electricity bills of replacing households, or aggregate annual savings of €201,800.

We proceed as follows: In the following Section 2.2 we provide the necessary background on the Refrigerator Replacement Program. In Section 2.3, we explain the data on which the analysis is based. Section 2.4 lays out the empirical challenges and the empirical strategy. In Section 2.5, 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 2.6. In Section 2.7, we compare the effects of subsidy and procedural change to each other and compute the effects of the alternative, untried regime on program success. Section 2.8 concludes.

2.2 The Refrigerator Replacement Program

Since January 2009, the Refrigerator Replacement Program (RRP; German: *Kühlgeräte - Tauschprogramm*) has been offering cash vouchers to households on federal income support[15] 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 "Energy-Saving-Check" (SSC, German: *Stromspar-Check*) that provides support to low-income households for reducing their energy and water consumption by conducting home energy audits. These 'SSC households' constitute the pool of households that are screened for participation in the RRP. 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 was scaled up to its current size with the start of the second funding cycle of the SSC ("SSC plus") in April 2013. The last funding cycle that we observe in the data started on April 1, 2019 and lasted until early 2022 ("SSC active", German: *Stromspar-Check*

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

[15] To qualify, the household needs to receive one of the following types 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. In 2020, more than 7 percent of German households qualified on this basis (Bundesagentur für Arbeit, 2020).

Aktiv). Recruitment of qualified households into the SSC's home energy audits takes place through a variety of channels.[16] 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 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 not more than two weeks, with the schedule driven by program logistics and under control of the local branch. During the first visit, the "energy advisors" inventory all electric devices and their usage in the household, collect utility bill information, assess the electricity consumption of refrigerators and freezers, and educate the household on electricity-saving behavior. Upon return to the branch office, the advisors use the inventory and electricity consumption assessment 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.[17] Non-eligible households then exit the SSC. For eligible households, the second visit contains additional components through which they are specifically targeted for enrolment in the RRP.[18]

The rationale for enrolling households in the RRP is the large contribution, roughly 25 percent (BDEW, 2019), that refrigeration appliances make to the electricity consumption of the average German household.[19] Differences in refrigerator efficiency can therefore significantly impact residential 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 a replacement with the most energy-efficient class of devices on the market.[20] The expected savings are part of the information shared with the household at the occasion of the second visit. As we explain below, the specific enrolment procedures changed between 2017 and 2018, but enrolment always concludes with the receipt of a voucher. Enrolled households, i.e., households

[16] SSC and RRP are actively promoted in many employment and social assistance agencies across the country through printed and audiovisual material. They are 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 SSC also maintains a website where information is available about the RRP in eleven languages. Additionally, recruitment takes place directly through the local branches.

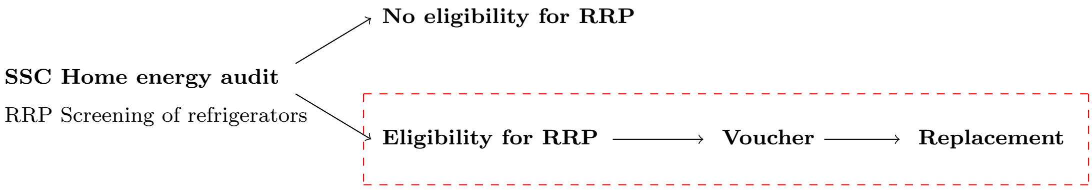
[17] The kit features LED light bulbs, switchable socket strips, TV standby cut-off switches, timers and water flow regulators. These items are directly installed by the two advisors.

[18] Completion of the first visit is necessary to become eligible for the RRP.

[19] "Refrigerator" refers to refrigerators, freezers, and combination units within the program.

[20] 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+++ predominantly in the classes B, C and D.

Figure 2.1: Sequence of procedures in SSC and RRP



Notes: Schematic flow chart of the sequence of procedures in Stromspar-Check (SSC) and Refrigerator Replacement Program (RRP). The red box (dashed lines) defines the sample used for the analysis in this paper.

in possession of a voucher, 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+++; and provide proof that the original refrigerator has entered the recycling chain.[21] 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. Figure 2.1 provides a flow chart of the sequence of procedures in the SSC and RRP. The red box in dashed lines marks the sample of interest for our analysis.

The RRP is the only federal scheme for replacing refrigerators in low-income households. At the same time, complementary programs exist in four of the sixteen states (*Länder*) and in a number of municipalities.[22] This coexistence of programs is one feature of the policy landscape that requires an appropriate empirical strategy. Another feature are cyclical dynamics at the federal level that are driven by the starting and ending of the funding cycles: 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.

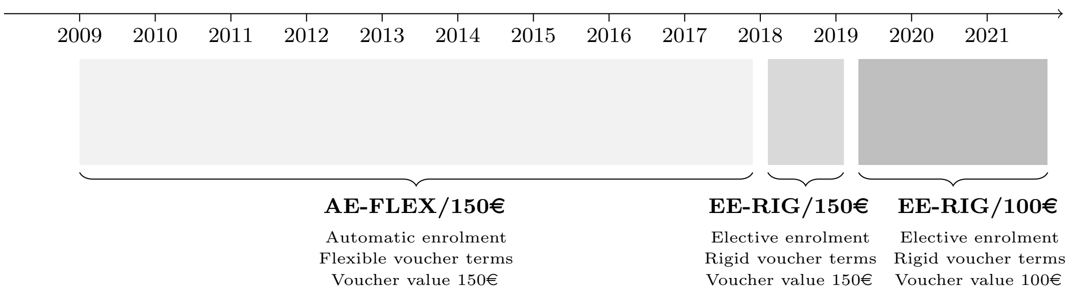
The policy landscape also gave rise to two unexpected changes in the RRP, one on January 1, 2018 and one on April 1, 2019 (see Figure 2.2 for a timeline). The first change, within the third funding cycle of the SSC, simultaneously affected specific procedures of the program, namely the enrolment mode via which households enter the RRP, and the terms of the voucher. The enrolment mode switched from automatic

[21] A further requirement during the first and second funding cycle up to March 2016 was that the volume and type of the new refrigerator had to be identical with the original refrigerator.

[22] At the level of the federal states, Berlin has offered 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 has complemented 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.

enrolment (AE) until the end of 2017 to elective enrolment (EE) from 2018 onward. Under AE, all eligible households automatically received the RRP voucher on the second visit. Under EE, on the second visit eligible households have been receiving an *invitation to request* a voucher from the local branch. EE hence requires households to take the active step of contacting the local branch and asking for the voucher to be mailed to them. In addition, the terms of the voucher changed: 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. From 2018 onward, the voucher has been valid for two months, without the option to renew. The reason for the change from flexible three-month renewable (FLEX) to rigid two-month non-renewable (RIG) terms in January 2018 was the discovery in late 2017 that the 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'.

Figure 2.2: Timeline of changes in program design



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.[23]

Taken together, the voucher-based subsidy scheme has therefore experienced three distinct regimes since 2009 (see Figure 2.2): A regime *AE-FLEX/EUR150* with automatic enrolment, flexible terms, and a subsidy of €150 up to December 2017, a regime *EE-RIG/EUR150* with elective enrolment, rigid terms and a subsidy of €150 up to

[23] As a result, the subsidy covered 37% of the purchasing price of the average new refrigerator before and 26% after the change, or 44% and 29%, respectively, when also considering the complementary state programs. See Appendix Figure A2.2.

January 2019,[24] and – finally – a regime *EE-RIG/EUR100* with elective enrolment, rigid terms, and a subsidy of €100 from February 1, 2019 onwards.

2.3 Data

Our data include more than 360,000 households that participated in an SSC audit between January 2009 and December 2020 (repeated cross-section). Of these, about 77,000 households were eligible for the RRP, the sample of interest for our analysis (see Figure A2.1 in the Appendix for the distribution of audits over the program period). About 20,000 households actually replaced their refrigerator. The take-up rate is therefore around 26 percent (see Table 2.1: 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 location by postal code. Documentation from the audit includes the date of the first and second visit, the responsible local branch, auditor IDs, annual electricity consumption and price paid per kWh. For the RRP, status of eligibility, enrolment (i.e. voucher request) and voucher redemption after refrigerator replacement is available. So is information on the old refrigerators, such as age, kWh consumption as measured during the audit, and volume. Finally, the data contain information on the newly purchased refrigerator, including the purchasing price, volume and kWh consumption as specified by the manufacturer.

To prepare the data for the analysis, we remove implausible observations.[25] In households where more than one appliance is marked eligible for replacement we use the first of those as "old refrigerator". For households whose audits were administered by only one advisor (10,505 observations) we introduce an ID in place of the second advisor ID to not lose these observations when introducing auditor fixed effects in the analysis. We also recode implausible refrigerator characteristics as missing. The

[24] 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 that underwent a home energy audit during the interim period could request a voucher no sooner than April 1 and thus they received a voucher for €100. Therefore, we set the day for the regime change on February 1, 2019 in our analysis.

[25] For example, we drop observations that report 0 inhabitants in the household (1,884 observations) and observations with a date of the second visit prior to January 1, 2009 (45 observations).

database truncates values of some variables at a maximum cutoff. This cutoff changes for a few variables over time. We harmonize truncation and set a consistent maximum value over the sample period. Table 2.1 presents descriptive statistics for the prepared dataset on the sample of eligible households, 77,305 observations in total. On average, eligible households consist of 2.8 household members which live on 69 square meters.[26] 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.

Table 2.1: Descriptive statistics

	Observations	Mean	Median	Std. Dev.	Min	Max
RRP variables						
Total no. of eligibile households (2009-2020)	77,305					
– Automatic enrolment yes/no (2009 - 2017)	49,182	0.99	1	0.04	0	1
– Elective enrolment yes/no (since 2018)	28,123	0.40	0	0.49	0	1
Voucher redemption yes/no	77,305	0.26	0	0.44	0	1
– Regime AE-FLEX/EUR150 (2009 - 2017)	49,182	0.26	0	0.44	0	1
– Regime EE-RIG/EUR150 (2018 - January 2019)	14,945	0.32	0	0.47	0	1
– Regime EE-RIG/EUR100 (February 2019 - 2020)	13,178	0.19	0	0.39	0	1
Federal subsidy rate (share of purchase price)	19,909	0.35	0.32	0.16	0.07	1
– Subsidy rate €150 (2009 - January 2019)	17,428	0.37	0.34	0.16	0.09	1
– Subsidy rate €100 (February 2019 - 2020)	2,481	0.24	0.21	0.10	0.07	1
Household variables						
Number of inhabitants	77,305	2.79	2	1.73	1	10
Electricity price per kWh	77,270	0.28	0.28	0.02	0.03	0.90
Living space in square meter	77,305	69.38	65	24.65	11	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	77,305	17.30	16	4.69	11	45
Volume in liters	77,305	239.29	238	76.88	37	600
Estimated savings from replacement in kWh	77,305	336.07	286	166.93	0	5,736

AE-FLEX denotes the automatic enrolment mode with flexible voucher terms and EE-RIG denotes the elective enrolment mode with rigid voucher terms. The federal subsidy rate is the share that the federal subsidy accounts for in the purchase price for the new refrigerator. Appendix Figure A2.2 shows the distribution of the subsidy rate summing up the federal and, if applicable, the respective complementary state subsidy.

Of the eligible households, 35 percent live together in families with at least one child in

[26]An average German household consists of 2.03 members (Destatis, 2020) and lives on 93 square meters (Destatis, 2018).

the household; more than a third of these families have more than two kids. 29 percent in the sample are single households, with about a third retired. 14 percent are single parent households with one or more children and 6 percent are retired couples. The remaining 16 percent in the sample have a different household composition. Virtually all eligible households are on some type of federal income support: 75 percent receive unemployment benefits, 12 percent basic income,[27], 5 percent a housing allowance,[28] and 4 percent other public benefits.

The RRP measurably reduces the energy bills of households that take up the program: Their average estimated reduction in annual energy consumption between 2009-2020 amounts to 336 kWh (see Table 2.1), with little trend across the observed period (see Appendix Figure A2.5). Replacement refrigerators grow in size over the sample period (see Appendix Figure A2.9) while the electricity price paid by target households increases from an average of €0.205 in 2009 to €0.289 in 2020 (see Appendix Figure A2.8), mirroring the general trend in Germany. As a result, average savings in electricity bills of replacing households increase from €70 in 2009 to €99 in 2020.[29]

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 (Destatis, 2021) and we weight each index according to the share of each RRP category in all newly purchased durables within the program.[30]

2.4 Empirical strategy

2.4.1 Identification

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). We consequently observe eligible households making replacement decisions in three distinct regimes. Our identification strategy therefore translates into a pre/post analysis of the procedural

[27]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.

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

[29]In January 2022, the average price per kWh paid in Germany further increased to €0.362 (BDEW, 2022) resulting in average annual savings of €123. At these rates at an average purchase price of €478 less the program grant of €100, the investment amortizes after about three years.

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

and the subsidy change while controlling for as many confounding factors as possible around the regime change. Our main analysis relies on two different econometric approaches – OLS and RD-in-time – that are suitable for such a setting, but address the empirical challenges in different ways.

The OLS analysis provides a comparison in means before and after a regime change, considering all observations in the full sample. The approach controls for potentially time-varying observable confounding factors by considering household characteristics as well as time and local fixed effects. The RD-in-time approach considers observations located close to both sides of the regime change within a certain bandwidth. It can allow for a more flexible form of the underlying model that accounts for the temporal distance of individual observations to the threshold. By restricting the analysis to observations within an appropriate time window around the regime change, confounding factors are expected to be less likely to vary significantly.

The continuity requirements of both OLS and RDD-in-time are threatened by possible selection effects. For continuity to hold, households need to have been quasi-randomly assigned to the three regimes of the program. We have three reasons for a justified belief that selection effects do not compromise our analysis. The first 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 the federal subsidy at the end of 2018 (see Section 2.2 for a detailed description). The second reason is empirical: To test formally for evidence that households strategically selected out of or into regimes around program changes, we look for bunching and discontinuities in household observables around the cutoff points. These tests reveal no visual clues for bunching around the thresholds (see Appendix Figures A2.10 and A2.11), and based on McCrary tests, we cannot reject the hypothesis that there is no bunching around the thresholds (see Appendix Tables A2.1 and A2.2). We also do not find any discontinuities in household observables (see Appendix Figures A2.12 and A2.13). The third reason is the dynamic nature of the program: New households become continuously eligible for enrolment into the program as their refrigerators 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 strongly over time (see Appendix Figures A2.3 to A2.7). This supports the notion that there is no evidence for clear selection effects and that observations can be treated as independent.

Irrespective of whether OLS or RD-in-time is used, the empirical strategy has to take into account that the conditions under which households take the replacement decision can vary over time and space. Our approach accounts for a range of temporal and spatial factors: Persistent trends such as rising electricity prices during the sample period, cyclical effects such as seasonal variations in refrigerator prices and seasonally varying household liquidity. Changing conditions on the German refrigeration appliance market do not pose an obvious threat.[31] We nevertheless control for short-term fluctuations in purchase prices using a retail price index by the German Federal Statistical Office. We also account for the presence of complementary programs at the state and municipal level that coexist with the RRP. In addition, temporal and spatial factors inside the program affect replacement decisions: One example are differences between local branches in program practices and differences in audit quality between advisors, even at the same branch. Interim periods between funding cycles and around unexpected program changes similarly need to be accounted for. The relevance of such interim 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. The drop can be explained by a significant share of eligible households being denied enrolment. At the same time, the share of redeeming households among eligible households inches higher, especially around the procedural change (see Appendix Figure A2.14).[32] Both observations suggest a potential bias of selection towards households with a high propensity to replace their refrigerator in the interim period.[33] By controlling for a broad range of factors, we are confident that the assumption of constant treatment effects important for identification in RD designs holds in our setting.[34]

[31]Unit sales (ZVEI, 2023) and purchase patterns (Destatis, 2023) have not changed perceptibly from year to year. Sales figures remain constant between 2015 and 2018, with data for 2019 unavailable. Purchase data from the representative Household Income and Consumption Survey conducted by the German Federal Statistical Office records virtually constant amounts of cooling appliances bought between 2016 and 2019, with data for 2018 unavailable. This is consistent with the absence of reported institutional or regulatory changes on the appliance market.

[32]In the interim period starting around two months before and ending around two months after the implementation of the procedural change, 6,000 households that fulfilled the eligibility criteria did not receive an invitation to join in the program and to request a voucher (consisting of 2,423 eligible households before the design change and 3,577 households after, and making up 63 percent of all households that fulfill the eligibility criteria during this period). In the interim period 2 months around the change in subsidy levels, 2,676 eligible households did not receive an invitation to join in the program (consisting of 1,888 households before the change and 788 households after, and making up 53 percent of all households that fulfill the eligibility criteria during this period).

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

[34](Hausman and Rapson, 2018) discuss the assumption in the context of RD-in-time and conclude that the threat of violation of the assumption tends to be larger in designs that use bandwidths of several years around policy changes. As we restrict the bandwidth in our estimation to under a year, the threat of violation in our setting should be less severe. As an empirical test, we do not find that the choice of bandwidth strongly affects our estimates.

2.4.2 Specifications

First, we estimate an OLS model that includes the full set of observations (2009-2020) before and after the policy change. The OLS approach both takes into account a set of control variables and fixed effects for energy advisor ID, branch, month and 2-year indicators. We add relevant controls which could influence the individual replacement decision of households, such as the price paid per kWh, 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.[35] We also add a refrigerator price index as control for changes in refrigerator purchase prices over time.

We estimate the following model separately for the subsidy and procedural variations:

$$Outcome_{it} = \beta_0 + \beta_1 Regime_t + \beta_2 X_i + \gamma_t + \delta_b + \zeta_a + \varepsilon_{it} \tag{2.1}$$

Regime indicates the current regime as a binary treatment variable: for the change in voucher value, the variable is coded 1 for a €150 and 0 for a €100 subsidy (automatic enrolment and flexible voucher terms in both regimes); for the procedural change, the variable is coded 0 for automatic enrolment and flexible voucher terms, and 1 for elective enrolment and rigid voucher terms (€150 subsidy in both regimes). X is a vector of controls. The subscripts t and i denote time in days and individual households, b denotes the local branch and a the advisor the audit is administered from. We cluster standard errors at the branch level.

Second, we employ an RD-in-time within a bandwidth of nine months.[36] As a robustness check, we illustrate in bandwidth plots how the choice affects the coefficient of interest for both RD estimations for all outcomes in the Appendix (see Figures A2.19 to A2.22). The running variable *DayCount* counts the number of days from the program change in both directions. To allow the slope of the linear time trend to vary on both sides of the threshold, we interact the treatment indicator *Regime* with the running variable. We also add location fixed effects at the branch and advisor level. The first

[35] Table A2.3 provides a comparison in means for relevant covariates before and after each program change (subsidy level and procedures). The imbalances in some of the variables capture secular changes in the economic environment of the RRP that are exogenous to the program.

[36] To test for optimal bandwidth in our setting we follow (Calonico et al., 2017). The test indicates an optimal bandwidth of 206 days (about 7 months) for the estimation of the subsidy effect and 178 days (about 6 months) for the procedural effect. However, we cannot explicitly test the optimal bandwidth taking into account the Donut design we use in our third specification. Therefore, we add two months of bandwidth on top for both RD specifications, which leaves us with a bandwidth of around 9 months for the subsidy estimation and around 8 months for the procedural estimation. For symmetry and comparability, we apply a bandwidth of 9 months for both effects for both RD specifications. Nine months is also the upper bound of available bandwidth for the subsidy effect estimation as the start of the first SARS-CoV-2 lockdown that may have confounded replacement rates puts a ceiling to bandwidth choice.

RD specification is estimated according to the following equation:

$$Outcome_{it} = \beta_0 + \beta_1 Regime_t + \beta_2 DayCount_t + \delta_b + \zeta_a + \varepsilon_{it} \tag{2.2}$$

Standard errors are clustered at the branch level.

Third, we present results of a RD specification that adds a Donut design as proposed by Barreca et al. (2011) and an Augmented Local Linear (ALL) design as proposed by (Hausman and Rapson, 2018). When applying RD to a setting which is prone to irregularities in the observations closely around the policy change, observations in this period might be better excluded from the sample on each side of the threshold, creating a "Donut hole".[37] We construct a donut that excludes two months of observations on each side of the threshold (program change) as we observe a drop in eligibility rates in this period.[38] The design controls for a potential bias in the selection towards households with a high propensity to redeem the voucher during the interim periods. We additionally apply an ALL design to adjust our outcome variable for location effects, thereby increasing the precision of our estimation. In a two-step approach, we first regress the outcome of interest on 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 RD estimation (Hausman and Rapson, 2018).[39] We apply ALL using a set of spatial indicators. We control for different practices at the local branches as well as for complementary programs by states, municipalities and local energy providers by including branch indicators and for audits conducted by different advisors by including fixed effects for each of the two advisors who conducted the audit. Combining ALL with the RD specification also mitigates the need for a flexible functional form and diminishes potential concerns for overfitting as the use of higher order polynomials puts high weight on observations far away from the cutoff (Gelman and Imbens, 2019; Hausman and Rapson, 2018).

We estimate the second stage of the ALL-RD as shown in the following equation:

$$Outcome_{it}^{Residuals} = \beta_0 + \beta_1 Regime_t + \beta_2 DayCount_t + \varepsilon_{it} \tag{2.3}$$

where the outcome uses the residuals from the first ALL stage that adjusts for location effects. We bootstrap standard errors to account for the ALL two-step approach, using 500 repetitions.

[37] Examples for applications are Ost et al. (2018); Kim and Koh (2020), and Gillingham and Huang (2021).

[38] We observed that a significant share of households participating in audits in the interim periods around the program changes does not receive the information letter as invitation to join the replacement scheme, even though these households fulfill the eligibility criteria. We examine the sensitivity of results to this choice in section 5.

[39] Examples for applications of ALL are Li et al. (2020) and Gillingham and Huang (2021).

We estimate equations 2.1 to 2.3 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 at the second stage of the program out of all enrolled households. The variable of interest is the binary decision to redeem the voucher and replace the refrigerator, estimated on the sample of enrolled households. We only observe this outcome for the period as of 2018.

In robustness checks, we systematically test the impact of different specification choices on the RD estimates along four dimensions: inclusion of the Donut design, inclusion of the ALL approach, bandwidth at lower bound of 6 months or upper bound of 9 months, as well as inclusion of an interaction term between the treatment indicator and the running variable. Results are presented in Figures A2.27 to A2.30 in the Appendix. We find the estimates to be robust in sign and statistical significance across all specifications.

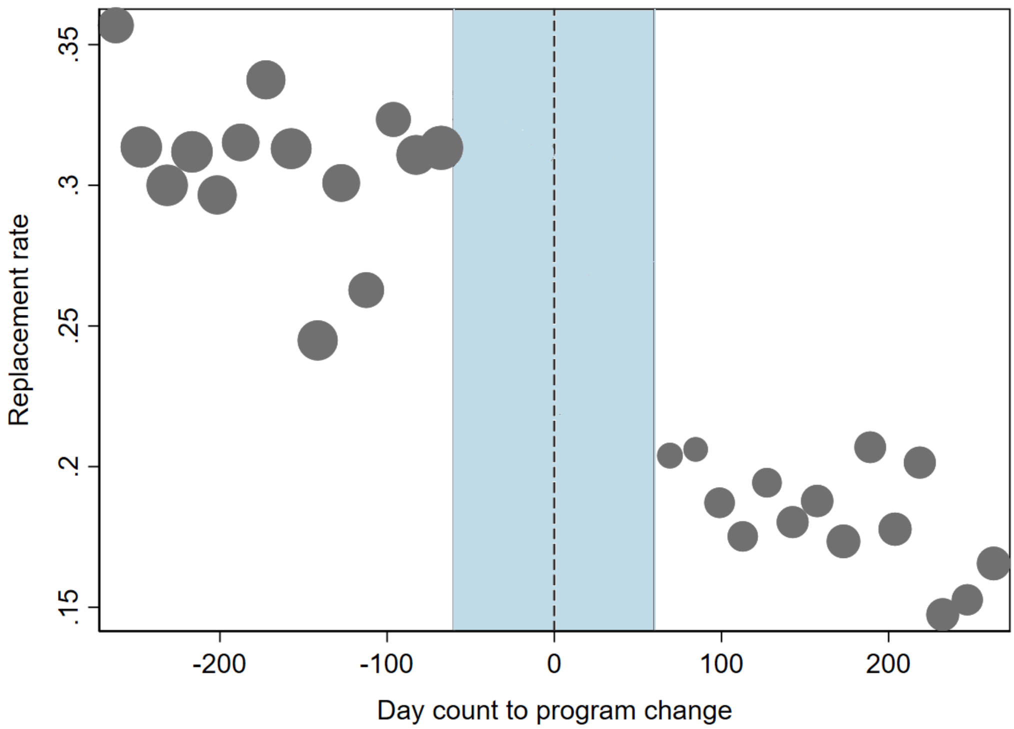
2.5 Main Results

2.5.1 Subsidy variations

We first investigate to what extent replacement decisions among eligible households respond to a large relative variation in the voucher-based subsidy. Figure 2.3 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 15-day interval, with larger bubbles signifying more observations.The blue-shaded area marks the interim period of two months around the change. We exclude data points that fall in this period from the analysis in 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 roughly a 10 percentage point reduction in the share of eligible households replacing their refrigerator. A simple comparison in means for the average replacement rate in the two regimes results in a reduction of 13 percentage points (from 0.32 to 0.19, see Table 2.1).

Figure 2.3: Subsidy variation, replacement rate: Discontinuity



Notes: This figure shows the rate of households that successfully complete refrigerator replacement out of all households that are eligible for the RRP around the reduction of the voucher value by € 50 on April 1, 2019. Replacement rates are binned and averaged over 15 days in a bandwidth of 270 days around the program change. The blue-shaded area marks the interim period of 2 months around the change. We exclude data points that fall in this period from the analysis in a RD Donut Design (see Figure A2.15 in the Appendix for the raw data plot).

Table 2.2 provides our estimation results across three specifications. All models indicate the treatment indicator of subsidy variation (= 1 for the subsidy of € 150, 0 for € 100) to be positive, confirming the visual impression of Figure 2.3 and the difference in means: Households react to prices, leading to a lower replacement rate after the reduction of the subsidy level to € 100. In our preferred specification (column 3) that includes both the Donut design as well as the ALL approach we estimate the replacement rate to be 4.9 percentage points higher for a voucher that has a € 50 higher value (p= 0.010). In the more simplified RDD without Donut and ALL, the effect is very similar at 4.8 percentage points (p= 0.087). In the basic OLS specification, the effect is slightly larger in magnitude at 7.3 percentage points (p=0.008).[40] In Figure A2.19, we show

[40]Appendix Table A2.27 provides robustness check results in a specification chart and Figure A2.26

how the treatment effect changes as function of the bandwidth. For models 2 and 3, the effect ranges from 3 (185 days bandwidth, model 2) to 8 percentage points (225 days bandwidth, model 3). In other words, a 33 percent lower subsidy level is associated with a likelihood of appliance replacement that is around 3 to 8 percentage points lower.

Table 2.2: Estimated effect of subsidy variation on the replacement rate

	1	2	3
Subsidy (€150 = 1)	0.073	0.048	0.049
	(0.027)	(0.028)	(0.019)
Day count		yes	yes
Day count×Subsidy		yes	yes
Location fixed effects	yes	yes	
Time fixed effects	yes		
Controls	yes		
ALL			yes
Donut			yes
No. observations	70,426	16,832	14,890

Notes: Standard errors in parentheses, clustered by branch or bootstrapped (ALL). Location fixed effects include energy advisor ID and local branch. Time fixed effects include month and 2-year indicators. Augmented Local Linear includes advisor IDs and branch indicators. The Donut design excludes 2 months around the program change. Column 1 uses the sample over the sample period 2009-2021. RDD estimates in columns 2 and 3 use the sample of eligible households in a bandwidth of 9 months around February 1, 2019.

We compare whether households that replace their refrigerators with a €150 versus a €100 subsidy change in terms of their observable characteristics (see Table A2.5). Households that successfully complete the replacement with the €100 subsidy are smaller but their old refrigerators are larger. The difference seems to be driven by

provides placebo tests.

households in NRW that receive additional funding by the state government. Potentially, the lower federal subsidy makes the replacement less attractive for larger households in NRW which had a small own contribution to the purchasing price when the federal subsidy was set at €150.

2.5.2 Procedural variations

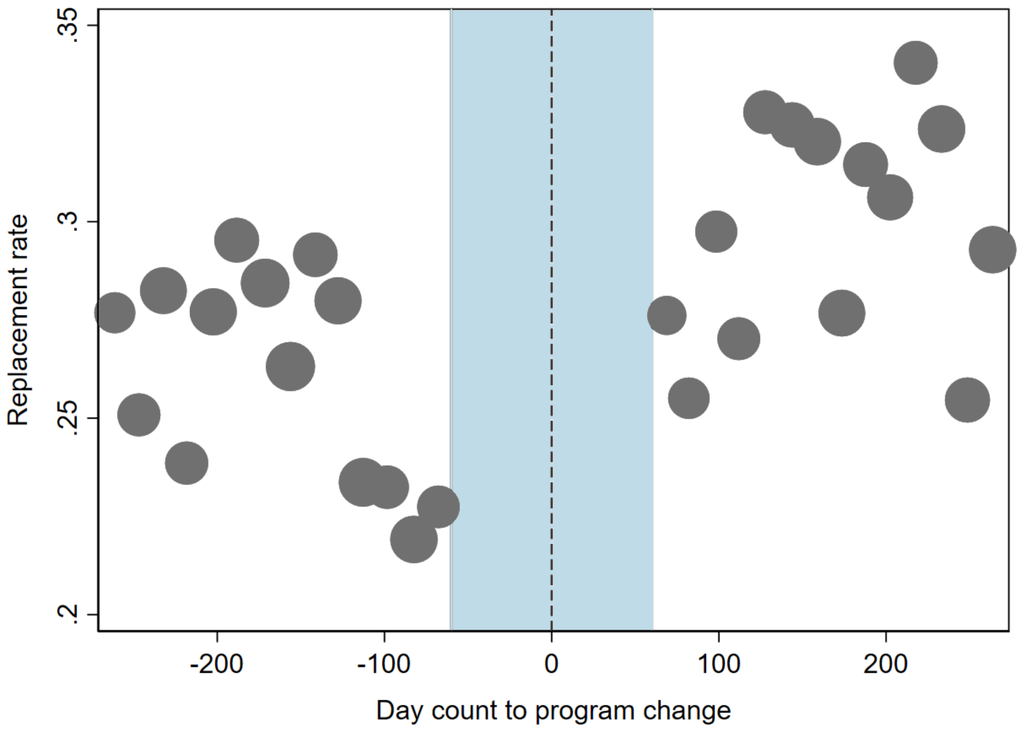
Figure 2.4 shows the replacement rate around the simultaneous procedural changes from automatic to elective enrolment and from flexible to rigid voucher terms. Day 0 is January 1, 2018. Negative day counts cover the period when enrolment was automatic and voucher terms flexible, positive day counts the period when enrolment was elective and voucher terms rigid. As before, each bubble captures the average replacement rate within a 15-day interval. The blue-shaded area marks the interim period of two months around the change. We exclude data points that fall in this period from the analysis in the Donut design. By inspection, the average replacement rate lies around 0.25 before the interim period: About a quarter of automatically enrolled eligible households redeem the €150 voucher upon replacing their refrigerator. After the interim period, the replacement rate rises to around 0.3: Around a third of eligible households elect to enrol in the RRP and successfully redeem the €150 voucher with rigid terms. A simple comparison in means for the average replacement rate in both regimes shows an increase of 6 percentage points (from 0.26 to 0.32, see Table 2.1).

Table 2.3 provides our estimation results. The specifications are analogous to the estimation of the subsidy effect. We estimate a positive coefficient that is statistically significant in all three specifications ($p=0.008$, $p<0.001$ and $p=0.015$ respectively), confirming the visual impression and the difference in means. In our preferred specification (column 3) including both the Donut design and the ALL, we estimate the effect of changing procedures at 4.2 percentage points.[41] In the more simplified RDD without Donut and ALL (column 2), the effect is larger in magnitude at 15 percentage points.[42] In the basic OLS specification (column 1), the effect is, at 4.6 percentage points, similar to column 3. Figure A2.20 shows how the treatment effect changes as function of the bandwidth. For models 2 and 3, the effect ranges from 3 (235 days bandwidth, model 3) to 19 percentage points (170 days bandwidth, model 2). 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 the effects of procedural to those of subsidy variation in a back-of-the-envelope calculation illustrates the merits

[41]The 2-month bandwidth of the Donut returns conservative estimates of the effect sizes: Narrower bandwidths lead to higher effect estimates that capture more of the transient noise and adjustments around the change. Wider Donut bandwidths also have elevated point estimates and greater variance.

[42]Appendix Table A2.28 provides robustness check results in a specification chart and Figure A2.26 provides placebo tests.

Figure 2.4: Procedural variation, replacement rate: Discontinuity



Notes: This figure shows the rate of households that successfully complete refrigerator replacement out of all households that are eligible for the RRP around the change in program procedures on January 1, 2018. Replacement rates are binned and averaged over 15 days in a bandwidth of 270 days around the program change. The blue-shaded area marks the interim period of 2 months around the change. In our main specification, we exclude data points that fall in this period from the analysis in a RD Donut Design (see Figure A2.16 in the Appendix for the raw data plot).

of alternative program designs: The procedural variation had a positive effect on the adoption of energy-efficient appliances that measured 0.7 to 2.5 times that of a €50 increase in the subsidy.[43] We do however not find both coefficients significantly different from each other at any bandwidth (see Figure A2.24 in the Appendix).

Table 2.3: Estimated effect of procedural variation on the replacement rate

	1	2	3
Procedural change (EE-RIG = 1)	0.046 (0.012)	0.150 (0.022)	0.042 (0.017)
Day count		yes	yes
Day count×EE-RIG		yes	yes
Location fixed effects	yes	yes	
Time fixed effects	yes		
Controls	yes		
ALL			yes
Donut			yes
No. observations	70,426	21,534	18,406

Notes: Standard errors in parentheses, clustered by branch or bootstrapped (ALL). Location fixed effects include energy advisor ID and local branch. Time fixed effects include month and 2-year indicators. Augmented Local Linear includes advisor IDs and branch indicators. The Donut design excludes 2 months around the program change. Column 1 uses the sample over the sample period 2009-2021. RDD estimates in columns 2 and 3 use the sample of eligible households in a bandwidth of 9 months around January 1, 2018.

We compare whether households that replace their refrigerators under automatic enrolment and flexible voucher terms versus elective enrolment and rigid voucher terms change in their observable characteristics (see Table A2.4). Mean comparisons suggest that households that successfully complete replacement under elective enrolment/rigid voucher terms are slightly larger (3.0 vs. 2.9 household members) and possess slightly larger (249 vs. 241 liters) and older (18.4 vs. 17.9 years) refrigerators. Their new

[43] The range is based on specification 3. The procedural effect is smallest in comparison to the subsidy effect at 185 days bandwidth and largest at 225 days.

refrigerators are also larger (268 vs. 259 liters) and consequently a bit more expensive (€476 vs. €462) with a higher kWh consumption (142 kwh vs. 139 kwh). Even though these mean comparisons are statistically significant, the magnitude of the differences is small from an economic point of view.[44]

2.6 Mechanisms

2.6.1 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 2.5 shows a discontinuity graph for the enrolment stage similar to Figure 2.4 for the replacement rate. 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 interim period (blue-shaded area): Around 40 percent of eligible households elect to enrol in the RRP for a subsidy of € 150. After the change in the subsidy and the interim period, the enrolment rate settles around 0.3: Roughly 30 percent of eligible households elect to enrol for a subsidy of € 100. During the interim period, enrolment rates are elevated.[45]

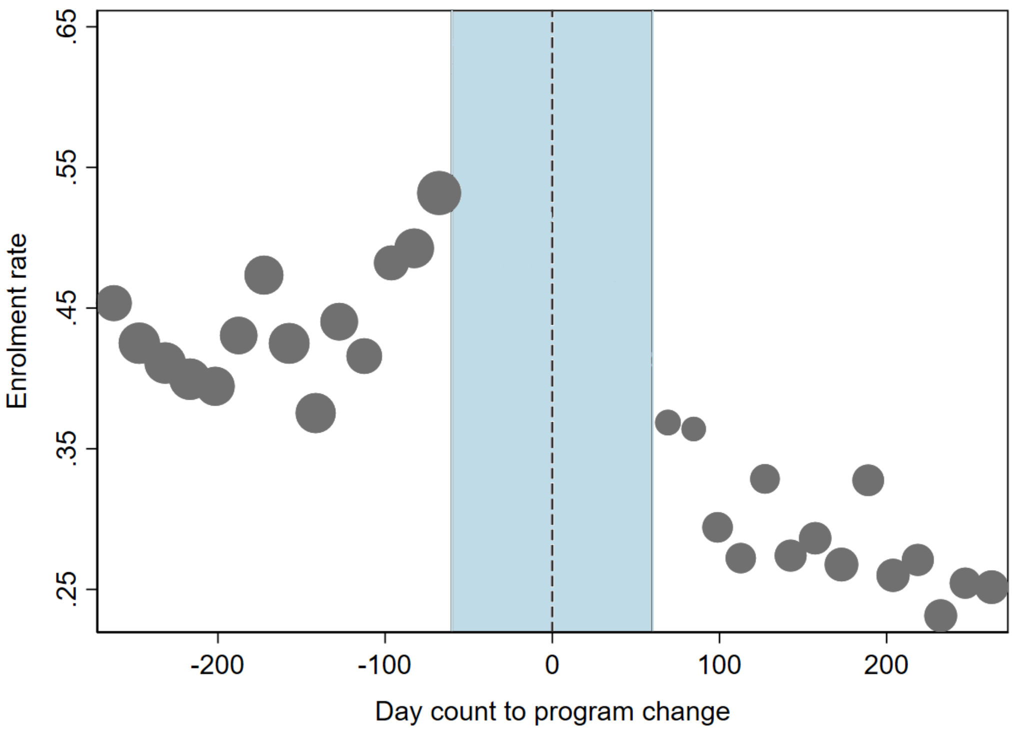
Table 2.4 provides estimation results, using the same specifications as for the replacement rate in Section 2.5. All specifications show a positive significant coefficient (p=0.002, p=0.002 and p=0.003 respectively), mirroring the results of our descriptive analysis.[46] In our preferred specification in column 3, we estimate the enrolment rate to be 8.8 percentage points higher for a € 50 higher voucher value. The other two models produce slightly larger estimates at 9.4 and 10.1 percentage points. In Figure A2.21, we show how the treatment effect changes as function of the bandwidth.

[44]We also check whether procedural change was accompanied by changes in program timing. Comparing the distributions of days passed between the first and second visit under the two regimes (see Figure A2.25, the patterns are statistically indistinguishable. This speaks against the conjecture that households managed to alter local branches' timing of the second visit after EE-RIG procedures had been introduced.

[45]An important factor in the elevated levels are irregularities in the issuance of the invitation letters to households during the interim period: Despite fulfilling the eligibility criteria, there is evidence of invitation letters being withheld (see the eligibility ratio in Appendix Figure A2.14 and explanations in Section 2.4). This has the effect of decreasing the denominator of the enrolment rate, driving up the enrolment rate.

[46]Appendix Table A2.29 provides robustness check results in a specification chart.

Figure 2.5: Subsidy variation, enrolment rate: Discontinuity



Notes: This figure shows the rate of households that request a voucher for refrigerator replacement and enrol in the program out of all households that are eligible for the RRP over time around the reduction of the voucher value by € 50 on April 1, 2019. Enrolment rates are binned and averaged over 15 days in a bandwidth of 270 days around the program change. The light-grey shaded area marks the interim period of 2 months around the change. In our main specification, we exclude data points that fall in this period from the analysis in a RD Donut Design (see Figure A2.17 in the Appendix for the raw data plot).

For a higher subsidy, we observe significantly more households electing to enrol in the program.

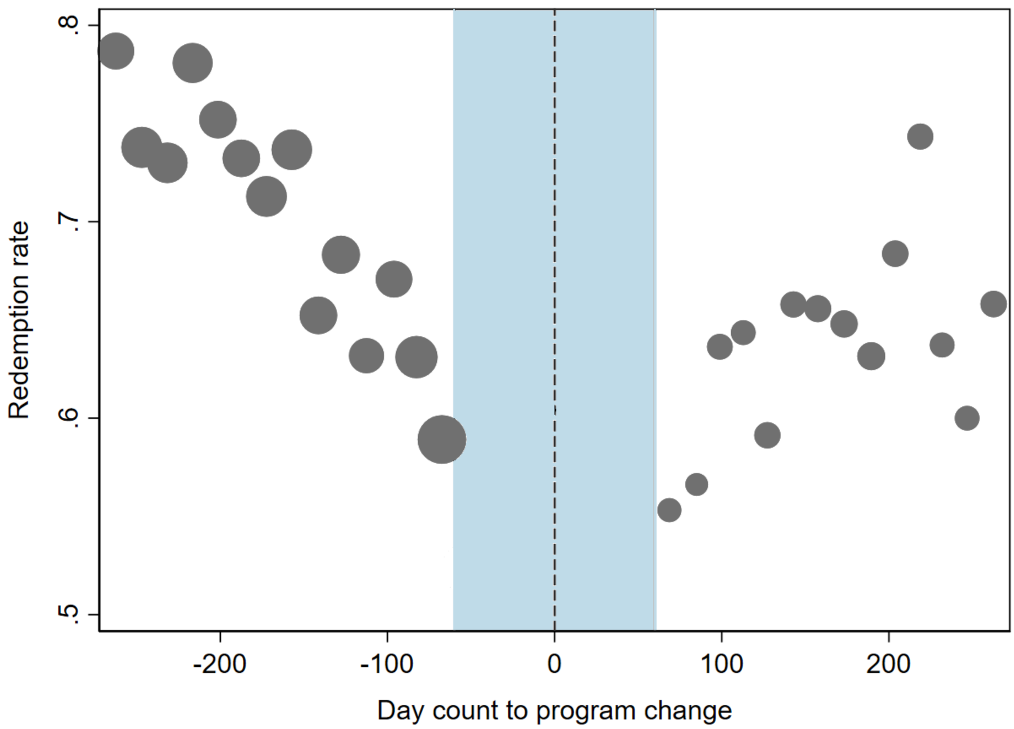
Table 2.4: Estimated effect of subsidy variation on the enrolment rate

	1	2	3
Subsidy (€ 150 = 1)	0.101 (0.032)	0.094 (0.029)	0.088 (0.025)
Day count		yes	yes
Day count×Subsidy		yes	yes
Location fixed effects	yes	yes	
Time fixed effects	yes		
Controls	yes		
ALL			yes
Donut			yes
No. observations	70,426	16,832	13,959

Notes: Standard errors in parentheses, clustered by branch or bootstrapped (ALL). Location fixed effects include energy advisor ID and local branch. Time fixed effects include month and 2-year indicators. Augmented Local Linear includes advisor IDs and branch indicators. The Donut design excludes 2 months around the program change. Column 1 uses the sample over the sample period 2009-2021. RDD estimates in columns 2 and 3 use the sample of eligible households in a bandwidth of 9 months around February 1, 2019.

The redemption stage of the replacement process is captured in the discontinuity graph of Figure 2.6. 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, and after the change in voucher value. By inspection, they lie in the range between 0.5 and 0.8 up to 270 days before the change and 0.5 to 0.66 up to 100 days before: One half to two thirds of enrolled households 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 households redeem their € 100 voucher. As a result, there is no clear effect visible at the redemption stage.

Figure 2.6: Subsidy variation, redemption rate: Discontinuity



Notes: This figure shows the rate of households that successfully replace their refrigerator out of all households that have requested a voucher and enroled in the program over time around the reduction of the voucher value by €50 on April 1, 2019. Replacement rates are binned and averaged over 15 days in a bandwidth of 270 days around the program change. The blue-shaded area marks the interim period of 2 months around the change. We exclude data points that fall in this period from the analysis in a RD Donut Design (see Figure A2.18 in the Appendix for the raw data plot).

Table 2.5 reports the formal estimation results, using the same specifications as in the previous models. The simple OLS model produces a small insignificant positive coefficient of 2.2 percentage points (p=0.529), while the two RDD specifications produce small negative coefficients of -2.1 and -4.5 percentage points (p=0.757 and p=0.297 respectively) that are insignificant as well.[47] In Figure A2.22, we show how the treatment effect changes as function of the bandwidth. For both specifications (columns 2 and 3), the effect remains insignificant across all bandwidths. We find a clear null effect for a higher subsidy value in the second-stage decision, conditional on voucher request. The intuition that households holding a voucher worth €150 rather than €100 are more likely to replace their refrigerator and redeem the voucher has therefore little empirical support.

Combining these insights, the effect of varying the subsidy estimated in Section 2.5 can be ascribed exclusively to a recruitment effect at the enrolment stage. This finding is relevant from a program management perspective, as we discuss in Section 2.7.

[47] Appendix Table A2.30 provides robustness check results in a specification chart.

Table 2.5: Estimated effect of subsidy variation on the redemption rate

	1	2	3
Subsidy (€150 = 1)	0.022 (0.034)	-0.021 (0.067)	-0.045 (0.044)
Day count		yes	yes
Day count×Subsidy		yes	yes
Location fixed effects	yes	yes	
Time fixed effects	yes		
Controls	yes		
ALL			yes
Donut			yes
No. observations	54,434	5,356	18,406

Notes: Standard errors in parentheses, clustered by branch or bootstrapped (ALL). Location fixed effects include energy advisor ID and local branch. Time fixed effects include month and 2-year indicators. Augmented Local Linear includes advisor IDs and branch indicators. The Donut design excludes 2 months around the program change. Column 1 uses the sample over the sample period 2009-2021. RDD estimates in columns 2 and 3 use the sample of households that have requested a voucher in a bandwidth of 9 months around February 1, 2019.

2.6.2 Procedural variations: Behavioral effects

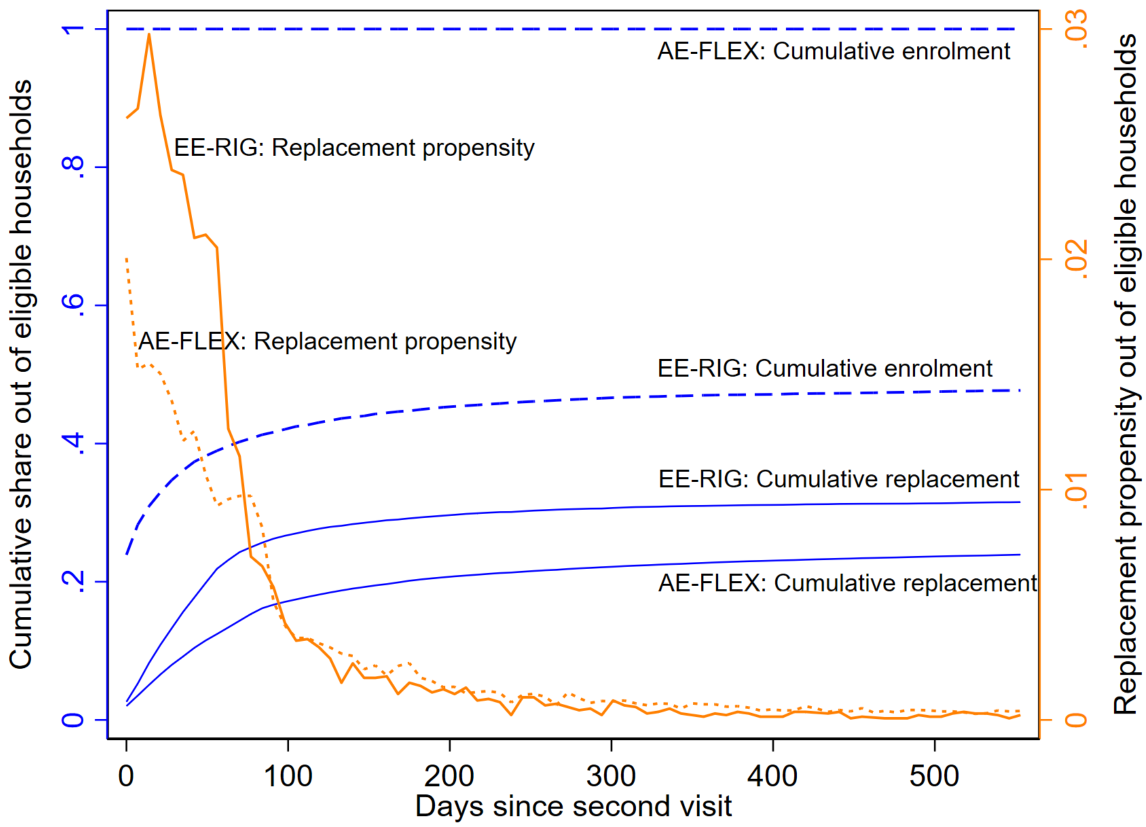
To understand more about the mechanisms behind the effect of procedural variation on the success rate of the RRP, we take a closer look at how the behavioral patterns before and after the procedural changes compare.

Figure 2.7 shows, as a function of days passed since the second home visit, three temporal patterns, two cumulative (in blue, left scale) and one intensive (in yellow, right scale), under the regimes AE-FLEX/EUR150 and EE-RIG/EUR150. 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): Cumulative enrolment of eligible households (blue, left scale) mechanically jumps to 100 percent on the day of the second visit. After the change, enrolment is elective (EE-RIG): Cumulative enrolment starts at around 20 percent of eligible households that enrol on the day of the second visit and grows at a slowing rate to top out at 44 percent. 90 percent 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 plausibly be traced to selection effects driven by 'hassle', 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 enrolment 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 indicator 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 32 percent 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, 24 percent of eligible households replace their refrigerator up to 550 days after the second home visit, 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 are an obvious factor and can explain why cumulative replacement under EE-RIG is *not lower* than under AE-

Figure 2.7: Behavioral patterns of cumulative replacement and replacement propensities by procedural regime



Notes: This figure shows the cumulative enrolment and replacement rates on the left-hand y-axis and replacement propensity on the right-hand y-axis for the automatic enrolment mode with flexible voucher terms (AE-FLEX) and the elective enrolment mode with rigid voucher terms (EE-RIG) respectively as a function of days passed since the second home visit. The data for AE-FLEX and EE-RIG cover the periods January 2009 to December 2017 and January 2018 to January 2019 respectively.

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 percent 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 percent as the first voucher approaches the end of its 90-day validity period. 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 deadline approaches. 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 percent 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 counterparts. 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. Rigid terms 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 change to rigid terms could therefore demotivate those eligible households that 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 a 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 could be thought of as strengthening the implementation intention by supporting 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 from straying from the intended path. In the RRP, some households that would not have completed the replacement within 90 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 a spike in the redemption probability (see Appendix Figure A2.23). These insights highlight the potential to use behaviorally informed procedural changes, such as goal setting, in the future in an effort to target more narrowly the motivation effect detected here.

2.7 Policy assessment

A key finding of our empirical analysis is that the procedural changes had a similar magnitude of impact on program performance as a €50 subsidy change.[48] This is subject to the qualification that the empirical setting only allows us to measure the impact of the subsidy change after the procedural change has been introduced. A conservative back-of-the envelope calculation of the subsidy increase needed to generate an effect equivalent to the procedural change suggests an additional €35 per household, using the estimate from the least favorable bandwidth choice.

To be useful for policy-makers and program managers, a comparison of effect sizes needs to be extended to considering not only their relative benefits, but also their costs or savings. An exhaustive assessment requires information often not available to the researcher.[49] As an indication, however, consider that lifting the average replacement rate by 5 to 7 percentage points by raising the voucher value from €100 to €150 not only means a higher productivity of each costly home visit conducted, but also raises the cost per replacement by 50 percent. One reason for the limited productivity is the lack of impacts at the redemption stage, as seen in the previous section. The procedural changes, on the other hand, not only boost the replacement rate by 4 to 15 percentage points, but do so at negligible and arguably even negative costs since fewer vouchers have to be issued and kept on the balance sheet.

The estimated effect sizes can also be used for counterfactual program scenarios. One scenario of interest is an alternative setting in which elective enrollment and rigid terms would have been introduced in 2013, right when the RRP was scaled up to its current size. Our point estimates in section 2.5 suggest at least 420 additional refrigerator replacements for every 10,000 invitations letters issued to eligible households. Between 2013 and 2017, 48,615 households were found eligible for replacement. Extrapolating the Local Average Treatment Effect from the RD estimation of our main specification while assuming a constant effect over time (4.2 percentage points, 95% CI: [3.9 pp; 4.5 pp]), we conjecture that bringing the procedural change forward to 2013 would have led to at least 2,042 (= 0.042 x 48,615) [95% CI: 1,896; 2,188] additional refrigerator replacements and additional aggregated savings of €201,800 per year.[50]

[48] The relative effect size depends on the choice of bandwidth and is smallest at 0.7 with 185 days of bandwidth and largest at 2.5 with 225 days (see Figure A2.24).

[49] As an illustration, we were not granted access to detailed information about important cost components of the RRP, such as salaries, measurement costs and database management costs

[50] This calculation uses the average electricity price of €0.289 in 2020 and average annual savings of 342 kWh. This would have led to average additional annual savings of €99 in electricity bills for replacing households, or aggregated annual savings of €201,800.

2.8 Conclusion

Our paper shows in the context of investment subsidies to low-income households for energy-efficient appliances that even seemingly small and arguably cost-saving procedural changes can substantially substitute for monetary inducements. In this, it adds to a growing literature in behavioral public policy that demonstrates how program design affects program performance. As a novel element, the empirical opportunity of the setting allows to express the impact of these procedural changes in a money metric.

The rich data made available by RRP management offer a rare glimpse into the 'black box' of consumer durable replacement decisions among the poor. Their analysis allows us to report three main findings. One is the subsidy elasticity of replacement decisions: A 50 percent higher subsidy increases the likelihood of refrigerator replacement by 5 to 7 percentage points, attributable to more households enrolling in the RRP. The second is how replacement rates are affected by procedural changes. These rates are 4 to 15 percentage points higher under elective enrolment and rigid terms than under automatic enrolment and flexible terms, with patterns that are consistent with self-selection and a behavioral mechanisms such as goal setting and time management by households. Our third main finding is that conservatively comparing the subsidy and the procedural variation, the arguably accidental changes in procedures were equivalent – in terms of replacement rates – to raising the subsidy by €35. These numbers give an 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 2,000 refrigerators that could have been replaced if the new procedures had been in place from 2013 onward. 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.

Future research can build on these findings in three ways. One is to (re-)evaluate existing programs with a view to uncovering more effects of procedural changes on program performance. Many small changes happen for reasons other than deliberate program optimization. The RRP is a case in point. 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 richer understanding of how and why procedures matter for program success.

The second way is to make progress towards theoretically and empirically informed procedural changes. Rather than accidental or driven by expediency, 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 vouchers 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 third way is to explore the optimal integration of economic incentives and procedures for program design. Design optimization was not part of the agenda in RRP. 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.

Chapter 3

Transporting behavioral insights to low-income households: A field experiment on energy efficiency investments

Coauthor(s): Timo Goeschl, Martin Kesternich and Madeline Werthschulte

Abstract: Many industrialized countries have recognized the need to mitigate energy cost increases faced by low-income households by fostering the adoption of energy-efficient technologies. How to meet this need is an open question, but "behavioral insights" are likely components of future policy designs. Applying well-established behavioral insights to low-income households raises questions of transportability as they are typically underrepresented in the existing evidence base. We illustrate this problem by conducting a randomized field experiment on scalable, low-cost design elements to improve program take-up in one of the world's largest energy efficiency assistance programs. Observing investment decisions of over 1,800 low-income households in Germany's "Refrigerator Replacement Program", we find that the transportability problem is real and consequential: First, the most effective policy design would not have been chosen based on existing behavioral insights. Second, design elements favored by these insights either prove ineffective or even backfire, violating 'do no harm' principles of policy advice. Systematic testing remains crucial for addressing the transportability problem, particularly for policies targeting vulnerable groups.

Publications: This chapter is available as ZEW Discussion Paper No. 24-079.

Acknowledgments: We would like to thank the SSC staff for the partnership and John List, Rob Metcalf, Daan van Soest, Alessandro Tavoni, participants of the EAERE 2023, AFE 2023 and 2024, NOeG 2024, and seminar participants at Bochum, Bologna, Grenoble, Heidelberg, IfW Kiel, Mannheim, and Würzburg for valuable feedback. Max Stellbrink and Carina Haller provided excellent research assistance. The experiment was pre-registered at the AEA RCT Registry as trial AEARCTR-0010757. This work has been funded by the Fritz-Thyssen-Foundation.

3.1 Introduction

Rapidly rising energy prices in many developed countries have highlighted to policy-makers the need to design and implement targeted policies for increasing energy efficiency among low-income households. While all households are negatively affected by higher expenditures for residential energy consumption, low-income households are particularly exposed. They already spend a large share of their disposable income on energy and their energy demand tends to be even less elastic compared to that of an average household (Schulte and Heindl, 2017). Many countries have recognized this urgent need. In the US, the Green and Resilient Retrofit Program funded under the Inflation Reduction Act has recently strengthened its effort to improve energy and water efficiency in low-income families (US Department of Housing and Urban Development, 2024). Similarly, in Europe, the EU Directive 2023/955 obligates Member States to submit plans to "prioritize energy efficiency improvements for vulnerable customers, low-income households, and individuals in social housing". As public funds become increasingly scarce, these plans are likely to feature increased attention on low-cost policies: Managers are looking for effective but low-cost policy options, making behavioral interventions ("nudges") an obvious candidate for future policy approaches.

Designing behaviorally-guided programs to achieve sizeable effects among a specific target group poses particular challenges for evidence-informed policy making. While policy-makers may be informed on "behavioral insights" (Halpern and Sanders, 2016; Gopalan and Pirog, 2017) emerging from an expanding set of carefully executed studies, that evidence base is often silent on the estimated effects of candidate policies on a specific subgroup like low-income households. The reason is that these households are typically absent or, at a minimum, systematically under-represented in untargeted programs. Consequently, applying these evidence-based insights to economically disadvantaged households raises concerns of "transportability": Can a policy-maker or program designer be confident that insights from specific interventions tested successfully elsewhere "will also work" for a new target group (Halpern and Sanders, 2016; Hallsworth, 2023)? Failure of transportability can lead to disappointing policy outcomes or, worse, to negative impacts on (vulnerable) target populations, possibly violating ethical rules of 'do no harm' in policy advice (Harrison et al., 2020).

This problem is common in evidence-based policy-making. Health policies, for example, increasingly rely on empirical evidence from Randomized Controlled Trials (RCTs). RCTs in this domain tend to be unrepresentative of the patient population (Goldstein et al., 2019) and often fail to feature patients with certain characteristics, such as ethnic minorities (Duma et al., 2018).[1] When policies subsequently implement measures

[1]The reasons for certain subgroups not being included can be both intentional and unintentional.

to reduce health disparities, for example by expanding access to certain treatments, results often differ from expectations because the treatments turn out to perform less well among the new subgroups (Essien et al., 2021; Degtiar and Rose, 2023). Likewise, in development economics, a policy-maker may want to implement a program to improve the nutritional status of pregnant women in her country. Yet, the only available evidence may come from other behavioral intervention programs that happened not to include pregnant women (Duflo et al., 2007).

Transportability is increasingly attracting the attention of researchers in many policy-related fields (Pearl and Bareinboim, 2011; Westreich et al., 2017; Dahabreh and Hernán, 2019; Degtiar and Rose, 2023). Narrowly speaking, it refers to the degree to which internally valid evidence on the effects of an intervention derived for a study population can be extended to infer its effects on a particular target population of political interest that were not part of the original study population. More broadly speaking, it refers to the degree to which results from a policy experiment in context A can be expected to hold in context B (Marchionni and Reijula, 2019; Francesconi and James, 2021). Beyond these narrow and broader conceptualizations of transportability lie the possible welfare effects ('do no harm') of not taking poor transportability into account when designing or changing policies.

Current calls for energy efficiency assistance programs targeted at low-income households also have to contend with problems of transportability. Most evidence on the effectiveness of policies to increase household energy efficiency derives either from observational data or RCTs from untargeted programs in which low-income households are absent or, at a minimum, systematically under-represented. Household income appears to be a relevant dimension for transportability of behavioral insight: Studies in other areas, for example health insurance (Domurat et al., 2021) and consumer debt (Holzmeister et al., 2022), have found that low-income households respond differently to seemingly well-established behavioral insights, sometimes even in the opposite direction. Deploying behavioral insights to the energy efficiency investment choice of low-income households could therefore also give rise to null effects or even negative outcomes.

In this paper, we empirically examine the transportability of behavioral insights in the context of an initiative aiming to improve one of the world's largest energy efficiency assistance schemes. One program within this scheme, the Refrigerator Replacement Program (RRP), offers cash incentives to low-income households for replacing old and

Medical RCTs, for example, intentionally exclude patients with co-morbidities, existing prescriptions, drug abuse and unintentionally struggle to recruit from marginalized groups. Likewise in education research, it has become apparent that rigorously establishing the effects of a policy such as charter schools and remedial training in one setting does not guarantee that the policy will have the same effects in another setting (Banerjee et al., 2007; Cohodes and Parham, 2021).

inefficient refrigeration appliances with new, highly efficient ones. Improvements to the RRP were sought against a backdrop of modest program performance as seen by management and sponsor: A take-up rate of about 25 percent among all eligible households was regarded as improvable given that program eligibility requires passing an individualized cost-benefit test and that verifying eligibility is a major cost factor for the program. The performance of the RRP has proven to be responsive to small changes in its design: In the past, unsystematic procedural changes had substantial impact on program take-up (Chlond et al., 2025), opening a window for targeted experimentation. The RRP's management agreed to partnering with researchers in order to estimate treatment effects of several candidate program improvements and compare those with the average treatment effect of management's own baseline, rather than bringing in an outside consultancy to receive a single "behaviorally informed" improvement proposal. The possible shape and form of candidate improvements was limited, however, by stringent design constraints common to the public sector, such as adherence to administrative procedures and not adding expenses to program administration (Della Valle and Bertoldi, 2021).

The paper reports on the resulting study, which takes the shape of a co-designed RCT that compares the effects of eight treatments. Six treatments are candidate improvements to be tested for transportability, one is the management's own program baseline, and one is a legacy design. The candidate improvements had two targets: One target was the presentation of the appliance replacement opportunity in the "program information letter", a critical feature in many public programs (Bhargava and Manoli, 2015; Finkelstein and Notowidigdo, 2019; Hotard et al., 2019; Linos et al., 2022). This letter is handed to households on the occasion of an audit visit and is intended to make them enrol in the RRP based on a solid understanding of the program. A review of the literature identified three main variations that had a track record of positive treatment effects on households' energy savings elsewhere[2] while being consistent with the design constraints. One was a visual enhancement of the economic effects of the replacement opportunity (Allcott and Greenstone, 2017; Stojanovski et al., 2020). In part, this variation was already included in the management's own baseline. The second variation was the introduction of loss framing (Gonzales et al., 1988; Homar and Cvelbar, 2021) in how the replacement opportunity is presented. The third variation was the use of peer experience with the appliance replacement (Allcott, 2011; Andor et al., 2020) when presenting the RRP in lieu of an individualized forward projection. Notwithstanding their general track record, these variations are essentially untested in the context of low-income households because such households are typically under-represented in this literature (Allcott, 2011; Fowlie et al., 2015).

[2]See Khanna et al. (2021) for a recent meta analysis.

The other target of candidates for improving the RRP was post-visit engagement with the low-income households. Here, a review of literature identified the use of reminders through letters, text messaging, and visual cues left in the household as a likely program improvement. This was based on amassed evidence on the positive effects of reminders on program take-up (Guyton et al., 2016; Gravert, 2021) and a literature that shows that low-income households are particularly likely to be subject to salience problems and cognitive scarcity (e.g., Shah et al. 2012; Haushofer and Fehr 2014), therefore potentially benefiting from repeated engagement some time after receiving the information letter. As before, however, the studies documenting the effectiveness of reminders contained few to no observations from low-income households.

Our experimental design does not only allow for an examination of the transportability of tried and trusted candidate improvements to the RRP, but also generates effect estimates that themselves make progress against established criteria of external validity, such as the 'SANS' (selection, attrition, naturalness and scalability) conditions (List, 2020). We randomly selected 21 trial sites for participation in our natural field experiment (Harrison and List, 2004). All treatment variations had to survive a demanding co-design process involving program management and program staff to ensure effortless administration, high naturalness, and full scalability to all 150 program sites in the country, including a pilot trial at one site. Throughout the entire experimental process, we continually raised awareness among program and site managers to disclose potential attrition. No attrition was reported to us among treated units.

Based on the observed behavior of 1,803 low-income households over the course of one year, we conclude that the transportability problem is both real and consequential when trying to design energy efficiency assistance programs targeted at low-income households. The unique opportunity afforded to us by program management succeeds in testing – with respect to take-up – candidate program improvements that are consistent with the program constraints, are scalable, and have a high degree of external validity. That test shows, however, that none of the candidate improvements outperforms management's own baseline, a visually enhanced presentation of the replacement opportunity without a reminder. Worse, some candidate improvements significantly decrease program take-up, with unexpectedly bad performance for loss framing the replacement opportunity. Unexpectedly, adding reminders can also backfire. The evidence points to the most vulnerable households among the low-income households as those program participants for whom transportability breaks down to the point of reversing the direction of treatment effects. The evidence on framing is consistent with the idea that in terms of prospect theory, low-income households tend to use reference points for assessing gains and losses from investing in energy efficient appliances that are significantly shifted to the left of that used by the average household behavioral insights have been obtained from (Kőszegi and Rabin, 2006). It is also consistent with evidence that loss tolerance is more widespread than previously considered and more prevalent among those with experience of adverse financial shocks (Chapman et al., 2024). The evidence on reminders is consistent with the idea that the announcement of a reminder leads to lower task completion because anticipated reminders undermine own investment in imperfect memory (Ericson, 2017). This may be particularly relevant to cognitively stressed low-income households. Our results add important texture to the demands of "Do No Harm" policy making: Should it be the case that the most vulnerable groups are most at risk from poor transportability, then this poses additional challenges for policies supposed to address their needs because additional safeguards and tests could restrict options and cause delays.

The paper is structured as follows. Section 3.2 highlights our contribution to the existing literature. Section 3.3 describes our experimental design, providing detailed information on the procedures of the energy efficiency assistance scheme, the RRP and the subsidy voucher (Section 3.3.1), on the treatment development and rationale (Section 3.3.2), and on the selection and training of the program sites where we conducted the RCT (Section 3.3.3). We next describe our household sample in Section 3.4 and our empirical strategy in Section 3.5. The presentation of the results follows three steps. First, we focus on the effects of the treatments targeting the information letter in Section 3.6.1, and second, we focus on the effects of the reminder effects targeting the post-visit engagement in Section 3.6.2. Third, Section 3.6.3 focuses on heterogeneous treatment effects by the type of federal income support received, which we use as a proxy of vulnerability and aspiration differences among our low-income sample. Section 3.7 concludes.

3.2 Related literature

Our study lies at the intersection of three strands of literature. First, our paper contributes to the emerging literature investigating the effects of behavioral interventions on the take-up of governmental welfare and public assistance programs. Recent studies investigate take up of Earned Income Tax Credit (EITC) benefits (Linos et al., 2022; Bhargava and Manoli, 2015), unemployment benefits (Bruckmeier et al., 2021), the SNAP food stamp program (Finkelstein and Notowidigdo, 2019), claims for tax refunds (Bronchetti et al., 2013), waivers for citizenship applications (Hotard et al., 2019), medicare insurance (Brot-Goldberg et al., 2023) or energy subsidies (Hahn and Metcalfe, 2021). As a common finding of these evaluations, take-up rates are usually rather modest. For example, only 14 percent claim the EITC benefits (Bhargava and Manoli, 2015), 6 percent of eligible low-income households enrol in the SNAP

(Finkelstein and Notowidigdo, 2019) and 4.3 percent claim energy subsidies (Hahn and Metcalfe, 2021). At the same time, this literature reports mixed evidence on the effectiveness of different behavioral interventions including presentational changes in program description in letters and reminders on program take-up among low-income households. Bhargava and Manoli (2015) observe that enhancing information letters by simplifying the explanation of process and salient program benefits can increase EITC take-up from 14 to 31 percent. In the context of fee waiver applications, Hotard et al. (2019) show information nudges to be effective as they increase application rates by 8.6 percentage points. Similarly, Finkelstein and Notowidigdo (2019), find positive effects of an information letter on SNAP take-up, in particular if the information letter further includes assistance information. Looking at the impact of different mail reminders on tax compliance, Hallsworth et al. (2017) report that the reminders increase tax compliance among delayed tax payers. While there is no significant difference between gain- and loss-framed reminders in the whole sample, loss-framed messages had a particular strong impact for those with the largest debt. Hahn and Metcalfe (2021) investigate the impact of different information letters on enrollment rates into an energy subsidy program. While receiving a letter increases enrollment rates significantly, the particular behavioral content of the letters appears to be less important with, e.g., no differences in enrollment rates between a gain and a loss frame. Likewise, Linos et al. (2022) report null-effects of information letters that vary content, design, messenger and mode on EITC take-up. Some studies additionally examine the role of default sets: While Brot-Goldberg et al. (2023) show that default rules can have large and persistent effects on enrollment and drug utilization in a voluntary drug benefits program, Bronchetti et al. (2013) find no significant default effect among low-income tax filers. We contribute to this literature in several ways: One is the setting, which brings international evidence from Germany to bear on the question of how take-up can be improved. Another is the nature of the assistance program: The RRP does not provide support in cash or in kind. Instead, it incentivizes investment decisions with medium-term cash-flow benefits to the household. Such programs raise new issues that merit attention due to their potential for sustained improvements in household finances, but also due to non-trivial welfare aspects of diverting cash from consumption to investment.

Second, the paper contributes to the literature on transportability of empirical evidence across study settings, in particular in public policy (Pearl and Bareinboim, 2011; Dahabreh and Hernán, 2019; Degtiar and Rose, 2023). One manifestation of a transportability problem is when priors for expected effect sizes that derive from existing empirical evidence are subsequently confronted with divergent evidence from a new study population. The inconsistent results reported by the literature on take-up of assistance programs discussed above are a case in point that transportability cannot be assumed. Like generalizability, transportability is an aspect of the external validity of extending inferences beyond the study sample, but is distinct from the former (Westreich et al., 2017): The problem of generalizability arises when the study sample is a strict and possibly non-random subset of the target population. The problem of transportability arises when the study sample is not a subset of the target population (Duflo et al., 2007).[3] When there is insufficient overlap, an internally valid sample average treatment effect (SATE) of a policy may not allow valid inference to its specific target average treatment effect (TATE) (Goldstein et al., 2019). Understanding more about the extent to which policies designed for particular target groups can rely on insights from interventions tested elsewhere is urgently required (Halpern and Sanders, 2016; Hallsworth, 2023). This is particularly relevant when the specific target group consists of vulnerable people, given the possible ethical implications of a failure to transport (Harrison et al., 2020). We contribute to this literature by conducting an RCT in which we explore the transportability to low-income households of behavioral insights that have consistently performed successfully in the general population. By showing that transporting some of these insights not only fails to improve performance, but can lower it, we demonstrate the importance of transportability in behavioral public policy.

The third strand of literature to which our experiment contributes deals with the impact of behavioral interventions on energy savings and efficiency, but with a particular focus on low-income households. While information nudges perform well in the field of energy efficiency investments within broader population samples (Khanna et al., 2021), little is known on the effectiveness of behavioral interventions on investments in low-income households. One particular reason is that empirical evidence from non-targeted energy savings programs provide only limited insights as low-income households are more likely to drop out of those programs compared to an average household. For example, Löschel et al. (2023) find that low-income households are less likely to adopt a cost-free energy saving app. Allcott (2011) reports that low-income households are more likely to stop receiving the cost-free home energy reports (HERs), which aligns with follow-up work stressing a relatively lower willingness to pay for such reports among tenants (Allcott and Kessler, 2019). Closest to our work are studies by Fowlie et al. (2015, 2018) reporting low take-up of financial incentives among low-income US households for energy efficient building weatherization, even though the gains of doing so are high. Using a randomized encouragement design Fowlie et al. (2015) find that despite massive additional expenditures ($1,000 per audited household), audit take-up only moderately

[3] While agreeing on the lack of overlap between study and target population (Hotz et al., 2005; Allcott, 2015), the literature has not yet converged on a single definition of transportability. See Dahabreh and Hernán (2019) and Degtiar and Rose (2023) for a discussion.

increases from 1 percent in the control to 6 percent in the encouraged group. The energy efficiency literature has recently started to provide a more nuanced picture on the transportability of behavioral interventions to new study populations. For the case of HERs, Andor et al. (2020) report the effects of social comparison-based HERs on residential electricity consumption to be significantly lower for targets groups other than US residents. Similarly, Bonan et al. (2021) show that prime-augmented HERs may even backfire if they address customers who hardly engage in pro-environmental behavior. Our study contributes by providing additional evidence from low-income households on the heterogeneity of treatment effects in the context of energy efficiency. This evidence not only includes the presence of weak, but also the presence of negative effects on program take-up.

3.3 Experimental design

We implement our RCT within the largest energy efficiency assistance scheme in Germany, the "Energy-Saving-Check." In the following sections, we first describe the program (3.3.1) and then turn to the experimental variations and the hypotheses (3.3.2). After this, we explain the roll-out of the experiment in the selected local program sites (3.3.3).

3.3.1 The "Energy-Saving-Check" and the "Refrigerator Replacement Program"

The "Energy-Saving-Check" (SSC, German: *Stromspar-Check*) is a nation-wide program that aims at lowering the energy bills of low-income households in Germany by reducing their electricity and water consumption. The SSC is 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). Annual funding of around 10-15 million Euro 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. Within the SSC, the Refrigerator Replacement Program (RRP; German: *Kühlgeräte - Tauschprogramm*) has been offering cash vouchers to households on federal income support[4] in order to encourage replacing their old and inefficient refrigeration devices with modern,

[4]To qualify, the household needs to receive 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. In 2020, more than 7 percent of German households qualified on this basis (Bundesagentur für Arbeit, 2020).

highly efficient models. The RRP started on January 1, 2009 and was scaled up to its current size with the start of the second funding cycle of the SSC ("SSC plus") in April 2013.

The recruitment of qualified households into the SSC's home energy audit takes place through a variety of channels. The program is actively promoted in employment and social assistance agencies through both printed and audiovisual material. In addition, active staffers provide individualized descriptions of the program using pop-up booths in shopping streets and malls. Some local branches of the social assistance agency mandate the participation of households with excessively high energy bills. The SSC also maintains a website where information is available about the RRP in eleven languages. Additionally, recruitment takes place directly through 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 Euros 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. 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.[5]

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.[6] Differences in refrigerator efficiency can therefore significantly impact residential 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 a replacement with the most energy efficient class of devices on the market.[7] The expected financial savings are communicated to

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

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

[7]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+++

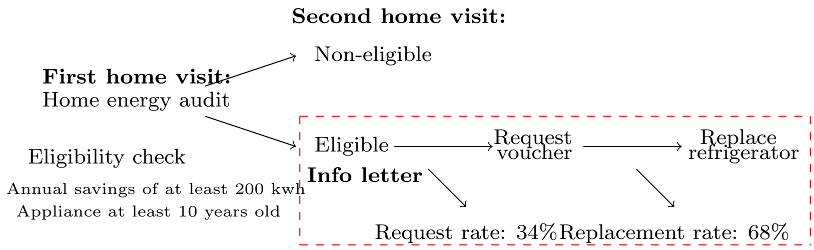


Figure 3.1: Procedure of the home audits.

the household via an information letter during the second visit. The investment of *every* eligible household passes a cost-benefit evaluation. Figure 3.1 summarizes the procedure of the two visits.

Figure 3.2 displays the letter that all eligible households receive during the second visit. It contains all relevant information and provides the basis for our experimental variations. First, it informs households that they meet the necessary eligibility criteria for the RRP and provides an estimate of the expected annual electricity savings (in EUR) from successful replacement (see green box in Figure 3.2). Second, it offers a step-by-step explanation on how to request and redeem the voucher: (1) eligible households *request* the voucher at one of the local sites, (2) they *replace* their refrigerator with a new model and (3) *redeem* the voucher in cash after successful replacement. To be able to redeem the voucher, a number of criteria have to be met: Households need to present their purchase receipt, document that the purchased device is of EU Energy Label class A, B, C or D; and provide proof that the old 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. Once requested, the voucher is valid for two months. Consequently, there exists a sharp 'deadline' on when the voucher expires.

Our experimental interventions vary the presentation of key information, as – common in many public policy assistance programs – the letter is a crucial bottleneck in the process. As Figure 3.1 shows, only 34 percent of all eligible households request the voucher. Conditional on the request, 68 percent of the households then successfully replace their refrigerator. The overall replacement rate of 23 percent consequently substantially falls short of the 100 percent replacement target that the program designers optimally would aim for considering the cost-benefit evaluation of the investment that every eligible household passed.

in classes A to F. Transitional arrangements were in place both in the retail sector and in the RRP.

Recipient
Street name, House number
Postcode City
Telephone number
Stromspar-Check

Mr / Ms / Family
Surname
Street name, House number
Postcode City

Important information from Stromspar-Check Aktiv on exchanging your refrigerator

Dear Mr/Ms/Family ...,

We measured your old refrigerator as part of the Stromspar-Check ('energy efficiency check') and, when analysing the values, found out that it would be worth exchanging your appliance.

If you purchase a new, highly efficient appliance, you could save around XX euros per year in electricity costs!

If you exchange your old appliance, you will receive a subsidy of 100 euros* for the purchase of your new appliance. We will need to issue you a voucher for this amount before you buy the new appliance. The number of vouchers is limited. Please contact your nearest Stromspar-Check location before you purchase.

It is extremely important that you note the following points:

- Your new refrigerator must have the energy efficiency class **A, B, C or D** of the new energy efficiency label (valid from March 1, 2021) and may not exceed energy consumption of **xxx kWh** per year.
- As soon as you have found a suitable appliance and you can finance it, call your nearest location (telephone number: xx/xxxxx). The Stromspar team will issue an individual voucher and send it to you. Please note that the **voucher is only valid for two months from the date of issue (up to a maximum of 11.02.2022).** It is not possible to extend this period if the voucher has expired!
- **Only buy your new refrigerator after you have received the voucher from us, because the purchase date has to be within the voucher's validity period.**

Please also note:

- Your old appliance has to be properly disposed of and we require written proof of disposal.

After purchase and disposal, you can redeem your voucher from the nearest Stromspar-Check location. To do so, you need to bring the following:

- the voucher
- the original proof (or receipt) of proper disposal
- the original purchase receipt of the new refrigerator (we will make a copy)
- the energy efficiency label of the new refrigerator (A, B, C or D sticker for the energy efficiency class)
- your ID card or passport.

Should you have any questions or require further information, feel free to contact us at your nearest location before your purchase.

Best regards,
Your Stromspar-Check Team

* = We cannot pay more than the purchase price of the refrigerator. The voucher does not cover, as applicable, ensuing disposal, transport and delivery costs.

Figure 3.2: Letter informing about voucher eligibility.

Note: Green highlights are added to the original letter. Red highlights are not added and part of the original letter.

3.3.2 Treatments and hypotheses

In a co-design process with the program officials and local site managers we jointly developed and pre-screened a set of treatment variations within or alongside the information letter. The aim of this procedure was to ensure effortless administration, high naturalness, and full scalability to all program sites in the country. Consequently, in April 2020, we started with a set of interactive online workshops where we jointly identified potential barriers for assessing the most economically relevant information provided in the letter and discussed a first set of potential interventions. It turned out that sensible language is a very important condition.[8] The treatments were then refined in a series of follow-up workshops and pre-tested in a pilot on site (in Frankfurt) from July to December 2021. Based on the lessons-learned from the pilot, we organized a final workshop in March 2022 to present the finalized experimental design and the procedures to the sites selected to participate in the RCT (see Section 3.3.3). The roll-out started in April 2022.

Inspired by well-established insights from behavioral economics, our interventions target three main dimensions. First, we distinguish whether we frame the reported annual savings in electricity cost as a financial GAIN or as a financial LOSS if households miss the opportunity of replacement. Here we differentiate between our first three treatments. The GAIN treatment simply reflects a legacy version of the information letter as shown in Figure 3.2. They key sentence in GAIN reads "If you purchase a new, highly efficient appliance, you could save around ... euros per year in electricity costs!" (green box). The RRP management planned to extend this legacy condition by adding a small purse pictogram with money falling in to the left of the text (see Figure 3.3a) and to introduce this extended version as the new baseline, replacing the legacy version. We refer to the management's own baseline version as GAIN^{+} treatment. In contrast to the management's baseline, a LOSS^{+} treatment points out the expected foregone savings in annual electricity costs from non-replacement, with the purse pictogram rotated by 180 degrees and money falling out (see Figure 3.3b). Here, the key sentence reads as follows: "If you do not replace your old refrigerator with a new highly efficient one, you will miss out on saving ... euros per year in electricity costs!"

As a second dimension, we randomize whether the reported annual savings in electricity costs stem from individual-level, appliance-specific engineering estimates (as in the legacy and baseline versions of the information letter) or from actual replacements being recently conducted in households with similar characteristics (i.e., with respect to the composition of household members). The basic idea of the first set-up is to

[8] E.g., we proposed a social norm intervention in the spirit of Allcott (2011), which was rejected by site managers. They were afraid of social pressure resulting from the letter and an additional burden placed on the low-income households.

(a) GAIN^{+} Frame.

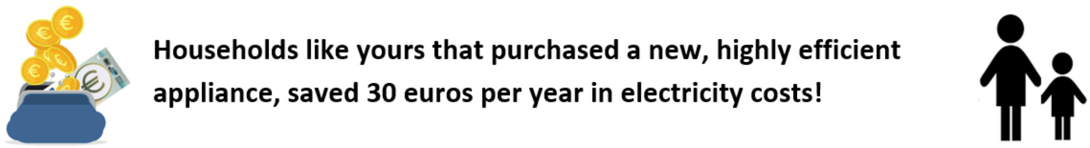
(b) LOSS^{+} Frame.

Figure 3.3: Treatment variations: Gain and Loss Frame.

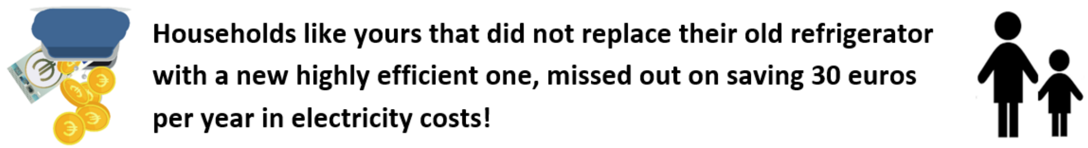
calculate the expected annual savings in individual electricity costs based on a comparison between the current actual electricity consumption levels of the old device and a hypothetical reference device of equivalent size and features that meets the necessary efficiency levels. For realizing these financial gains, it is assumed that households exactly follow the suggested protocol and purchase a model similar to the reference device provided, that they use it in an optimal sense and that there is no change in individual electricity prices. In contrast, in the second set-up, our 'peer experience' treatments display electricity savings based on realized monetary values based on actual replacements by a peer group with a similar household composition. We calculate the annual savings realized by determining the average difference in energy costs between the old appliance and the new appliance actually purchased for the following six household types, representing 89 percent of all households in our sample: single person (26 EUR savings); two adults (29 EUR savings); two adults and one child (30 EUR savings); two adults and two children (36 EUR savings); two adults and three or more children (41 EUR savings); single parenting (30 EUR savings).

We highlight the reference to peer-behavior in two ways. First, we add a pictogram of the respective household composition to the right side of the savings information to the letter. Second, we alter the text to read "Households like yours that purchased a new, highly efficient appliance, saved ..." (GAIN^{+} PEER) as opposed to "If you purchase a new, highly efficient appliance, you could save ..." (GAIN^{+}) (see Figure 3.4a). The alterations are similar in the loss frames (see Figure 3.4b).

As a third treatment dimension, we introduce different reminders for eligible households (see Figure 3.5). In a first variant, in accordance with the EU General Data Protection Regulation (GDPR), households are asked for their written consent to be recontacted 4-8 weeks after having received the information letter. Conditional on con-



(a) GAIN$^+$ PEER Frame.



(b) LOSS$^+$ PEER Frame.

Figure 3.4: Treatment variations: Peer experience Gain and Loss Frame.

sent[9] and depending on households' stated preferences, local site managers then send out a letter- or SMS-based reminder at the beginning of a new month (when households usually are more financially liquid; see Figure 3.5a for a translation of the reminder text). In a second variant, the energy advisor places a tag, which displays the logo of the Energy-Saving-Check program, inside the refrigerator during the second visit (see Figure 3.5b).[10]

We organize these three treatment dimensions in a "2x3 + Reminder" design, as displayed in Figure 3.6. The columns of the table capture the first treatment dimension, the variation in the framing. We distinguish between (1) the legacy GAIN frame, (2) the visually enhanced management's baseline frame, i.e., GAIN$^+$, and (3) the visually enhanced LOSS$^+$ frame. The rows of the table in Figure 3.6 display the second treatment dimension, i.e., the savings based on individual estimates or peer experience. While the legacy treatment is only combined with the individual savings estimate, we vary for both the GAIN$^+$ and the LOSS$^+$ frame whether they are combined with the individual savings estimate or the peer-experienced savings. Finally, we combine selected treatments with the third dimension, the reminders. That is, the orange fields in Figure 3.6 display treatment versions that we test both with and without reminders.[11]

[9]To our knowledge, every household provided consent.

[10]During the co-design process, we discussed whether the tag should be placed inside or outside the refrigerator. The program officials and site managers insisted on placing the tag inside the refrigerator as an outside-placement may give rise to stigma. Persons visiting the audited household would directly see the logo of the program directed to low-income households.

[11]Note, that this definition is agnostic with respect to the specific reminder types. Such differentiation between reminder types is only relevant for the GAIN$^+$ treatment, which we combined with (i) the SMS/letter reminder, (ii) the refrigerator tag and (iii) both reminder types. In Figure 3.6 and the regression analysis, we pool these three groups into one GAIN$^+$ REMINDER treatment. By contrast, both reminder treatments in the loss domain of Figure 3.6 are combined with the refrigerator tag, but abbreviated in our specifications as LOSS$^+$ REMINDER and LOSS$^+$ PEER REMINDER treatment, respectively. In robustness checks, we differentiate between the reminder versions.

Reminder for the refrigerator exchange in Stromspar-Check Aktiv

Dear Mrs./Mr./Family xxx,

we would like to remind you that you can request a voucher of 100€ for the purchase of a refrigerator of energy efficiency label class A, B, C or D as part of the Stromspar-Check. To do so, please contact your location.

With kind regards

Your Stromspar-Check Team

(a) SMS or letter reminder.

(b) Refrigerator tag.

Figure 3.5: Reminder treatments.

	Gain		Loss
Individual estimates	GAIN (legacy)	GAIN+ (baseline)	LOSS+
Peer experience		GAIN+ PEER	LOSS+ PEER

+

REMINDER

Figure 3.6: Treatment dimensions

In sum, we implemented a total of eight different treatments: (1) GAIN (legacy version), (2) GAIN$^+$ (management baseline), (3) LOSS$^+$, (4) GAIN$^+$ PEER, (5) LOSS$^+$ PEER, (6) GAIN$^+$ REMINDER, (7) LOSS$^+$ REMINDER, (8) LOSS$^+$ PEER REMINDER. Due to technical issues with the database, the legacy GAIN treatment could only be implemented starting at the end of July 2022. To meet our scalability targets, the randomization process was carried out automatically by the program database. Households are allocated to the different treatments with equal probabilities, except for (6) GAIN$^+$ REMINDER to which we over-sampled.[12]

With our experimental design, we test four main hypotheses. Our first hypothesis concerns the impact of using visual enhancement or visual aids in an information letter about energy efficiency in order to reinforce the verbal message (Stojanovski et al., 2020; Allcott and Greenstone, 2017). In our case, the visual enhancement consists of adding the purse pictogram to the letters, to emphasize the monetary consequences of program participation. We expect that such enhancements will be conducive to increasing attention to the message of the letter, in particular the reference to the euro-denominated savings in the message body.

H1: Refrigerator replacement rates will be higher in GAIN$^+$ compared to GAIN .

The second hypothesis concerns possible performance differences between the GAIN$^+$ frame and an alternative LOSS$^+$ frame. Such message framing, in particular the framing of the same material outcome as a gain or a loss by shifting the mental reference point, has received attention in the literature on enhancing household energy efficiency for some time. A starting point is the seminal study by Gonzales et al. (1988) who examined the effect of exposing 408 home owners who qualify for enrolment in a energy efficiency retrofit program to one of two different Home Energy Audit procedures. In one, home owners were visited by auditors trained to employ a gain framing by referring to the benefits of enrolling; in the other, by auditors trained to employ a loss framing by referring to the foregone benefits of not enrolling. They find significantly higher enrolment among home owners in the loss framing treatment. This finding seems to reflect a higher impact of loss framing on behavior in general (see Kühberger (1998) for a meta-study) and in the specific context of energy savings (see Homar and Cvelbar (2021) for a meta-study covering 61 studies).[13]

[12]The reason is that the GAIN$^+$ REMINDER treatment was originally designed to feature two treatments, one using an SMS reminder, the other a mailed letter reminder. Feedback during the pilot phase led to the decision to combine the two treatments into one, with the household choosing how the reminder would be provided. In addition, as discussed above, treatment group (6) includes participants who are reminded by the SMS/letter, participants who are reminded by the tag and participants who are reminded by both the tag and the SMS/letter.

[13]Recent studies by DeGolia et al. (2019); Park et al. (2023) point in the same direction. The finding

The cumulative evidence that favors loss over gain framing in the context of energy efficient behavior leads us to hypothesize that the visually enhanced loss frame will give rise to a higher probability that households will replace their refrigerators than the visually enhanced gain frame. Both treatments increase the salience of estimated savings via the added pictogram, but in line with much of the literature, we expect the loss frame to provoke stronger behavioral reactions.

H2: Refrigerator replacement rates will be higher in the LOSS$^+$ frame compared to the GAIN$^+$ frame.

One threat to H2 comes from the lack of specific evidence on how low-income households respond to different message framings. This is relevant because most of the cumulative evidence on message framing is derived on the basis of the average household, with little discussion on its "transportability", i.e., on whether the same patterns also apply to our specific low-income demographic. Their behavioral patterns could differ.[14]

Thirdly, we hypothesize a reference to peer experiences to both reduce uncertainty on the actual electricity savings from replacement and to increase the personal relevance of the program. First, when expected electricity savings are calculated based on individual projected estimates, it may not be entirely clear to the household what "*can* save you" (as indicated in the information letter) exactly means in this context. By contrast, the peer experience design is more specific in that it provides estimates of *realized* savings by similar households. Second, considering the large literature on peer effects in energy behavior (e.g., Allcott 2011, Andor et al. 2020), households may have higher trust in information stemming from households in similar socioeconomic circumstances.[15] Hence, the peer experience may decrease the degree of uncertainty and increase the attachment households place on the possible savings.

H3: Refrigerator replacement rates will be higher in the peer experience design compared to a design relying on individual projected estimates.

The fourth hypothesis focuses on the effect of reminders as a tool to overcome poten-

is not universal, however: Some studies also find no (e.g., Sussman et al., 2018) or the opposite (e.g., Chen, 2023) effect.

[14]For example, Mullainathan and Shafir (2013) show scepticism towards the effectiveness of loss frames among low-income individuals. Relatedly, Fehr et al. (2022) show that financial scarcity decreases the likelihood to exhibit an endowment effect, a mainstay of behavioral biases among experimental subjects.

[15]Likewise, also research in other domains provides evidence of social learning from peer behavior, e.g., Escobar and Pedraza (2023) in the context of stock trading and Abdulai (2023) in the context of agricultural farming technologies.

tial procrastination of, and inattention to, the replacement choice. There is previous evidence from the RRP that inattention and procrastination may have a role to play in explaining low levels of replacement among eligible households. Specifically, Chlond et al. (2025) find that inadvertent program changes in the RRP that also involved having to set households deadlines for replacing their refrigerators led to increased take-up. This is a pattern that would be consistent with households suffering from procrastination, which can be overcome by deadlines as a form of goal setting.

Reminders are a popular strategy for overcoming procrastination and inattention among targeted individuals. This popularity can be explained by the effectiveness of reminders across a wide range of behavioral contexts (see Gravert (2021) for a recent review), with even small reminders frequently resulting in sizeable effects. This also holds for the context of energy efficiency (Fang et al., 2023) and of low-income households (Karlan et al., 2016; Guyton et al., 2016). The behavioral economics of reminders emphasizes limited attention (Karlan et al., 2016) and the interplay of present bias and limited memory (Ericson, 2017) as drivers of why households pay insufficient attention to the future benefits of an action (here, the future saving from replacing the appliance), overemphasize the current cost (here, the outlay for replacing the appliance), leading them to postpone – and ultimately forego – an otherwise beneficial investment. Reminders intervene in this process by lowering the cost of attention and/or overcoming limited memory. On this basis, we predict that the effect of reminders on replacement behavior is positive.

H4: Refrigerator replacement rates will be higher in the reminder treatments compared to the non-reminder treatments.

Threats to H4 come from two different areas. One area is the fact that in our experiments, reminders are always combined with a gain or a loss framing. The reminder effect may therefore depend on the initial frame presented in the information letter, and thereby on the first-order effects of this framing (see H2). At least two issues arise as a result. One, the frame could affect the composition of the sample being recontacted for the reminder.[16] Two, the frame could determine the direction of the reminder effect, reinforcing or possibly reversing the effect.[17]

The other area of threat is the procedural requirement of compliance with the EU

[16]For example, if H2 holds and low-income households have higher replacement rates under a loss than under a gain frame, then the sample that has not redeemed the voucher yet and, hence, receives the reminder will not be the same across the two framings.

[17]For example, if H2 does not hold and households are discouraged through a loss frame, then a household can make sure that it will not be exposed to a loss frame again by replacing the refrigerator. Therefore, the added reminder may reverse the framing effect and increase the incentives for replacement.

GDPR in programs such as the RRP, ruling out collecting and using contact details for re-contacting household without legitimate cause and prior informed consent. Prior informed consent to being reminded could lead to an 'anticipation' effect that could exacerbate rather than mitigate procrastination (Ericson, 2017). The reason is that anticipating future reminders, the household will find it in its interest to allocate even less costly mental effort to acting on the decision situation than when not anticipating being reminded. This anticipation effect may indeed be negative and contribute to lower replacement rates in those reminder treatments that ask for permission to receiving a letter or SMS message. The treatment that uses a tag, on the other hand, would be expected to be unaffected from such anticipation effect since the tag is present from the beginning and for as long as the household leaves it in place.

3.3.3 Selection and training of sites

To address potential concerns of site selection and to support the scaling idea of our interventions, we randomly selected a number of local sites being invited to participate in our field experiment. To this end, we collected data on the number of long-term unemployment benefit recipients of the 158 municipalities in which an SSC field office is located. We determined weights that reflect the share of the number of benefit recipients in the service area of a site relative to the total number of benefit recipients in Germany. Higher weights were consequently assigned to sites which cover a greater number of benefit recipients. This weighting ensures that each individual benefit recipient has the same probability of being part of our study. The selected sites were then determined by a weighted randomized draw of a pre-arranged number of 30 out of 150 total sites.

The randomly drawn sites received invitations by the program officials to participate in our study in December 2021. 23 out of the 30 sites followed our invitation.[18] One site (Munich) had to be excluded from participation due to a large-scale investment program tested in parallel there. A second site (Groß-Gerau-Kreis) experienced an unexpected lack of site management, which is why this site was as well excluded from the study. Three additional sites (Jena, Weimar, Erfurt) as well as the pilot site (Frankfurt) showed interest and self-selected themselves into the experiment. To avoid resulting selection effects, we exclude these sites from our analysis.[19] Hence, our final sample consists of the following 21 sites: Anklam, six sites in Berlin, Bremen, Gelsenkirchen, Hamburg, Ibbenbüren, Cologne, Konstanz, Leipzig, Meißen, Mettmann (county), Minden, Offenburg, Osnabrück (county), Saarlouis, Wiesbaden. Figure 3.7

[18]The sites not signing up were Chemnitz, Delmenhorst, Essen, Recklinghausen, Moers and two sites in Dortmund.

[19]Results are however robust to including them.

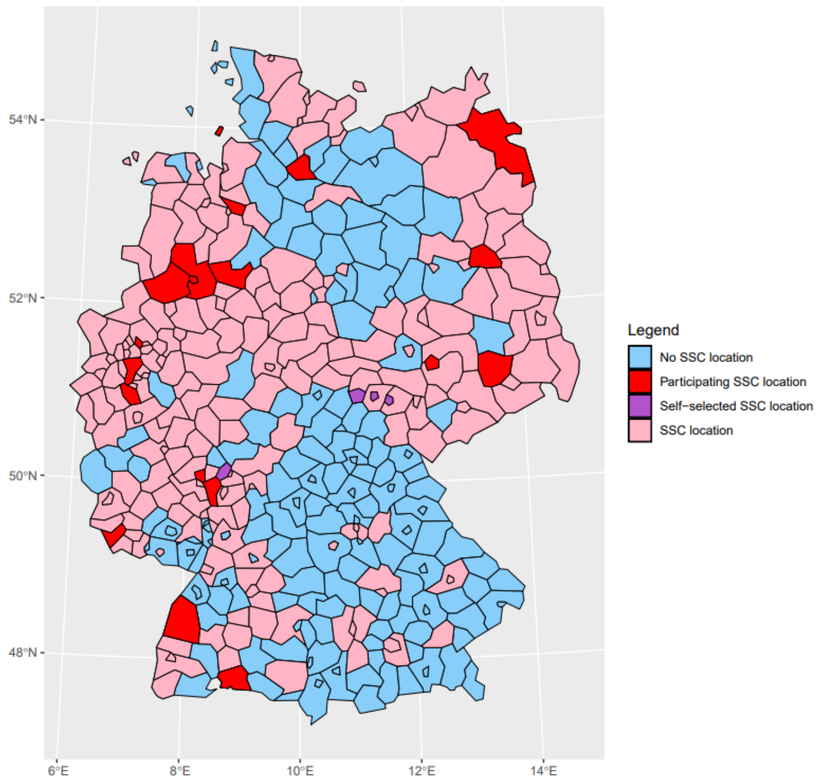


Figure 3.7: Program sites of the "Energy Saving Check".

displays all program sites (pink), the sites participating in our study (red) and the excluded self-selected sites (violet).

After the selection process had been determined, we organized various training sessions to familiarize the site managers with the treatments and their implementation. This was necessary as all treatments, except for the legacy GAIN treatment, were new to the site managers participating in the RCT. The managers of the selected sites are distinct from the site managers participating in the co-design process, and were not informed about the origins and discussions underlying the different treatments (including management's request for the baseline GAIN$^+$ frame). As a further important aspect of scalability, the program database was adapted to assign eligible households to the different treatments and automatically print out the correct version of the information letter. Thus, there is no leeway for site manager choices.

However, the site managers are responsible for sending the SMS and letter reminders, and for equipping the auditors with the refrigerator tag. For this reason, we prepared an online tutorial, Q&A sessions and information material for the site managers. The Q&A sessions were held from December 2021 to January 2022. From January 2022, the database programming was finished to include all randomized treatments and all study sites received treatment materials. The period between January 2022 and April 2022 thus served as an additional trial phase for site managers to inform, organize and

Table 3.1: Summary statistics of household data

	Mean	Std. Dev.	Min	p10	p25	p50	p75	p90	Max	N
Electricity price (Euro/kWh)	0.33	0.05	0	0.30	0.30	0.32	0.34	0.40	1	1,803
Energy consumption (kWh)	2926.10	1708.17	0	1250.00	1712.00	2572.33	3700.00	4957.83	17,501	1,762
Electric water heating: (1=yes)	0.37	0.48	0	0.00	0.00	0.00	1.00	1.00	1	1,803
Annual est. savings (kWh)	336.33	153.37	12	204.67	228.68	293.64	414.62	547.56	1,466	1,802
No. persons (count)	2.78	1.88	1	1.00	1.00	2.00	4.00	5.00	11	1,803
Living space (m^2)	71.08	27.60	20	43.00	52.80	65.00	83.00	107.00	300	1,803
Unemployment benefits (1=yes)	0.65	0.48	0	0.00	0.00	1.00	1.00	1.00	1	1,803

Note: Displayed are summary statistics for the variables electricity price (in Euros/kWh), household energy consumption (in kWh), an indicator for heating warm water with electricity, engineering estimates of savings from refrigerator replacement (in kWh), number of persons in household, living space, and an indicator for receiving long-term unemployment benefits.

train their local staff to the interventions.

3.4 Sample description

Our study sample consists of program participants, who (i) received their audit during the intervention period from April 1, 2022 to February 15, 2023,[20] (ii) were audited within one of our randomly selected intervention sites, and (iii) were found eligible for the refrigerator replacement program. In total, we observe 1,803 households that fulfill these criteria.

Table 3.1 displays summary statistics of household information being recorded in the program database after the first household visit. The average electricity price in the billing period 2021/2022 was 33 cents/kWh, which aligns with the average electricity prices paid nation-wide (Statistisches Bundesamt (Destatis), 2023b). Electricity consumption in our sample is notably lower compared to national statistics. While the participants of the Energy-Saving-Check consume on average 2,926 kWh per year, the German average is at 3,383 kWh per year (Statistisches Bundesamt (Destatis), 2023c). An important determinant of electricity consumption is whether warm water is produced using electricity which applies to 37 percent in the sample. The annual energy savings from refrigerator replacement as estimated by engineering estimates are on average 336.33 kWh. Together with the average electricity price, this maps into annual financial savings of 112.74 Euros on average. The average number of household members in our sample is 2.8 persons – slightly higher than the German average of 2.0 household members (Statistisches Bundesamt (Destatis), 2023a). Despite larger household sizes, the living space of our sample (71.1m^2 on average) is lower compared to the national average (96.2m^2, Statistisches Bundesamt (Destatis) (2023d)).

These differences to national statistics are highly plausible given that all of our study participants are recipients of federal income support. The majority of 65 percent are

[20]The program's funding period ended on March 31, 2023, but the local sites stopped voucher issuance on February 15, 2023, due to accounting procedures.

recipients of long-term unemployment benefits (Arbeitslosengeld II), the second largest fraction of 16 percent receive a basic pension (Grundsicherung) and 12 percent receive housing benefits (Wohngeld).

We next turn to comparing the summary statistics across treatment groups. Table 3.2 displays the mean difference in characteristics between the respective treatment and (the mean of) all other groups. Due to randomization, in expectation we should see no major differences in observable characteristics between treatments. Importantly, we compare the individual estimated energy savings in kWh, as measured during the first home visit, and not the financial savings as communicated to the households. We would thus not expect differences in the estimated kWh savings across treatment groups.

The number of observations per treatment is about equally split across groups except for two exemptions. First, the number of observations for the GAIN group is lower due to technical issues causing a delay in implementation (see Section 3.3.2). Data collection for this group only started by the end of July 2022. Second, the number of observations is higher for the GAIN$^+$ REMINDER group, since we summarize three initially implemented treatment groups that vary in the reminder version and we over-sampled the GAIN$^+$ SMS/Letter REMINDER group as explained in Section 3.3.2.

Comparing the means of the respective covariates across groups, we see slight differences in the electricity price paid and the estimated kWh savings from refrigerator replacement. As one example, subjects in the GAIN treatment on average pay 0.17 ct more per kWh and their expected savings from replacement are on average 37.36 kWh higher compared to the average household in all other treatments (Table 3.1, column 1). These differences become insignificant or only marginally significant once adjusting for multiple hypothesis testing (List et al., 2019). We also add controls for these variables in our robustness checks.

Figure 3.8 displays the mean and 95 percent confidence intervals of our main outcome, refrigerator replacement, by treatment group. Replacement rates, i.e. the share of households who were informed about the subsidized replacement opportunity during the second visit and actually replaced their refrigerator, range between 14 and 24 percent. We observe the highest replacement rate among households randomly placed in the baseline treatment GAIN$^+$. The LOSS$^+$ treatment has the lowest replace rate. Between these two endpoints are the treatments using peer experience to encourage take-up, repeating the pattern of gain versus loss framing, as well as the legacy design GAIN. As Figure 3.8b shows, replacement rates do not substantially improve by combining changes in the framing of the information letter with reminders. In fact, reminders could negatively affect take-up: The replacement rate among households in the GAIN$^+$ REMINDER treatment is 17 percent, below the replacement rate of 24

Table 3.2: Differences in summary statistics by treatment group

Variable	GAIN	GAIN+	LOSS+	GAIN+ PEER	LOSS+ PEER	GAIN+ REMINDER	LOSS+ REMINDER	LOSS+ PEER REMINDER
Energy price (ct/kWh)	0.017***	-0.008*	-0.004	0.008**	-0.005	-0.001	-0.007	0.007
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Energy consumption (kWh)	-52.260	219.960*	-60.091	-126.032	-84.370	77.374	-92.195	-79.227
	(169.694)	(132.055)	(142.873)	(137.835)	(145.405)	(83.596)	(143.276)	(144.543)
Electric water heating: (1=yes)	0.049	-0.022	0.025	0.056	-0.021	-0.030	-0.032	0.046
	(0.048)	(0.037)	(0.040)	(0.038)	(0.041)	(0.023)	(0.040)	(0.041)
Estimated savings (kWh)	37.361**	-0.413	-0.397	0.310	16.051	-18.097**	1.631	10.618
	(15.134)	(11.767)	(12.742)	(12.143)	(12.884)	(7.401)	(12.815)	(12.925)
No. persons (count)	-0.157	0.151	-0.001	-0.101	0.127	-0.009	0.017	-0.072
	(0.186)	(0.144)	(0.156)	(0.149)	(0.158)	(0.091)	(0.157)	(0.158)
Living space (m^2)	-2.150	2.861	0.768	-3.001	-3.551	1.723	-0.925	-0.009
	(2.727)	(2.116)	(2.292)	(2.184)	(2.318)	(1.333)	(2.306)	(2.326)
Transfer scheme (1-7)	0.382	-0.177	0.182	-0.129	-0.244	0.036	0.132	-0.101
	(0.314)	(0.244)	(0.264)	(0.252)	(0.267)	(0.154)	(0.266)	(0.268)
N per group	109	190	159	177	155	702	157	154
N total	1,803	1,803	1,803	1,803	1,803	1,803	1,803	1,803

Note: Displayed are the differences in means for the variables energy price, energy consumption, an indicator for heating warm water with electricity, estimated savings in kWh from refrigerator replacement, number of persons in household, living space, an indicator for receiving unemployment benefits and transfer type categories by treatment group compared to all other groups. Standard errors are in parenthesis. Significance levels: $^{*}: p < 0.10$, $^{**}: p < 0.05$, $^{***}: p < 0.01$.

percent in the GAIN+ treatment. The heterogeneity of replacement rates across treatments is already visible at the voucher request stage, which must precede replacement. The voucher request rates range between 26 and 37 percent.[21] The request rate is highest in the GAIN+ treatment and lowest in the LOSS+ treatment.

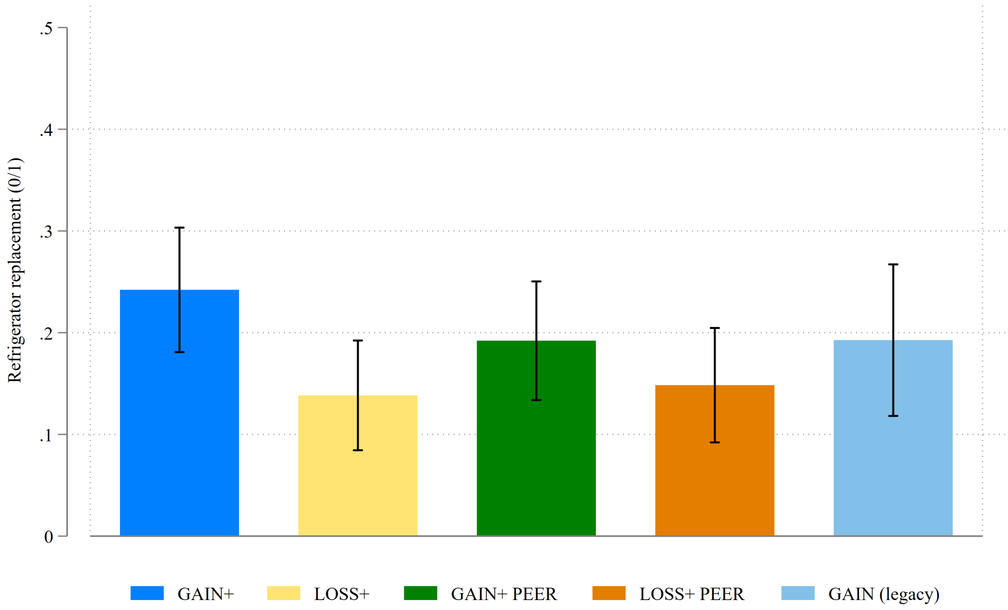
In numbers, we observe that 46 out of 190 possible refrigerators are actually replaced in the GAIN+ treatment. In the Loss treatment, only 22 out of 159 possible refrigerators are replaced. A first statistical comparison already indicates that the replacement rate of the GAIN+ treatment is significantly higher compared to the average replacement rate of all other treatments (two-sided t-test, $p = 0.0203$). We investigate these surprising patterns closer in the next sections.

3.5 Empirical strategy

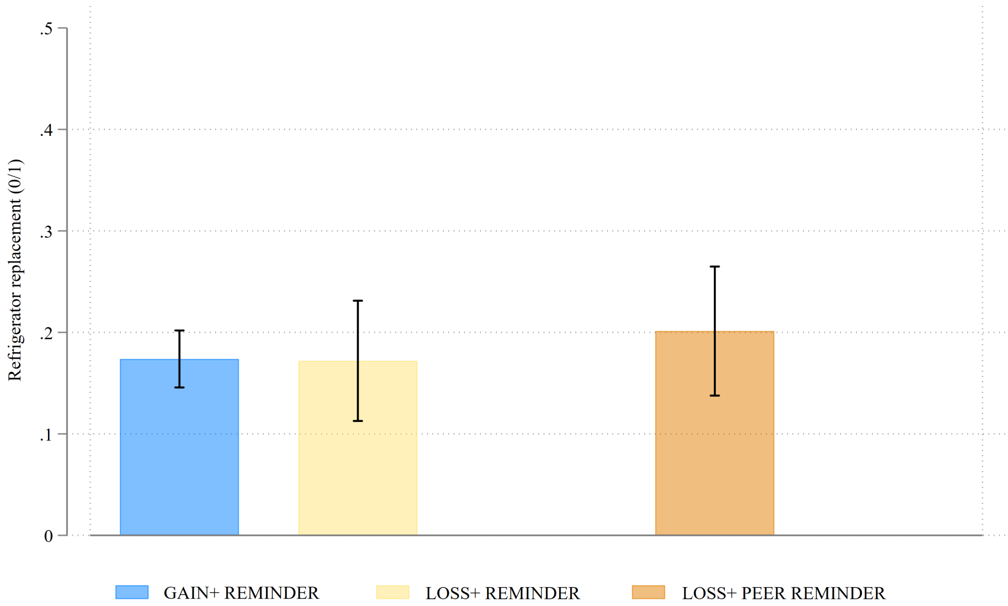
To analyze the effect of the different behavioral interventions on RRP performance, we compare our two outcomes (voucher request and refrigerator replacement) across the treatment groups. That is, we first assign an indicator variable for each treatment group and then run a regression of household (i) voucher request and (ii) refrigerator replacement choice on the treatment indicators as follows:

$$\begin{aligned} Y_i =& \beta_0 + \beta_1 GAIN_i + \beta_2 LOSS^+_i + \beta_3 GAIN^+ PEER_i + \beta_4 LOSS^+ PEER_i \\ &+ \beta_5 GAIN^+ REMINDER_i + \beta_6 LOSS^+ REMINDER_i + \beta_7 LOSS^+ PEER\ REMINDER_i \\ &+ Savings\ Info_i + \mathbf{X}_i + \mathbf{F}_i + \epsilon_i, \end{aligned} \tag{3.1}$$

[21] Figure A3.1 in the Appendix displays the corresponding voucher request rates.



(a) Information letter treatments.



(b) Reminder treatments.

Figure 3.8: Refrigerator replacement rates by treatment group.

whereby i denotes the individual household and Y_i refers to voucher request or refrigerator replacement choice. The treatment indicators equal 1 if the household is in the respective treatment group and 0 otherwise. We estimate equation (1) as a linear probability model (LPM).

We define the GAIN$^+$ treatment as the omitted treatment category as the program management designated this treatment to be the new program baseline. Moreover, using the GAIN$^+$ treatment as baseline provides greater statistical power compared to using legacy program baseline, the GAIN treatment. Due to the delay in implementing the GAIN treatment, the group has the fewest observations. Using the GAIN treatment as baseline would thus penalize the power of all comparisons. We pool the different reminder versions that are combined with the GAIN$^+$-frame (GAIN$^+$ REMINDER) to increase power for analysis. However, in an additional robustness check, we explore potential differences between the tag and the SMS/Letter reminders.

In all regressions, we control for the expected financial savings from replacement as communicated to the household, denoted by *Savings Info*. Importantly, this variable adjusts for differences in communicated savings that might otherwise bias the comparison between the individual estimates and peer experience treatments. In controlling for *Savings Info*, we allow the different framing to impact replacement choices, but hold constant the displayed monetary values.

Further, in subsequent specifications, we add control variables obtained from the program database. Specifically, the vector $\mathbf{X}_i$ summarizes the household's electricity price, electricity consumption, usage of electric water heating, the number of persons in the household, the living space, the type of social benefit transfer the household receives and the federal state the household lives in. Further, the vector $\mathbf{F}_i$ summarizes fixed effects for the local site, the two energy advisors auditing the household, month fixed effects and month-site fixed effects.[22] Finally, ϵ_i denotes the error term.

3.6 Results

We first discuss the results on the treatments that modify the information letter design in Section 3.6.1. Results on the reminder treatments are discussed in Section 3.6.2. Section 3.6.3 investigates heterogeneous treatment effect by the type of federal income support received.

[22]We include month-site fixed effects since some sites have complementary programs that increase the voucher value and vary over time. E.g., the Berlin sites introduced additional cash vouchers in November 2022. These complementary programs exist in at least four of the sixteen states and in a number of municipalities.

3.6.1 Information letter treatments

Table 3.3 displays the regression results for voucher request across the different treatment groups and Table 3.4 shows the corresponding results for refrigerator replacement. In both tables, the GAIN$^+$ treatment is the omitted baseline group.

Table 3.3 shows that both the legacy GAIN and the LOSS$^+$ frame lead to lower voucher request rates compared to the management's baseline GAIN$^+$ treatment. In column (1), these reductions are significant at the 10- and 5-percent level, respectively. Further, across the different models the lower request rate in the LOSS$^+$ treatment remains robust and significant at the 5-percent level. The impact of the frame on behavior is substantial: Framing the reported annual savings in electricity costs as a financial loss hampers the conversion rate of an information letter into a voucher by 10.5 to 11.5 percentage points. The exception is column (4), in which the LOSS$^+$ coefficient reduces to 5.4 percentage points and turns insignificant. Column (4) includes advisor fixed effects, and thus controls for the specific pair of advisors that visited the household. However, once including the advisor fixed effects, we loose variation in our outcome variable, which likely explains the loss of significance of some treatments. We thus view specification (5) as our preferred specification. It includes control variables as well as site, month and month-site fixed effects, but disregards the advisor fixed effects. Figure 3.9a graphically summarizes the estimates reported in column (5).

For the PEER-treatments, we find mainly positive coefficients, suggesting larger voucher request rates compared to the GAIN$^+$ treatment. Yet, these coefficients are not significant. Specifically, for the GAIN$^+$ PEER treatment, the coefficients in column (1)-(4) suggest 0.2 to 5.6 percentage points higher request rates but are not statistically significantly different from zero. In our preferred specification (5), the coefficient even turns negative but remains small and indistinguishable from zero. The coefficient of the LOSS$^+$ PEER treatment in specification (5) suggests a 1.4 percentage points higher request rate compared to the GAIN$^+$ treatment, but is insignificant from zero. In the other specifications, the LOSS$^+$ PEER coefficients range from –0.017 to 0.032 and are, yet again, not statistically different from zero.

As displayed in Table 3.4, for our main outcome of interest, refrigerator replacement, all estimated treatment coefficients display a negative sign and, thus, indicate that the GAIN$^+$ baseline yields the highest refrigerator replacement rates. Again, the difference to the LOSS$^+$ treatment is most robust, yielding a statistical significant difference at the 1- or 5-percent level, indicating that replacement rates are 10.2-11.7 percentage points lower as compared to the GAIN$^+$ treatment. As discussed above, the only exception is the specification in column (4). When adding the advisor fixed effects, the negative effect of LOSS$^+$ reduces to 6.2 percentage points and becomes insignif-

Table 3.3: Treatment effects on voucher request

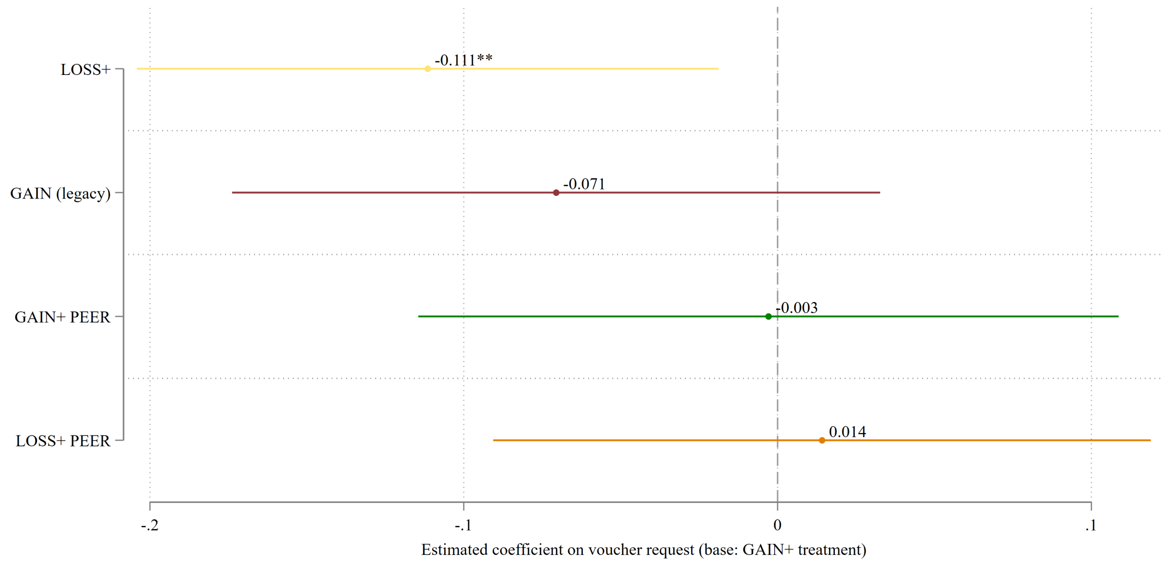
	(1)	(2)	(3)	(4)	(5)
	Voucher request (0/1)				
GAIN$^+$	REF	REF	REF	REF	REF
LOSS$^+$	-0.105**	-0.115**	-0.108**	-0.054	-0.111**
	(0.050)	(0.049)	(0.047)	(0.050)	(0.047)
GAIN$^+$ PEER	0.031	0.056	0.002	0.025	-0.003
	(0.053)	(0.054)	(0.055)	(0.058)	(0.057)
LOSS$^+$ PEER	-0.004	0.032	0.004	-0.017	0.014
	(0.054)	(0.055)	(0.053)	(0.055)	(0.053)
GAIN (legacy)	-0.098*	-0.085	-0.077	-0.040	-0.071
	(0.056)	(0.055)	(0.054)	(0.054)	(0.053)
GAIN$^+$ REMINDER	-0.065*	-0.062	-0.068*	-0.038	-0.059
	(0.039)	(0.039)	(0.037)	(0.037)	(0.036)
LOSS$^+$ REMINDER	-0.044	-0.048	-0.044	-0.029	-0.041
	(0.051)	(0.050)	(0.047)	(0.047)	(0.047)
LOSS$^+$ PEER REMINDER	0.031	0.071	-0.032	-0.017	-0.034
	(0.055)	(0.056)	(0.055)	(0.057)	(0.056)
Constant	0.301***	0.541***	0.309***	0.324***	0.415***
	(0.043)	(0.085)	(0.045)	(0.045)	(0.100)
Savings Info	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	No	Yes
Fixed Effects	No	No	Yes	Yes	Yes
Advisor FE	No	No	No	Yes	No
N	1802	1761	1785	1725	1745

Note: Linear probability models of voucher request (yes/no) on treatments. The treatments are included as indicators for the respective treatment group. The GAIN$^+$ treatment is the omitted reference treatment group. All regressions control for the communicated savings from replacement. Columns (2) and (5) add control variables for household's electricity price, number of persons in the household, past electricity consumption, living space, federal state, the social benefit transfer scheme, and whether the household heats warm water with electricity. Columns (3)-(5) include fixed effects for the respective intervention sites, the month of when the participant is informed about his/her replacement eligibility and a month-site interaction. Column (4) controls for the two advisors visiting the household. Robust standard errors are in parenthesis. Significance levels: $^{*}: p < 0.10$, $^{**}: p < 0.05$, $^{***}: p < 0.01$.

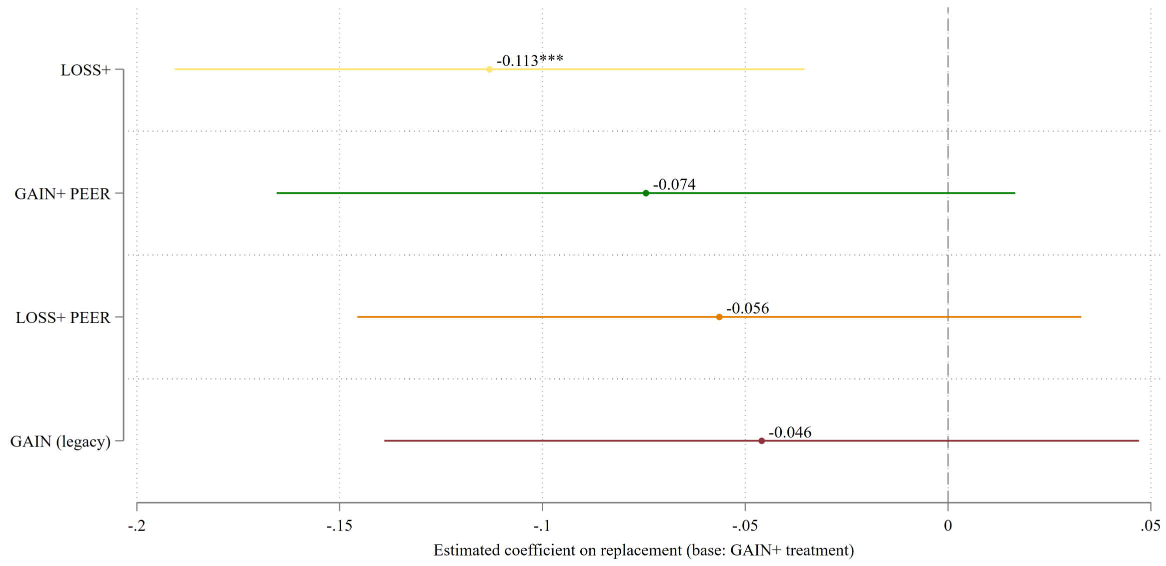
Table 3.4: Treatment effects on refrigerator replacement

	(1)	(2)	(3)	(4)	(5)
	Refrigerator replacement (0/1)				
GAIN$^+$	REF	REF	REF	REF	REF
LOSS$^+$	-0.104**	-0.117***	-0.102**	-0.062	-0.113***
	(0.042)	(0.041)	(0.040)	(0.044)	(0.040)
GAIN$^+$ PEER	-0.044	-0.045	-0.061	-0.037	-0.074
	(0.046)	(0.046)	(0.046)	(0.049)	(0.046)
LOSS$^+$ PEER	-0.088*	-0.065	-0.071	-0.065	-0.056
	(0.045)	(0.045)	(0.046)	(0.048)	(0.046)
GAIN (legacy)	-0.051	-0.051	-0.046	-0.025	-0.046
	(0.049)	(0.048)	(0.048)	(0.052)	(0.047)
GAIN$^+$ REMINDER	-0.069**	-0.072**	-0.075**	-0.051	-0.070**
	(0.034)	(0.034)	(0.032)	(0.035)	(0.032)
LOSS$^+$ REMINDER	-0.070	-0.078*	-0.082**	-0.066	-0.085**
	(0.043)	(0.044)	(0.042)	(0.044)	(0.042)
LOSS$^+$ PEER REMINDER	-0.035	-0.020	-0.090*	-0.070	-0.098**
	(0.047)	(0.048)	(0.046)	(0.049)	(0.047)
Constant	0.234***	0.292***	0.233***	0.226***	0.238**
	(0.037)	(0.072)	(0.039)	(0.040)	(0.096)
Savings Info	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	No	Yes
Fixed Effects	No	No	Yes	Yes	Yes
Advisor FE	No	No	No	Yes	No
N	1802	1761	1785	1725	1745

Note: Linear probability models of refrigerator replacement (yes/no) on treatments. The treatments are included as indicators for the respective treatment group. The GAIN$^+$ treatment is the omitted reference treatment group. All regressions control for the communicated savings from replacement. Columns (2) and (5) add control variables for household's electricity price, number of persons in the household, past electricity consumption, living space, federal state, the social benefit transfer scheme, and whether the household heats warm water with electricity. Columns (3)-(5) include fixed effects for the respective intervention sites, the month of when the participant is informed about his/her replacement eligibility and a month-site interaction. Column (4) controls for the two advisors visiting the household. Robust standard errors are in parenthesis. Significance levels: $^{*}: p < 0.10$, $^{**}: p < 0.05$, $^{***}: p < 0.01$.



(a) Voucher request.



(b) Refrigerator replacement.

Note: Displayed are the estimated coefficients of column (5) of Table 3.3 and column (5) of Table 3.4. The coefficients are sorted by effect size.

Figure 3.9: Coefficient plot of information letter treatments.

icant. We suspect a substantial decrease in the variation of replacement rates across treatments once conditioning on advisor pairs as explanation for this significance loss. We thus view column (5) as our preferred specification. Figure 3.9b visualizes the coefficients reported in column (5) in Table 3.4.

With respect to our hypotheses, we do not find robust evidence supporting H1. While replacement rates are 4.6-5.1 percentage points higher in the management's newly introduced GAIN$^+$ frame compared to the legacy GAIN frame, this difference is not statistically significant. Moreover, our findings clearly oppose our second hypothesis H2. The LOSS$^+$ frame significantly reduces replacement compared to a GAIN$^+$ frame. Here, our findings thus contrast with evidence on the effectiveness of loss frames (Laibson and List, 2015), but support recent literature raising doubt in the universal belief in loss frames (Gal and Rucker, 2018).

The comparison of Table 3.3 and Table 3.4 reveals a nuanced picture of providing information of peer experiences: While the mainly positive coefficients of the PEER treatments suggest an increase of voucher request rates relative to the GAIN$^+$ baseline, the peer experience seems to decrease refrigerator replacement rates. The coefficients of both the GAIN$^+$ PEER and the LOSS$^+$ PEER treatments are negative in Table 3.4. For the LOSS$^+$ PEER treatment, the negative coefficient is even significant at the 10-percent level in the specification (1) of Table 3.4.

Table A3.1 in the Appendix explores this linkage and displays the results of a regression of the replacement probabilities conditional on requesting the voucher. The coefficients of the GAIN$^+$ PEER and the LOSS$^+$ PEER treatments are significant, large and negative. This implies that although the peer experience tends to encourage households to request the voucher – potentially because they are more optimistic regarding their own replacement choice – these households finally fail to successfully conduct the replacement. In other words, peer experience information may increase the rate of requested vouchers, but it does not lead to more energy efficient appliances. From a program perspective, such peer information effect may be rather undesirable: More vouchers are in circulation, which increases the administrative effort and liabilities on the program's balance sheet, but no replacements follow.

In summary, the results displayed in Table 3.4 fail to confirm H3. The peer experience information in the GAIN$^+$ PEER treatment does not increase replacements as compared to the individual engineering estimates in the GAIN$^+$ baseline. The coefficients rather point to a negative effect but do not statistically differ from zero. Likewise, post-estimation tests both on column (1) and (5) fail to reject equality of the LOSS$^+$ and LOSS$^+$ PEER treatments ($p = 0.7111$ and $p = 0.1784$, respectively). Hence, despite the size of coefficients points towards the LOSS$^+$ PEER treatment realizing higher replacement rates compared to the LOSS$^+$ treatment, which would be confirmatory of H3, we do not see robust statistical evidence.

Overall, we find the GAIN$^+$ treatment to be the most effective information letter design. It tends to improve upon the simpler GAIN version and substantially improves upon the LOSS$^+$ version. Further, information about successful peer experience cannot foster technology adoption in our sample.

3.6.2 Reminder treatments

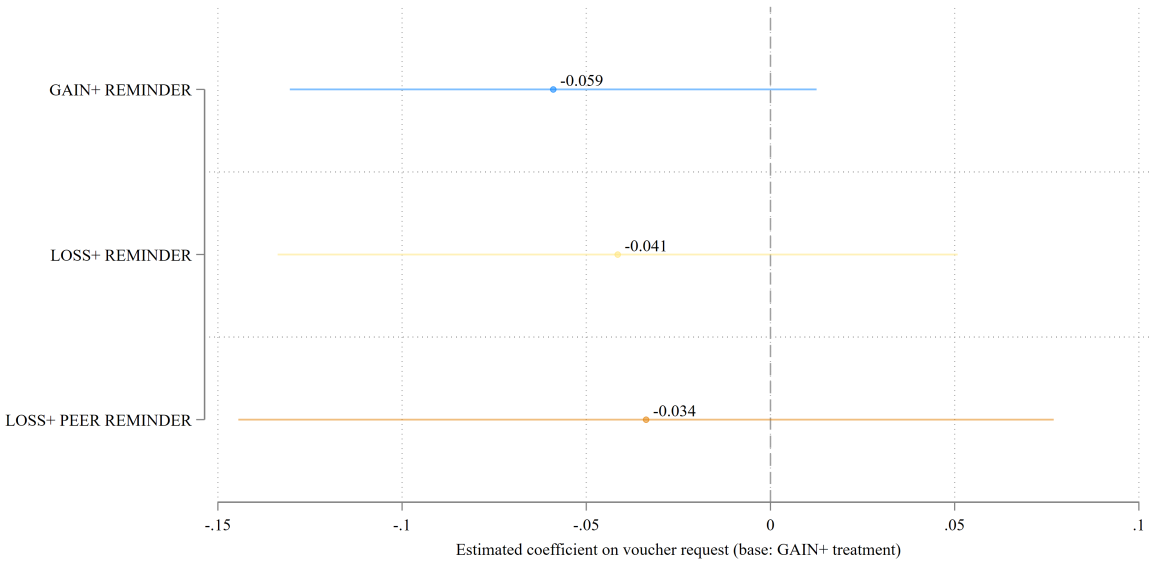
Figure 3.10 displays the estimated coefficients for the reminder treatments of column (5) of Table 3.3 in the upper panel and Table 3.4 in the lower panel. The estimated effects on voucher request are negative but mostly insignificant. More interestingly, for refrigerator replacement, we observe a significant and negative effect for the GAIN$^+$ REMINDER treatment. At the 5-percent significance level, the added reminder decreases energy efficiency investments by 6.9-7.5 percentage points compared to the GAIN$^+$ baseline in all specifications, except for the previously discussed model (4).

This finding is in stark contrast to our hypothesis H4. The reminder decreases the effectiveness of the GAIN$^+$ treatment. A potential explanation for such an empirical finding might be found in the negative anticipation effect as described in Section 3.3.2. The GAIN$^+$ participants may have anticipated receiving a reminder on the replacement choice, thus postponed the decision, allocated less mental effort to remember, and, ultimately, never followed through with the investment choice.

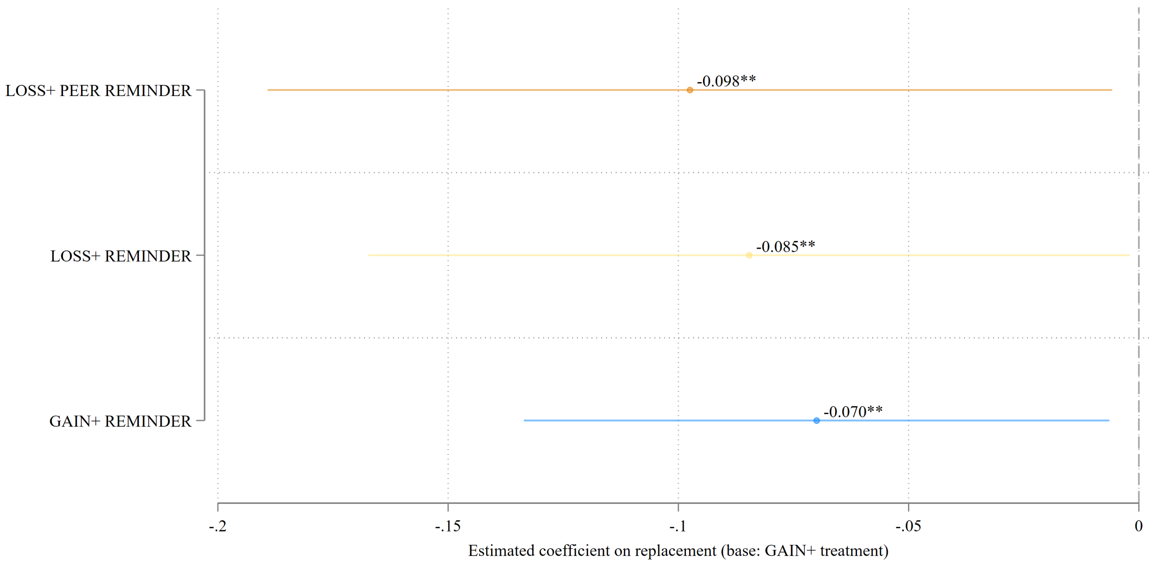
At first sight, this finding is similar for the combination of the loss framing and the reminder. As indicated in Figure 3.10 and in our preferred specification (5), both the LOSS$^+$ REMINDER and the LOSS$^+$ PEER REMINDER coefficient are negative and significant at the 5-percent level. Across our different specifications, the effect of the LOSS$^+$ REMINDER treatment ranges between a 7.0 and 8.5 percentage points decrease in replacement rates compared to the GAIN$^+$ baseline. The LOSS$^+$ PEER REMINDER treatment causes 2.0-9.8 percentage points lower replacement rates. However, the significance of these effects varies and depends on the inclusion of fixed effects.

Further, to evaluate the LOSS$^+$ REMINDER and the LOSS$^+$ PEER REMINDER coefficients with respect to H4, we need to compare them to the LOSS$^+$ and LOSS$^+$ PEER treatments. According to a post-estimation test on column (5) of Table 3.4, the LOSS$^+$ REMINDER treatment does not significantly improve upon the LOSS$^+$ treatment ($p = 0.4833$), and the LOSS$^+$ PEER REMINDER treatment does not significantly improve upon the LOSS$^+$ PEER treatment ($p = 0.3404$).[23]

[23] The respective p-values of post-estimation tests on column (1) of Table 3.4 are $p = 0.4094$ for LOSS$^+$ REMINDER vs. LOSS$^+$ and $p = 0.2205$ for LOSS$^+$ PEER REMINDER vs. LOSS$^+$ PEER.



(a) Voucher request.



(b) Refrigerator replacement.

Note: Displayed are the estimated coefficients of column (5) of Table 3.3 and column (5) of Table 3.4. The coefficients are sorted by effect size.

Figure 3.10: Coefficient plot of reminder treatments.

Hence, we likewise reject H4 for the loss treatments but, contrary to the gain treatments, do not find evidence consistent with negative anticipation effects. Importantly, one explanation for the significantly negative coefficient of the GAIN$^+$ REMINDER treatment and the lack of significant difference between the LOSS$^+$ REMINDER, the LOSS$^+$ PEER REMINDER and their respective non-reminder treatment versions may be the use of different reminder formats. The LOSS$^+$ REMINDER and the LOSS$^+$ PEER REMINDER treatments introduce the fridge tag. This is different for the GAIN$^+$ REMINDER treatment. This treatment group pools three reminder treatments, (1) a GAIN$^+$ Letter/SMS REMINDER, (2) a GAIN$^+$ Tag REMINDER and (3) a GAIN$^+$ Letter/SMS and Tag REMINDER.

Table A3.3 in the Appendix displays the replacement regression results for the different reminder versions.[24] In theses regressions, the GAIN$^+$ Letter/SMS REMINDER and the GAIN$^+$ Letter/SMS Tag REMINDER treatments robustly show significant and negative coefficients. The GAIN$^+$ Letter/SMS REMINDER treatment effect ranges between a 7.2-8.2 percentage point decrease in the likelihood to replace the refrigerator, which is significant at the 5-percent level. For the GAIN$^+$ Letter/SMS Tag REMINDER treatment, the coefficient displays a 8.8-10.4 percentage points decrease at the 1- to 5-percent significance level. By contrast, the GAIN$^+$ Tag REMINDER treatment does not significantly reduce replacement rates compared to the GAIN$^+$ baseline. Similarly, as discussed above, the LOSS$^+$ Tag and LOSS$^+$ PEER Tag REMINDER treatments do not significantly reduce replacement rates compared to the LOSS$^+$ treatment and the LOSS$^+$ PEER treatment, respectively. Hence, only the Letter/SMS reminders significantly reduce replacement rates and cause the rejection of H4. The tag reminders do not significantly affect replacement rates in comparison to their respective information letter-only version.

This finding is consistent with the anticipation effects proposed by Ericson (2017). Only the Letter/SMS reminder was announced to households, by contrast, the tag reminder was directly placed into the household fridge during the second home visit. Thus, only the Letter/SMS reminder might have caused households to anticipate the reminder, postpone the replacement choice and reduce the mental capacity spent on thinking about the replacement choice. Ultimately, this explains the backfiring effect that we observe in our data.

In the next section, we more closely investigate the role of our sample in an exploratory analysis. In particular, we explore the extent to which some of the treatment effects can be attributed to the specific decision-making processes of low-income households.

[24]Table A3.2 in the Appendix displays the corresponding results for voucher request.

3.6.3 Heterogeneous treatment effects

We delve deeper into understanding the role that the vulnerable situation of households in our sample plays in explaining households' (non)reaction to the tested interventions by investigating heterogeneous treatment effects by federal income support type.[25] As described in Section 3.4, 65 percent of our sample receive long-term unemployment benefits (Arbeitslosengeld II, Bürgergeld since January 2023). These unemployment benefits are set to provide a minimum subsistence level only,[26] and, hence, involve the lowest benefit payments we observe in our sample jointly with low pension supplements and benefits for asylum seekers. Households that receive a housing allowance or child supplements earn a small own income that is not sufficient to finance their rent and heating cost and additional expenses for children, respectively.

Sociologists, for example, argue that households receiving long-term unemployment benefits tend to differ in their mindset and attitudes towards society from other groups of welfare recipients. This could lead to a different behavioral response to interventions. In this literature, the difference in mindset and attitudes is traced back to heterogeneity among recipients with respect to the legitimacy of receiving public benefits, as perceived by others and by themselves. These perceptions have turned increasingly negative, in particular for long-term recipients (Dörre, 2015), leading to feelings of discrimination and of a lack of solidarity towards the long-term unemployed (Köster, 2023). This leads to a mindset of feeling left behind, denied access to respectable segments of society, and pitted against other welfare recipients (Dörre, 2015).

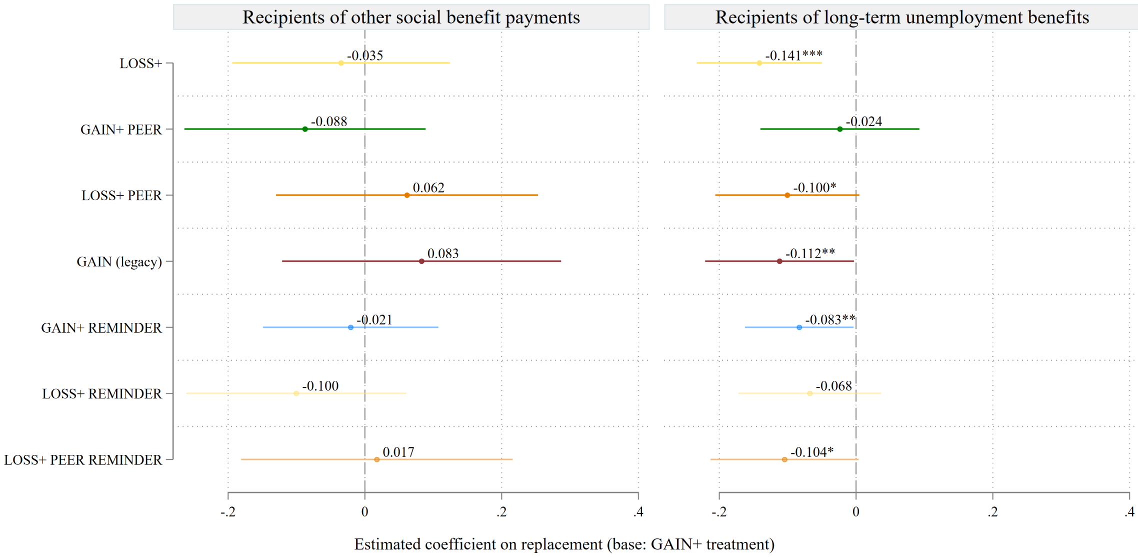
We compare the effectiveness of the treatments on refrigerator replacement rates between ALGII recipients and the recipients of other social benefit payments. To this end, we run regression (1) either among the sample of ALGII-recipients or among the recipients of other social benefit payments. Table A3.4 in the Appendix displays the corresponding regression results. Figure 3.11 presents the estimated coefficients of column (4) for the sample of recipients of other social benefit payments and of column (8) for the sample of recipients of long-term unemployment benefits (ALGII).

Figure 3.11 shows evidence of heterogeneous treatment effects by transfer type. We observe insignificant effects of the treatments for the recipients of other benefit payments than the unemployment benefits on the refrigerator replacement decision.

By contrast, for the recipients of long-term unemployment benefits (ALGII), the GAIN$^+$ baseline significantly outperforms the LOSS$^+$ treatment – a finding that we observed on average but seems to be driven by the program participants who live off long-term

[25]Please note that this analysis was not pre-registered and is exploratory.

[26]In 2022, the long-term unemployment benefits are set to 449 Euros per month for a single-person household.



Note: Displayed are the estimated coefficients of column (4) (left panel) and column (8) (right panel) of Table A3.4 in the Appendix.

Figure 3.11: Coefficient plot of heterogeneous treatment effects on refrigerator replacement rates by transfer type.

unemployment benefits. The LOSS$^+$ coefficient is significant at the 1-percent level, and shows a 14.1 percentage points lower likelihood to replace the refrigerator compared to the management's baseline.

A similar finding holds for the GAIN$^+$ REMINDER treatment. The negative coefficient compared to the GAIN$^+$ treatment, which Table 3.4 displayed on average, is only significant for the sample of ALGII-recipients. More specifically, among that sample, the GAIN$^+$ treatment outperforms the refrigerator replacement rates of the GAIN$^+$ REMINDER treatment at 5-percent significance and suggests 8.3 percentage points higher replacement rates.

Further, in this analysis, the new GAIN$^+$ baseline also significantly outperforms the legacy GAIN treatment. Among the recipients of long-term unemployment benefits, GAIN decreases the likelihood of replacement by 11.2 percentage points, which is significant at the 5-percent level. For this particular sample, we can thus confirm the hypothesis H1.

Overall, the analysis by transfer types reinforces our conclusions from the average treatment effect analysis. The GAIN$^+$ baseline yields investment rates that are significantly higher than for the LOSS$^+$ treatment, particularly for the marginalized group of the long-term unemployed. Prior evidence and prospect theory would have predicted higher replacement rates in the LOSS$^+$ treatment. A potential explanation is that our specific target group interprets the frames differently from the average household because of their experience and their resulting beliefs in their ability to complete an appliance replacement. In this context, the GAIN$^+$ frame entertains a notion of the household succeeding in their planned action, in contrast with much of these households' lived experience, positively affecting beliefs. The LOSS$^+$ frame, on the other hand, reinforces the households' beliefs in their likely failure to complete the replacement. These changes in beliefs and expectations could explain why the loss framing backfires for the target group, in particular for long-term unemployed households. For the latter, the GAIN and the GAIN$^+$ REMINDER treatment also perform significantly worse that the GAIN$^+$ baseline while there is no significant difference for other households. Both the positive effect of a visual enhancement as well as the backfiring of pre-announced reminders appears to be specific to this vulnerable group, highlighting its sociological specificity.

3.7 Conclusion

When trying to design or improve programs targeting specific groups in society, such as economically disadvantaged households, it is tempting for program managers to transport behavioral insights that have proven successful among the general population to their particular application context. This temptation is also present in the context of large-scale energy efficiency assistance programs that are regarded as underperforming, but whose performance has also been shown to respond favorably to small procedural changes. Yet, behavioral insights that have demonstrated their effectiveness for the general population may not generate the intended impacts among the target group, either in terms of magnitude or not even in terms of direction. When the target group is vulnerable or exposed, then the possibility that such transportation of insights has adverse impacts carries additional significance from an ethical perspective.

Partnering with the management to improve the Refrigerator Replacement Program, one of the world's largest energy efficiency assistance programs, we exploited the opportunity to test a set of alternative candidate improvements based on behavioral insights. These improvements needed to conform with the stringent requirements set by managers: Adherence to existing administrative procedures, a track record of effectiveness elsewhere, and zero or negligible cost. To be included among the alternatives were a baseline incorporating visual elements as well as a legacy design. In addition, we constrained ourselves to alternatives and procedures that allow us to fulfill emerging standards of scalability, the so-called SANS conditions. The process of co-designing and piloting resulted in six co-designed treatments, on top of the the baseline and the legacy design, that incorporated behavioral insights. Visual enhancement, loss framing, and using peer-based comparison were chosen to improve the information stage of the program while reminders were intended to improve the post-visit engagement of the target households. The evidence generated under conditions of randomized site selection, absence of attrition, demonstrated naturalness, and scalable interventions affirms that the transportability of behavioral insights to low-income households is under threat. While visual enhancements delivered largely as expected, peer experiences turned out to be ineffective as an improvement. Loss framing in fact back-fired in terms of program performance. Reminders were at best ineffective.

Our results can be seen to have implications for both researchers and policy-makers in the present, the short and the long run. In the present, they point to the need for researchers to disclose more actively the heterogeneity of treatment effects that certain behavioral interventions generate, to acknowledge the absence of evidence on treatment effects for specific groups, and to concede threats to the scalability of treatment effects derived without adherence to conditions such as SANS. Policy-makers for their part need to strive for a better understanding of the evidence that is used to derive average treatment effects and for a better understanding of how this evidence relates to their specific target group.

In the short run, we believe that our results make the case for a more systematic collection of evidence on groups under-represented in studies of behavioral public policy in order to build up the 'true' distribution of treatment effects in the population as a whole. Researchers will need the cooperation from policy-makers to carry out such systematic collection. This requires an openness among policy makers to piloting and – better yet – to conducting controlled experiments yielding internally valid estimates of effects.

If undertaken, such joint efforts by researchers and policy-makers have a chance of resulting in a convergent evidence base that narrows confidence intervals around inferential estimates of transported policies. This chance relies on the premise that certain regularities are likely to arise, which will enhance the inference of effect sizes for target groups as more testing is conducted in the short run. The target groups of political import in many OECD countries tend to be vulnerable sections of society, in particular economically disadvantaged groups. These groups exhibit a high degree of heterogeneity across and within. At the same time, their members also confront similar cognitive and affective challenges. A widening evidence base among behavioral economists on the commonalities in how such members respond to different instruments will put researchers in a better position to predict and advise on promising strategies and for policy-makers to make more informed choices than at present.

Chapter 4

Agents of Change? – Energy advisors and the success of energy efficiency assistance programs

Coauthor(s): Timo Goeschl, Martin Kesternich and Madeline Werthschulte

Abstract: Optimizing the human factor in assistance programs – frontline workers who implement the program on the ground – offers yet unexplored potential to increase the success of assistance programs. In the context of energy efficiency programs, energy advisors provide an important link between the program and its beneficiaries, offering advice tailored to individual households' needs and a personal point of contact. This paper provides first evidence on the characteristics of a successful energy advisor, delivering a comprehensive study of advisor determinants on households' investment decisions. Matching survey data (n = 113) collected from energy advisors that staff a nation-wide energy efficiency assistance program in Germany with data from energy audits in the homes of around 6,800 low-income households, we find significant heterogeneity in the performance between advisors to encourage energy efficiency investments in households. Advisors' preferences, attitudes and own investment choices impact households' investment decisions. While advisor demographic information does not directly impact decisions, it matters for household investment via peer effects. The results are insightful for policymakers seeking to optimize public assistance programs.

Publications: At the date of thesis submission, the paper has not been published in any form.

4.1 Introduction

Energy efficiency assistance programs are a widely used policy tool to improve building efficiency and support low-income households in times of high energy prices. The EU Energy Efficiency Directive lists targeted schemes to assess energy consumption and provide better information, access to grants and subsidies, and technical and administrative assistance as suitable instruments to address energy poverty (European Union, 2023). Most assistance schemes combine financial support and access to information to address monetary and informational barriers, thereby aiming to narrow the energy efficiency gap in low-income households. However, previous research on such programs has shown that uptake among eligible households remains disappointingly low – even when programs employ elaborate encouragement designs for participation in audits and offer highly subsidized or free-of-charge investments in efficiency improvements (Fowlie et al., 2018). Strategies to increase take-up in programs include an increase in monetary incentives, improved design of information materials and targeted encouragement via home visits and phone calls (Fowlie et al., 2018; Chlond et al., 2024, 2025).

The role of energy advisors – the human factor in assistance programs – has received little attention in the literature. However, advisors could be an important element in improving outreach to low-income households as they serve as the bridge between the program and its beneficiaries. Their role in encouraging program uptake and energy-efficient investment among eligible households could be crucial as they serve as the program's human face and provide a personal point of contact for households navigating an otherwise abstract and bureaucratic process. While program subsidies provide attractive monetary incentives and information campaigns raise awareness on energy efficiency, they do not directly address practical and psychological barriers to navigating administrative processes, applying for and claiming benefits. Energy advisors can provide personalized guidance, build trust, and tailor information to individual households' needs beyond the standard program procedures. Previous research suggests that energy advisors influence whether households follow up on energy audits (Palmer et al., 2015). However, the literature does not provide evidence on the specific factors that make advisors more or less successful – crucial information for program managers seeking to optimize their staffing decisions on frontline workers.

This paper provides first evidence on the characteristics of a successful energy advisor, delivering a comprehensive study of advisor determinants on households' investment decisions. We study energy advisors in the context of a large energy efficiency assistance program in Germany, where advisors administer home energy audits to low-income households. The program provides a favorable setting to study the role of advisors for household decision-making as it is based on a large amount of personal interactions between advisors and households. Our analysis examines the impact of advisors on households' decisions to replace inefficient refrigeration appliances and identifies which advisor characteristics and traits contribute to successful appliance replacement. We use data from a survey among 113 energy advisors in the assistance program that we link with around 6,800 audits conducted by the surveyed advisors in the period 2018 to 2023. We show that advisors are not deliberately assigned to advisees based on demographic characteristics, and their reported preferences and attitudes are uncorrelated with their audit experience. Given the random assignment of advisors and the assumption that their preferences and attitudes are exogenous, our analysis provides an unbiased estimate of the effect of advisor characteristics on household decision-making.

Our findings document significant heterogeneity in advisor performance – advisors matter for households' investment decisions. We identify advisors' economic preferences and their attitudes toward the program as relevant determinants of household decision-making. Additionally, advisors who had recently replaced their own refrigerators are more effective in convincing households to do the same. Surprisingly, most demographic characteristics of advisors, such as gender or age, do not matter. However, we document demographic peer effects: similarities between advisors and households —such as a shared migration background, age, or being a parent— affect households' investment decisions.

Our study makes three contributions. First, we systematically analyze the role of energy advisors in assistance programs, identifying key characteristics that drive household investment decisions. Second, we extend the literature on peer effects documenting vertical effects between benevolent advisors and advisees with shared traits. Third, we contribute to the literature on the energy efficiency gap, shifting the focus from household preferences and biases to the influence of external actors—advisors—on investment behavior. To our knowledge, we are the first to open the "black box" of the human factor in energy efficiency assistance programs, systematically studying the role of energy advisors and identifying advisor preferences and characteristics as determinants of audit success. While Palmer et al. (2015) document that heterogeneity in energy advisor quality matters for follow-up by households after home energy audits, the study does not provide evidence for the determinants of advisor quality. Additionally, related work on energy efficiency has demonstrated that the human factor matters in other domains: Christensen et al. (2023) find that heterogeneity in contractor performance in weatherization projects significantly impacts achieved efficiency improvements, while Giraudet et al. (2018) show that contractor performance varies by the day of the week. We contribute to this literature by taking a systematic perspective in evaluating the role of energy advisors, considering a broad set of personal characteristics, preferences and attitudes.

Additionally, our findings link to broader evidence on factors influencing household investment decisions in energy efficiency and other contexts. Own actions by frontline workers seem to matter. For instance, community organizers who install solar panels themselves are more persuasive in encouraging others to do the same (Kraft-Todd et al., 2018). Social connections also play a role in investment decisions: Peer effects have been documented in a number of energy- and non-energy related choices, such as PV adoption (Bollinger and Gillingham, 2012; Graziano and Gillingham, 2015), water consumption behavior (Bollinger et al., 2018), education outcomes (Sacerdote, 2001; Zimmerman, 2003), choice of financial saving or health care plan (Mugerman et al., 2014; Sorensen, 2006), and use of welfare schemes (Bertrand et al., 2000). These studies look at the effects of horizontal interactions between neighbors, college roommates and co-workers. Fewer studies look at effects of vertical interactions in advisor-advisee or sales representative-customer settings, for instance, in the context of collecting charitable donations and providing financial advice (List and Price, 2009; Bucher-Koenen et al., 2021; Stolper and Walter, 2019). We add evidence on a novel setting in the vertical dimension, where advisors are benevolent and purely intrinsically motivated, documenting the existence of peer effects.[1]

By analyzing determinants of households' energy efficiency investment choices we also contribute to the large literature on the energy efficiency gap (Allcott and Greenstone, 2012; Gillingham and Palmer, 2014; Gerarden et al., 2017). A growing literature concerns the extent to which individual preferences and behavioral biases affect the decision to invest in energy-efficient technologies (Qiu et al., 2014; Newell and Siikamäki, 2015; Bradford et al., 2017; Schleich et al., 2019; Heutel, 2019; Fischbacher et al., 2021). These studies document significant correlations between households' loss aversion, patience, risk attitudes, pro-sociality as well as pro-environmental attitudes with investments in energy efficiency. This literature has studied preferences and behavioral traits of the household taking the investment, however, not of the advisor. We provide evidence on this factor, documenting that advisors' characteristics shape households' investment decisions.

The rest of the paper is organized as follows: Section 4.2 provides background on the energy efficiency assistance program. Section 4.3 describes the data. Section 4.4 discusses the identification of advisor effects on household investment choices. Section 4.5 presents the results, followed by a discussion in Section 4.6. Section 4.7 concludes.

[1]Advisors in the program studied do not receive commission-based incentives for successful refrigerator replacements by advised households. The "product" they "sell" to households – replacing their old inefficient refrigerator and claiming the program subsidy – is always in the economic interest of households as eligibility to the subsidy is conditional on passing a cost-benefit calculation.

4.2 Stromspar-Check energy efficiency assistance program for low-income households

Stromspar-Check (German, "Electricity-saving-check") is a Germany-wide initiative that aims to support low-income households to reduce their electricity consumption. The program is operated jointly by the German Caritas Association, one of the largest social welfare organizations in Germany, and the Association of Energy and Climate Protection Agencies (eaD) since 2009 with currently more than 100 local branches throughout the country.[2] Stromspar-Check consists of two key instruments to help low-income households to save electricity: A free home energy audit for all households that register in the program and a Refrigerator Replacement Program for a subset of the households that have participated in the audit and have been found eligible for refrigerator replacement.

To qualify for participation in a home energy audit, households need to be recipients of one of various types of federal income support or their income must be below the income limit for attachment.[3] Recruitment of qualified households into participation in home energy audits takes place through a variety of channels. The program is actively promoted in many employment and social assistance agencies across the country through printed and audiovisual material. The local branches also advertise their services with pop-up booths in shopping streets, malls and in front of the employment agencies and social assistance agencies. The program also maintains a website where information on the services offered is available in eleven languages. Program officials have no systematic understanding of how its different channels contribute to overall recruitment.

Since the program started in 2009, more than 450,000 households have participated in a free-of-charge home energy audit. Audits are administered by the local branch where a household registers. The typical home energy audit consists of two visits to the household by a two-person team within a period of around three weeks. During the first home visit, the energy advisors make an inventory of all electric devices and their usage in the household and measure water consumption. They also record the age and electricity consumption of refrigerators and freezers. Based on the inventory, the energy advisors assemble an energy-saving kit to help households save electricity,

[2]The project is funded by the Federal Ministry of the Environment. In 2022-2023, the Federal Ministry for Economic Affairs and Climate Action was the sponsor of the program.

[3]Eligible types of federal income support include unemployment benefits ("Arbeitslosengeld II"), housing allowances ("Wohngeld", "Sozialhilfe"), low pensions ("Grundsicherung"), child supplements ("Kinderzuschlag") or benefits for asylum seekers ("Leistungen nach Asylbewerberleistungsgesetz"). In 2020, more than 7 percent of German households qualified on this basis (Bundesagentur für Arbeit, 2020).

including items such as LED light bulbs, water-saving shower heads and switchable socket strips. Using the records collected on refrigeration appliances, the advisors determine the eligibility of households to join the Refrigerator Replacement Scheme of the program. During the second home visit, households receive a report with advice on electricity-saving behavior, items from the kit are installed, and the household learns about eligibility for subsidized refrigerator replacement.

The Refrigerator Replacement Scheme encourages households to replace their old inefficient appliances with modern efficient models. In Germany, refrigerators and freezers account for roughly 25 percent of the electricity consumption of the average household so that an upgrade to an efficient appliance provides considerable saving potential in the electricity bill. The vast majority of households in Germany own their household appliances, which means that they are responsible for their replacement and repair. Eligibility for the scheme is determined based on the age and measured electricity consumption of owned appliances. Households are eligible for a replacement subsidy if their refrigerator or freezer is older than 10 years and the replacement with an efficient appliance would save the household at least 200 kWh annually.[4] During the second home visit, eligible households receive an information letter that informs them about the scheme and invites them to request a replacement voucher worth 100 Euro.[5] Besides the federal replacement subsidy, complementary voucher schemes are in place in various states.[6]

Upon receiving the voucher, households face a two months deadline to purchase a refrigeration appliance and redeem the voucher. To be able to redeem their voucher in cash with their local Stromspar-Check branch, households need to meet a number of criteria: They need to present the purchase receipt; document that the purchased device is of the required energy efficiency standard,[7] 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

[4]The savings expectations are based on engineering estimates: Based on the inventory data, energy advisors use a custom database to calculate expected savings according to a comparison between the current device and a reference device of equivalent size and features that fulfills the required efficiency standard.

[5]Before April 2019, the federal subsidy was set at 150 Euro.

[6]Hamburg has offered a top-up subsidy of 100 Euro since September 2010, Saxony-Anhalt of 75 Euro since May 2020, and Berlin of 50 Euro since December 2020. Since July 2016, North Rhine-Westphalia has offered a graded top-up subsidy according to household size: single households receive 50 Euro, 2-person households 100 Euro, and 3-person households 150 Euro. Households with 4 or more persons receive 200 Euro. This same graded scheme was adopted by Berlin in October 2022. A different grading system was introduced by Saarland in May 2020 and by Hamburg in April 2023, single households receiving 50 Euro, 2-person households 100 Euro, 3- to 4-person households 150 Euro, and 5 or more persons receiving 200 Euro on top.

[7]Up to February 2021, the new appliance needed to comply with the A+++ standard. After the readjustment of the European label categories implemented in March 2021, new appliances now need to at least comply with energy class D.

fulfills the requirements, pre-financing the purchase, and organizing the logistics of delivering the new and of disposing of the old refrigerator.

The program is based on a large amount of personal interactions between advisors and households. Advisors spend a good amount of time in the households: Depending on the size of the dwelling, the first visit usually takes between 60 and 90 minutes. During the second visit, advisors spend another around 20 minutes in the household. As the program procedures force the households to welcome the advisor into their homes, the households need to have a trusting attitude toward the advisor and the program. During the audit, a hierarchical relationship between advisor and household is established as the advisor inspects all rooms of the dwelling and records details of all energy-consuming activities, including intimate details about the households' lifestyle, such as personal hygiene routines.[8] The extensive interactions and particular relationship between advisors and households make the hypothesis that advisors have scope to affect investment decisions of advised households a plausible one.

4.3 Data

To determine the contribution of energy advisors to households' investment decisions, we analyze data from the Stromspar-Check database on audits that took place between January 2018 and February 2023. For each audit, we observe household information and which advisors administered the home visit. The analysis focuses on two outcome variables of interest. In the first stage, households request a replacement voucher, i.e., enrolling into the program and confirming their intentions to replace their refrigerator. In the second stage, households redeem the voucher and replace their refrigerator against a new, efficient one. We look at outcomes (1) voucher request and (2) voucher redemption and appliance replacement. Voucher request is an intermediate outcome and can be seen as a proxy for the intention by households to replace their old refrigerator. Voucher redemption directly tracks households' investment decisions. Advisors may influence both stages of household decision-making via their interactions with households, providing information to the households, explaining them the details of the Refrigerator Replacement Scheme and motivating them that the energy savings will be worth the effort. While the letter informing households about their eligibility to the program is standardized, advisors have freedom in how they communicate – and advertise – the scheme and the costs and benefits of replacing the refrigerator to households.

[8] For instance, advisors ask how frequently household members shower, take a bath and use the toilet.

4.3.1 Households

For the period January 2018 to February 2023, the database contains information on 127,147 audits. Each audit has information on the local Stromspar-Check branch which administered the audit and the date of the first and the second home visit. The database also records the audit details, such as eligibility for the refrigerator replacement voucher and the records concerning the age and electricity consumption of old refrigerators and freezers in the household. Moreover, the database records whether households request a voucher to enroll into the program and whether they replaced their appliance and redeemed their voucher. Demographic information on the households contains the number of inhabitants, their age (child under 18, adult, pensioner), the type of government benefit that the household receives (see Footnote 3 for further information), the state and the ZIP code of the household's location.

For the purpose of analyzing the role of advisors in households' investment decisions, we only keep the observations of households who were found eligible for the replacement voucher. This leaves us with 41,196 audits. If households own more than one cooling appliance, for instance, a freezer besides a refrigerator, the database contains several appliance observations per audit. Our data records 52,647 appliance observations. For preparation for the analysis, we clean the data. We drop values below the eligible age of 11 years of the old appliance and below 200 kWh for estimated savings from replacement, leaving us with 48,701 appliance-level observations. Table 4.3 provides summary statistics for the prepared data in Panel (i). Out of all eligible appliances, households request a voucher for 39 percent and redeem the voucher for 26 percent. On average, three persons live in a households with old appliances on average 18.3 years old. Households pay on average 0.29 Euro per kWh electricity and the estimated savings from replacing the appliance are at 357 kWh. This means that the average annual savings potential of households is at 104 Euro during the observation period. Figure A4.1 in the Appendix plots how these characteristics evolve over time. The number of inhabitants is stable over time, except for a drop in 2021. Estimated savings from replacement only increase in 2023. The age of old appliances is more volatile, however without a trend visible. The electricity price per kWh paid increases over time, rising strongest during the last two years. This coincides with a general increase of retail electricity prices for German households following the energy crisis after the Russian invasion in Ukraine. In the analysis, we control for these characteristics at the household level to account for variation over time.

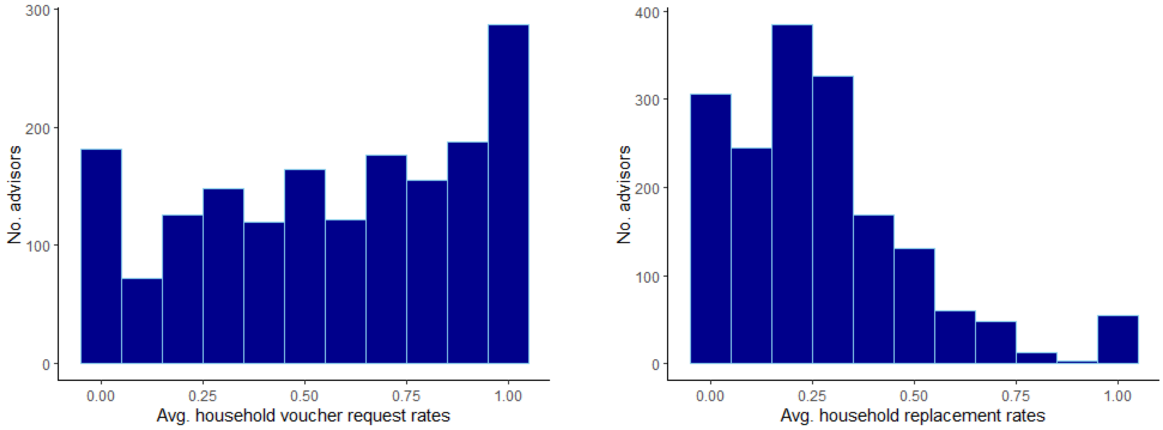


Figure 4.1: Average voucher request and replacement rate of households by advisor
Notes: This figure shows the distribution of average voucher request and appliance replacement rates of households by the advisor who administers the audit. The left-hand side shows voucher request rates and the right-hand side shows appliance replacement rates.

4.3.2 Energy advisors

Information for each audit in the database also contains the IDs of the pair of energy advisors that conducted the home visits. We observe 2,315 different advisors in the period 2018 to 2023, for whom we know when they visited which households. To get a first idea of the heterogeneity in the performance of advisors, we calculate the performance of each advisor in terms of the average voucher request and appliance replacement rates of advertised households. As shown in Figure 4.1, heterogeneity for both outcomes is high. Voucher request rates by advisor (left-hand plot) are relatively uniformly distributed – with slightly more weight at the upper end of the distribution, while for refrigerator replacement rates (right-hand side plot) advisors with low and very low rates are clearly in the majority. Only for a small group of advisors, at least three out of four advised households replace their appliances.

4.3.3 Survey among energy advisors

To further inspect the role of advisors, we conducted a survey among the energy advisors at Stromspar-Check in March and April 2022. The survey elicited the advisors' opinions regarding and assessment of the Refrigerator Replacement Program and their general satisfaction with the working environment at Stromspar-Check. We also elicited advisors' risk, time and altruism preferences on a ten-point Likert scale, following the established measures of Falk et al. (2018). We moreover surveyed standard socio-economic and demographic characteristics and asked whether they had replaced their refrigeration appliances within the last 10 years. Invitations to participate in the survey were sent out via email to the local branch managers who distributed the invitation to their energy advisors. In total, 136 advisors participated in the survey which relates to a share of more than 20 percent of the advisors active in the program during the survey period.

Table 4.1: Survey summary statistics - Program variables and socio-demographic characteristics

Program variables			
Contract	N	*Joined program*	N
Management	20	2009 - 2012	22
Short-term	74	2013 - 2015	17
Long-term	14	2016 - 2018	28
Others	16	2019 - 2021	52
		2022	5
Socio-demographic characteristics			
Gender	N	*Age in years*	N
Female	19	20 - 29	2
Male	93	30 - 39	15
		40 - 49	28
		50 - 59	47
		60 - 69	29
Education	N	*Native language*	N
University degree	19	German	104
Vocational training	39	Other native language	16
High school degree or equivalent	52		
Family status	N	*Children*	N
In relationship	41	Yes	21
Single	70	No	85

Notes: The table provides summary statistics on the survey answers of the 136 surveyed advisors. "High school degree or similar" refers to persons who report to have obtained either *Fachabitur*, *Mittlere Reife* or *Hauptschulabschluss*.

Table 4.1 reports information on contract type and the year when advisors joined the program as well as socio-demographic characteristics. The majority of advisors have short-term contracts of up to two years, joining the program as part of a labor market reintegration measure. Most advisors joined in the period 2019 to 2021. We also observe a significant share of advisors who joined the program in earlier funding periods and a few who only joined a few weeks before we conducted the survey. The majority of advisors is male, German is their native language, they are single and have no children. Most advisors are middle-aged between 40 and 59 years old. In terms of education, most advisors have obtained the German equivalent of a high school degree. A significant share completed a vocational training, stressing the technical nature of the job.

Table 4.2: Survey summary statistics - Economic preferences and attitudes toward the program

	N	Mean	Std. Dev.	Min	Median	Max
Economic preferences						
Altruism	110	7.600	2.267	0	8	10
Risk	107	5.860	2.405	0	6	10
Patience	105	6.543	2.245	0	7	10
Attitudes towards the program						
Program is a good idea	132	1.614	0.561	0	2	2
Right households profit	129	1.302	0.645	0	1	2
Households save energy	132	1.712	0.560	-2	2	2
Households protect climate	132	1.561	0.621	-2	2	2
Voucher worth the effort	130	0.931	0.873	-2	1	2
Own refrigerator replacement						
Replacement $\leq$ 10 years ago	103					
Replacement $>$ 10 years ago	15					
Replacement timing not reported	11					

Notes: The table provides summary statistics on economic preferences, attitudes towards the program and last replacement of the own refrigerator by the 136 surveyed advisors. Economic preferences were elicited on a ten-point Likert scale, following established measures of Falk et al. (2018). Attitudes towards the program were elicited on a five-point scale ranging from strong disagreement (-2) to strong agreement (2) regarding the statements about the nature of the program.

Table 4.2 gives an overview of the surveyed advisors' economic preferences and attitudes toward the program. On average, advisors agree most with the statement that the program helps households to save energy and to protect the climate, and that the program is generally a good idea. Lower average agreement is found for statements regarding the right households profiting and, particularly contended, the replacement voucher being worth the administrative effort. Moreover, we survey how many years in the past advisors have last replaced their own refrigerator. The majority of advisors has done the replacement 10 years ago or more recently.

4.3.4 Merge of audit and survey data

We combine the survey data with the information from household visits in the Stromspar-Check database in the period 2018 to 2023.[9] For the purpose of our analysis, we create a dataset that connects the survey answers of the advisors to each appliance observation of any audit that they conducted. We merge these two data sources by the advisor ID, which the advisors reported as part of the survey and which is recorded in the database indicating the advisors who administer an audit. As each audit is conducted jointly by two advisors and we are interested in the link between each advisor's traits and audit outcomes, we reshape the database data from wide, reporting two energy advisors per audit and appliance screening, to a long format where the unit of observation is the advisor level. Each audit consists then of two advisor-audit observations that record different main advisors and the respective other advisor as co-advisor.[10] The team of advisors generally conducts all tasks concerning an audit jointly.[11] For households that own more than one cooling appliance the data restructuring from wide to long means that we create for each appliance observation in the audit two observations with different main advisors. Following the merge of survey and database data as described above, we observe 10,380 appliance observations in 6,851 audits by 113 advisors who answered the survey. The merged dataset includes 18 percent of all appliance observations and 17 percent of all audits from the database. We can connect 83 percent of all survey participants, losing 23 advisors who answered the survey but have no corresponding observations in the database.[12] The remaining 113 advisors are distributed over 88 different local program branches.

We use three datasets for different purposes in the analysis, (i) the original database

[9]We have information on household visits going back to 2009 but we choose the period from 1 January 2018 due to two reasons. First, a new set of program procedures was introduced on this date. Before 2018, households eligible for refrigerator replacement directly received a replacement voucher during the second home visit instead of receiving the information letter. Choosing the period since 2018 holds program procedures constant in our analysis. Second, we face a trade-off between sample size and temporal proximity of audit data to the survey date. We do not consider information on households' investment decisions that originates from audits that took place a long time ago. Hence, we only take into account audits that took place at most four years before and one year after the survey took place.

[10]Whether both of these audit-advisor observations are included in our merged dataset depends on whether both advisors participated in the survey. We only observe a few of these cases in our merged data.

[11]Only few advisors conduct audits by themselves without a co-advisor. These are advisors with additional expertise that have completed an additional technical training and usually have worked in technical professions related to energy efficiency. In our data, this concerns 9,073 audits. For these audits, we only have one audit-advisor observation per audit and appliance screening. We introduce a joint value for these cases in the variable that indicates the co-advisor ID so that these observations remain part of the analysis in specifications in which we include co-advisor fixed effects.

[12]Potential explanations are that these advisors are management staff that do not conduct household visits, that they have reported an incorrect ID number in the survey, or they did not conduct audits in households eligible for appliance replacement.

Table 4.3: Overview of datasets

	N	Mean	Std. Dev.	Min	Median	Max
(i) Database data (wide format)						
Voucher request	48,701	0.386	0.487	0	0	1
Appliance replacement	48,701	0.259	0.438	0	0	1
No. inhabitants	48,699	2.993	1.855	1	3	15
Electricity price per kWh (€)	48,695	0.290	0.037	0.030	0.290	0.930
Old appliance age (years)	48,701	18.269	5.226	11	18	45
Est. savings from replacement	48,701	357.131	181.819	200.000	301.130	4,284.000
(ii) Database data (long format)						
Voucher request	97,402	0.386	0.487	0	0	1
Appliance replacement	97,402	0.259	0.438	0	0	1
No. inhabitants	97,398	2.993	1.855	1	3	15
Electricity price per kWh (€)	97,390	0.290	0.037	0.030	0.290	0.930
Old appliance age (years)	97,402	18.269	5.226	11	18	45
Est. savings from replacement	97,402	357.131	181.819	200.000	301.130	4,284.000
(iii) Merged data: Database and survey data (long format)						
Voucher request	10,380	0.345	0.475	0	0	1
Appliance replacement	10,380	0.230	0.421	0	0	1
No. inhabitants	10,380	2.926	1.850	1	3	14
Electricity price per kWh (€)	10,380	0.303	0.051	0.130	0.300	0.900
Old appliance age (years)	10,380	18.530	5.518	11	18	45
Est. savings from replacement	10,380	367.355	199.752	200.000	299.720	2,943.330

Notes: The table provides summary statistics on outcomes and household characteristics for the three datasets used in the analysis with observations at the appliance level. The time period covered in each dataset stretches from January 2018 to February 2023.

data in wide format at the appliance level, (ii) the database data reshaped to long format at the advisor-appliance level, and (iii) the merge between database and survey data in long-format, also at the advisor-appliance level. First, the original database excerpt with observations at the household-appliance level is used for checking whether advisors have an impact on households' investment decisions. Second, the data reshaped to long format is used to further investigate the heterogeneity in advisor performance and to extract advisor-level fixed effects. And third, the merged database-survey data containing only observations of advisors who participated in the survey is used to determine the role of advisor characteristics and peer effects between advisors and households on household investment. See Table 4.1 for summary statistics of household characteristics across the three datasets.

4.4 Identification of advisor effect on household investment

Before investigating the relevance of advisors and their preferences and attitudes for households' investment decisions, we discuss the necessary conditions that need to hold in order to estimate an unbiased effect of advisors on household voucher request and replacement rates. Advisors' characteristics and preferences need to be exogenous to any factors that determine household decision-making. For exogeneity, two conditions need to hold: fixed advisor characteristics and quasi-random assignment of advisors to households.

First, advisors' characteristics and preferences need to be fixed, so that advisor characteristics affect household investment, so that there is no effect of households on advisors. Demographic characteristics, such as gender, age, education, native language, and family status, are most plausibly determined pre-audit and not affected by households. Economic preferences and attitudes toward the Replacement Scheme are elicited in the survey conducted in March and April 2022. Economic preferences can be thought of as relatively stable over time. Attitudes toward the program could theoretically change over time, as advisors understand the program more in depth and collect more experience. If preferences and attitudes are not stable over the observation period and change with the audit experience collected, it could introduce bias into the estimation.

To address these concerns, we have formulated survey questions in a general way following established measures for elicitation of economic preferences (Falk et al., 2018) and asking advisors questions about their general attitudes toward the program so that specific audits are unlikely to affect their answers.[13] Moreover, we conduct empirical tests whether, first, collected experience in advising is correlated with household appliance replacement and, second, collected experience up to the survey period is correlated with preferences and attitudes reported. Evidence for either of these relationships would hint at preferences, attitudes or other unobserved variables changing over time, putting the exogeneity assumption under threat. The first test checks whether there is learning in the program that improves advisors' performance over time. For this purpose, we number all replacement audits from first to last by advisor.[14] In different

[13] The timing of the survey supports this assumption as the program was on a break between mid-February and beginning of April between two funding cycles. Most advisors who took the survey had been on a break from home visits for a few weeks, providing them with some time to take a step back and facilitating a more general assessment of the program as a whole rather than taking the survey in close temporal proximity to having delivered an audit.

[14] For numbering each audit by advisor, we make use of the full dataset going back to the start of the replacement program in 2009 to get the accurate numbering for advisors who joined the program before 2018. We only take into account audits with for appliance replacement eligible households as these audits provide potential learning for advisors regarding the outcomes of interest.

specifications, we test whether the number of audits administered at the advisor level is related to the household investment decision. Table 4.4 shows that the number of audits administered by an advisor is not related to households' replacement decisions across specifications. Apparently, advisors do not improve their performance over time, no learning seems to take place. The second test checks whether over time, as advisors collect more experience, their preferences and attitudes change. To this end, we count the number of audits each advisor who participated in the survey had administered until February 2022 (the month before the survey was conducted). We test whether the total number of audits administered up to the survey date is correlated with reported preferences and program attitudes. Table 4.5 shows the relationship between the total number of audits administered until February 2022 and preferences and attitudes toward the program. Economic preferences are not significantly related to the total number of audits (see column (1)). Out of all measures of attitudes, only agreement with the statement "Generally, the program is a good idea" is significantly related to the number of audits administered (see column (2)). However, when controlling for economic preferences in column (3), the relationship turns insignificant. We conclude that there is no robust evidence for the experience of advisors to affect either households' investment decisions or advisors' preferences and attitudes.

The second condition is that assignment of advisors to households needs to be quasi-random. If advisors are systematically assigned to households based on observable or unobservable characteristics. This would mean that advisor-household matches are not random and estimates of the effect of advisor characteristics on households' investment may be biased. The effect of advisor characteristics on audit outcomes would be conditional on a certain pre-selection scheme by branch managers. With active matching, the advisor-household matches that we observe would be a selected subset out of all possible combination and we would be much more likely to observe certain combinations of advisors and households than others. For instance, if branch managers systematically assigned patient advisors to households who struggle with understanding the benefits of refrigerator replacement or with grasping the administrative procedures of claiming subsidies, the estimated effect of patience on household investment decisions would be downward biased. The managers at the local branches coordinate advisors and household visits. As part of their management tasks, they assign advisors to households. If managers at the local branches have beliefs regarding the optimal matching of their advisors with advised households, they may attempt to actively select the advisor-household matches that they expect would optimize audit outcomes.[15] For active matching, managers require (i) information on characteristics

[15]Note that while the refrigerator replacement decision of households is one dimension of program success, the main focus of the home visits is on delivering electricity savings via the installation of the saving kits and advice on behavioral adjustments to take. Active matching based on achieving better

Table 4.4: Relationship between household investment and advisor experience

Dependent Variables:		Replacement	
Model:	(1)	(2)	(3)
No. audit	7.71×10^{-5}	1.68×10^{-5}	
	(0.0001)	(0.0001)	
No. audit squared		1.3×10^{-7}	
		(1.77×10^{-7})	
No. audit Ref. cat. [0,25]			
(25,50]			-0.0010
			(0.0075)
(50,100]			-0.0082
			(0.0096)
(100,200]			-0.0103
			(0.0121)
(200,500]			-0.0069
			(0.0186)
(500,Inf]			0.0273
			(0.0399)
Household controls	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes
Observations	85,536	85,536	85,536
R^2	0.28256	0.28257	0.28259

Notes: Standard errors in parentheses are clustered by advisor and co-advisor. Household characteristics include the number of inhabitants, electricity price per kWh, old refrigerator age, estimated savings from refrigerator replacement and the governmental transfer type received. Fixed effects account for ZIP code, year-month, branch-year, advisor and co-advisor. The data used for this analysis is in the form of the database data in long format (see Section 4.3). p-values ***: 0.01, **: 0.05, *: 0.1.

of their advisors and (ii) information on households who have registered for a home visit. Managers plausibly know their advisors relatively well and have information on their demographics as well as softer characteristics. However, managers have quite limited information on households before the audit which they could use for an optimal advisor assignment. Before the audit, branch managers usually know the name of the family, the address as well as the type of income support received by the household. Hence, based on the pre-audit information of branch managers the potential for active matching is rather limited. Moreover, in practice active matching is restricted by the

outcomes in these dimensions is only reason for concern if advisors' performance in the replacement scheme is correlated.

Table 4.5: Relationship between advisor economic preferences and attitudes and experience

Dependent Variable:	Total no. audits		
Model:	(1)	(2)	(3)
Economic preferences			
Altruism	-8.495		-12.18
	(7.000)		(7.966)
Risk	6.569		8.324
	(7.231)		(7.665)
Patience	-10.83		-10.65
	(6.981)		(7.347)
Attitudes towards the program			
Program is a good idea		-46.14*	-39.74
		(23.82)	(30.47)
Right households profit		8.520	25.27
		(19.39)	(28.81)
Households save energy		19.98	38.61
		(27.62)	(38.96)
Households protect climate		21.85	22.93
		(23.70)	(31.37)
Voucher worth the effort		0.2890	-11.26
		(14.38)	(20.47)
Observations	80	111	79
R^2	0.07802	0.04408	0.12727

Notes: The data used for this analysis is the count of the total number of audits administered by 28 February 2022 by advisor merged with survey information about advisors' economic preferences and attitudes toward the program. p-values ***: 0.01, **: 0.05, *: 0.1.

timing of audits and availability of advisors.

Nevertheless, we conduct a short test on whether the data provide evidence for active matching based on observable characteristics of advisors and households. We look at matching of advisors and households that share the same demographic characteristics. This is the only information in our data we can use to check for non-random assignment of advisors. However, it is also the most plausible information, branch managers could base optimal matching upon. If managers' beliefs align with empirical evidence from other contexts, they would identify homophily as a potentially relevant factor that may facilitate advisor-advisee interactions and, therefore, positively affect audit success. We check for active matching by comparing the expected share of matches predicted by the product of the share of advisors with a certain demographic characteristic and the share of households with that same characteristic with the realized share of advisor-household matches in the data. If the observed share of matches is higher than the expected one, this would be evidence for assignment of advisors to households based on demographics. We observe three demographic characteristics that we can attribute to both advisors and households:

- Age: We test matching between advisors of at least 60 years of age and households with all inhabitants of age 65 years or older.
- Migration: We test matching between advisors that report a different native language than German and households that are seeking asylum in Germany.[16]
- Parental status: We test matching between advisors that report to have children with households living with children.

Table 4.6: Matching of demographic characteristics at household level

	Share advisors	Share households	Expected matches	Observed matches	Difference	χ^2-Statistic
Age	38.0%	8.7%	3.3%	3.4%	-0.1	0.129
Migration	16.3%	1.1%	0.2%	0.1%	0.1	2.062
Children	16.1%	62.5%	10.1%	8.5%	1.6	15.196

Notes: The table provides results for one-sided proportion tests to compare the expected versus the observed share of demographic matches between advisors and households at the audit level. The null hypothesis of each test is no difference between expected and observed matches. The alternative hypothesis is the observed share of matches being greater than the expected share. The χ^2 statistic for each test is reported in the last column.

Table 4.4 provides the results for the assignment test. The first two columns report the share of audits in which advisors and households which the respective characteristic are observed. Based on these percentage shares, the expected share of matches is calculated in the third column. The fourth column reports the share of advisor-household matches observed in the data. The fifth column takes the difference between observed and expected share, and the sixth column reports the chi-squared test statistic for a one-sided proportion test, checking whether the difference is significantly greater than 0. The observed share of matches is greater than the expected share only for one out of three studied characteristics: Advisors at least 60 years old visit households with all inhabitants 65 years or older more often than predicted. However, the difference

[16]We identify households via their type of social transfer scheme received: “Asylbewerberleistungs-gesetz” benefits.

is small at 0.1 percentage points and not statistically significant (p=0.360). Expected and observed share of matches between advisors with another native language than German and asylum-seeking households are also very close, the expected number of matches being slightly higher than the observed number (p=0.925). And the observed share of matches between advisors and households with children is around 2 percentage points below the expected share (p=1.000). We conclude that there is no evidence for systematic assignment of advisors to households based on demographic characteristics. As we do not find evidence against either of both exogeneity conditions, we conclude that the estimated effects of advisor characteristics on household investment are plausibly exogenous and unbiased.

4.5 Results

Results are grouped according to the dimensions of impact in Section 4.5.1 – investigating the relevance of advisors for households' investment decisions, determinants in Section 4.5.2 – focusing on the role of specific advisor characteristics, and peer effects between advisors and advised households in Section 4.5.3.

4.5.1 Impact: Relevance of advisors for household investment decisions

We test whether advisors at Stromspar-Check have a significant impact on households' decision-making. We do so by estimating a model of both measures of investment, (1) voucher request and (2) refrigerator replacement, and compare the explanatory power of such a model with and without advisor fixed effects. The basic model is as follows:

$$Y_i = \beta_0 + \beta_1 X_i + \gamma_z + \delta_m + \zeta_b + \varepsilon_i$$

We control for important household characteristics X, such as the number of persons in the advised household, the electricity price paid, the type of governmental transfer received and the age of the old appliance to be replaced. We add fixed effects that account for spatial and temporal variation. γ accounts for ZIP code-level effects, δ for year-month and ζ for branch-year effects. Subscript i denotes the household level, m the year-month, and b the branch-year.

In a next step, we add in advisor fixed effects, co-advisor fixed effects, and advisor pair fixed effects controlling for the combination of advisors that visit a household jointly. We observe how the explanatory power of the models changes when incorporating information on the advisors. For this estimation, we use the database data (in wide format, before merging with survey characteristics).

Table 4.7: Relevance of energy advisors for household decision to replace refrigerator

Dependent Variable:	Refrigerator replacement				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Inhabitants	0.0126***	0.0115***	0.0117***	0.0113***	0.0104***
	(0.0019)	(0.0019)	(0.0019)	(0.0020)	(0.0021)
Electricity price per kWh	-0.0632	-0.1106	-0.0658	-0.1121	-0.1630
	(0.0976)	(0.1017)	(0.1024)	(0.1110)	(0.1224)
Refrigerator age	0.0013***	0.0016***	0.0015***	0.0017***	0.0013***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0005)
Transfer: AsylbLG	-0.0088	0.0099	0.0005	0.0147	0.0049
	(0.0189)	(0.0189)	(0.0199)	(0.0203)	(0.0215)
Transfer: None	0.0623***	0.0669***	0.0635***	0.0691***	0.0795***
	(0.0164)	(0.0158)	(0.0164)	(0.0155)	(0.0183)
Transfer: Child supplements	0.0373	0.0372	0.0507*	0.0438	0.0502
	(0.0278)	(0.0288)	(0.0287)	(0.0287)	(0.0325)
Transfer: Pension supplements	0.0442***	0.0434***	0.0461***	0.0463***	0.0440***
	(0.0111)	(0.0107)	(0.0118)	(0.0113)	(0.0120)
Transfer: Basic income	0.0255	0.0285	0.0267	0.0277	0.0304
	(0.0193)	(0.0179)	(0.0201)	(0.0183)	(0.0223)
Transfer: Housing allowance	0.0835***	0.0847***	0.0846***	0.0875***	0.0833***
	(0.0156)	(0.0151)	(0.0166)	(0.0163)	(0.0175)
Est. savings from replacement	$8.2 \times 10^{-5***}$	$7.91 \times 10^{-5***}$	$7.65 \times 10^{-5***}$	$7.64 \times 10^{-5***}$	$8.57 \times 10^{-5***}$
	(1.65×10^{-5})	(1.7×10^{-5})	(1.73×10^{-5})	(1.79×10^{-5})	(1.98×10^{-5})
Fixed-effects					
ZIP code	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
Branch-year	Yes	Yes	Yes	Yes	Yes
Advisor		Yes		Yes	Yes
Co-advisor			Yes	Yes	Yes
Combination of advisors					Yes
Fit statistics					
Observations	48,686	48,686	48,686	48,686	48,686
R^2	0.21295	0.26047	0.26697	0.30838	0.39304

Notes: Standard errors in parentheses are clustered by advisor and co-advisor. p-values ***: 0.01, **: 0.05, *: 0.1.

Table 4.7 shows the results for the estimations for the outcome refrigerator replacement. The number of persons in the household as well as the age of the old appliance and the engineering estimate for expected savings from replacement are positively associated with the replacement decision; the electricity price paid per kWh is not significantly associated. The type of social security transfer received also affects the household decision. As compared to recipients of unemployment benefits, recipients of pension supplements and a housing allowance as well as households with an income below the level for attachment are more likely to replace their appliances. Model (1), without any controls for advisor-specific effects, has an R2 of 0.21. Explanatory power of the model increases when adding in advisor fixed effects in Model (2) (R2 0.26) or co-advisor fixed effects in Model (3) (R2 0.27). Including both sets of fixed effects jointly raises the R2 to 0.31 in Model (4). Additionally adding advisor pair fixed effects in Model

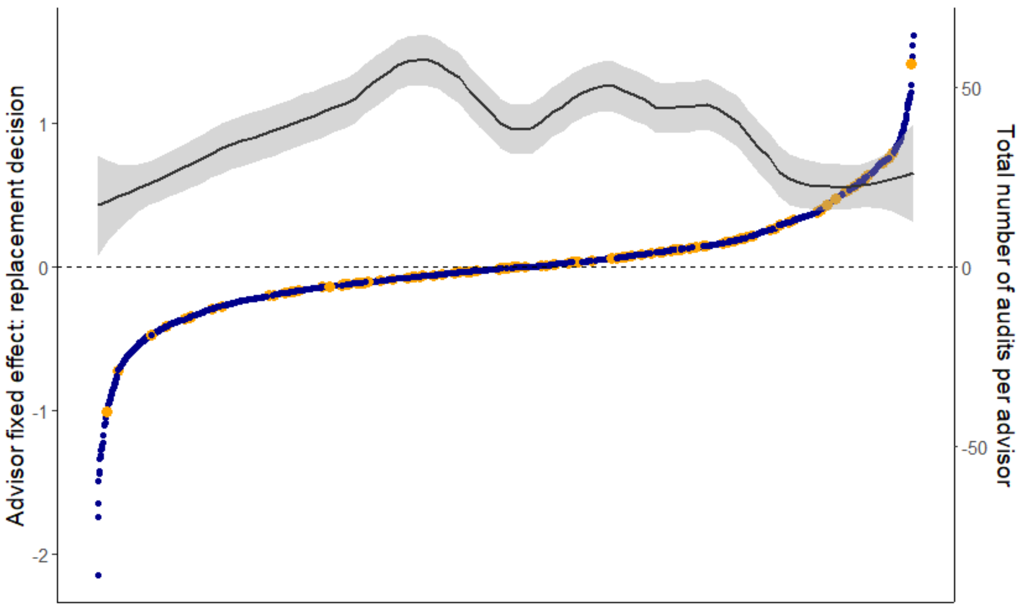


Figure 4.2: Advisor-fixed effect coefficients for the household investment rate
Notes: This figure shows the coefficients for advisor fixed effects sorted by sign and magnitude on the left-hand y axis in dark-blue. Orange observations mark the advisors who participated in the survey. On the right-hand y-axis, the corresponding total number of audits by advisor is plotted in a smoothed curve in black with gray 95% confidence intervals.

(5) further increases R2 to 0.39. Using Likelihood-Ratio tests, we find the increase in the explanatory power from each of the models compared to the baseline Model (1) to be statistically significant at the 0.1-percent level. Results for the first-stage outcome voucher request are qualitatively similar (see Table A4.1 in the Appendix). This exercise confirms the findings by Palmer and Walls (2015) who show that advisors are important for audit follow-up by the advised households and that an idiosyncratic component exists in audit success that is affected by advisors. In our setting, the result that advisor fixed effects increase the explanatory power of the model imply that advisors are relevant for households' investment decisions, both at the first and the second stage of decision-making.

To closer investigate the heterogeneity in the contribution of advisors to households' investment, we extract the advisor fixed effects estimated in column (4) in the specification above. We again use the full database data, but for this purpose we reshape the data to a long format to observe each advisor in all audits administered by her as main advisor. Figure 4.2 plots the coefficients of the advisor fixed effects for the regression with household appliance replacement as outcome in ascending order. The fixed effect coefficients (dark-blue dots) range from around -200 to 150 percentage points. At a mean replacement rate of 26 percent, the heterogeneity appears to be substantial. The same figure shows a smoothed curve of the total number of audits administered by advisor on the secondary y-axis (black line with gray confidence intervals). Advisors

Table 4.8: Summary statistics: Advisor fixed effects

Statistic	N	Mean	Std. Dev.	Min	Median	Max
All advisors						
Voucher request	2,314	0.020	0.470	−2.219	−0.007	2.428
Refrigerator replacement	2,314	0.007	0.361	−2.145	−0.007	1.609
Surveyed advisors						
Voucher request	113	0.109	0.433	−0.729	0.012	2.266
Refrigerator replacement	113	0.069	0.307	−1.007	0.020	1.414

Notes: The table provides summary statistics on the coefficients of advisor fixed effects. Fixed effects for both the full sample of advisors and the subsample of surveyed advisors are extracted from an estimation of model (4) in Table 4.7.

with a lower total number of administered audits are grouped towards the lower and upper tail of the fixed effects distribution – with fewer observations per advisor, the variance increases. To account for this effect and to check whether explanatory power of advisor fixed effects is not exclusively driven by these extreme data points, we re-conduct the Likelihood Ratio Test using the sample of audits conducted by advisor whom we observe in at least 50 audits – where coefficients lie in the range between around negative to positive 20 percentage points. We find advisor information remains significantly adding to the explanatory power of households' investment choices.

Table 4.8 presents summary statistics on the advisor fixed effects. The upper panel shows them for the full advisor population. The standard deviation of advisor fixed effects on the voucher request decision is higher than the one on the refrigerator replacement decision. Advisors seem to have more scope to influence the decision whether households request a voucher or not as compared to whether they purchase a new refrigerator. This is in line with expectations: the decision to request a voucher is much easier to implement for households, needs less effort and has no economic consequences. The first decision is also chronologically closer to the home energy audit so that the visits by advisors may be in fresh memory and therefore more impactful as compared to a few weeks later. The lower panel shows summary statistics for only the surveyed advisors – extracted from the estimation with the full sample of advisors. For both outcomes, mean and median are above the values for the full population, and the standard deviation is below. This suggests that the sample of surveyed advisors on average performs slightly better and the difference between the best and poorest performing advisors is lower. However, the distribution of orange data points in Figure 4.2 demonstrates that the surveyed advisors are a rather representative sample in terms of performance, spreading relatively evenly between the 25th and 75th percentile with slightly more weight above the median. In terms of the number of audits, these surveyed advisors tend to be the ones who have administered a large total number of audits. However, we also observe some advisors in the survey who we only observe in a few audits. We conclude from the impact analysis that the heterogeneity in the performance of advisors is substantial and that the assignment of advisors has an impact on households' investment decisions. Generally, our surveyed advisors seem to be a rather representative sample of the full advisor population.

4.5.2 Determinants: Effect of advisor characteristics on household investment decisions

We investigate whether socio-demographic characteristics, economic preferences, advisors' own energy efficiency choices and their attitudes towards the program as elicited in the survey affect households' decisions to request a voucher and replace their refrigerator. For this analysis we use the merged data in long format. We regress both outcomes of interest – voucher request and appliance replacement – on all advisor characteristics jointly:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_a + \beta_3 P_a + \beta_4 I_a + \beta_5 A_a + \beta_6 M_{ia} + \gamma_z + \delta_m + \zeta_b + \eta_c + \varepsilon_i$$

D represents socio-demographic characteristics, P economic preferences, I own investment by advisors in an efficient refrigeration appliance, and A attitudes toward the program. These characteristics all vary at the advisor level a. M provides a score for the demographic similarity of advisor and advised household, varying at the advisor-household level. In addition to ZIP code, month-year and branch-year effects, we include co-advisor c indicators η. Standard errors are clustered at the advisor level. For clarity, we only report the coefficients on the characteristics of interest from the joint regression in each of the following subsections. Table A4.2 in the Appendix provides an overview of all coefficients jointly.

Socio-demographic characteristics

In the survey, we ask advisors about their role in the program, i.e., whether they are management, long- or short-term advisors or have another type of contract, and about the year when they started working at Stromspar-Check. We also ask about demographic information, such as their gender, age, education, family status and children as well as whether their native language is German.

Results are shown in Table 4.9. The contract type only matters for voucher request, where advised households of advisors with a long-term contract show higher request rates. Additionally, the timing when advisors joined the program matters for both outcomes. As compared to advisors who have been active in the program since the first funding period (2008-2012), households advised by advisors who joined the program in the second and third cycle (2013-2015, 2016-2018) have by about 31 percentage points higher household voucher request rates. For appliance replacement rates, the effect of joining in the second or third cycle as compared to before is at 25 to 28 percentage points a bit lower than for the voucher request decision. In terms of socio-demographic characteristics, being a parent (30 percentage points higher) and being middle-aged (13 percentage points higher compared to young age group) is significantly associated with voucher requests. For appliance replacement, only education matters: Advisors with at least a high school degree have 14 percentage points lower replacement rates in advised households as compared to advisors who have completed a vocational training.

Table 4.9: Socio-demographic characteristics

Dependent Variable:	Voucher request	Refrigerator replacement
Model:	(1)	(2)
Program controls		
Position: Ref. cat. Management staff		
Position: Other	0.2318	0.1692
	(0.3273)	(0.2754)
Position: Short-term contract	0.3979	0.2003
	(0.3032)	(0.2562)
Position: Long-term contract	0.8317**	0.2202
	(0.4129)	(0.2672)
Joined program: Ref. cat. 2008-2012		
Joined program: 2013-2015	0.3111***	0.2496**
	(0.1110)	(0.0941)
Joined program: 2016-2018	0.3134**	0.2806**
	(0.1293)	(0.1223)
Joined program: 2019-2022	0.1681	0.2180
	(0.1849)	(0.1553)
Socio-demographic characteristics		
Gender: Female	0.0483	0.0401
	(0.0564)	(0.0510)
Age: Ref. cat. $<$ 40 years old		
Age: 40-59 years old	0.1339**	0.0569
	(0.0635)	(0.0441)
Age: $\geq$ 60 years old	0.1813	0.2188
	(0.1771)	(0.1402)
Education: Ref. cat. Vocational training		
Education: No high school degree	0.0339	0.0474
	(0.0616)	(0.0614)
Education: High school degree	-0.0767	-0.1361**
	(0.0917)	(0.0680)
In relationship	0.0907	-0.0095
	(0.0851)	(0.0658)
Children	0.3069**	0.1666
	(0.1412)	(0.1099)
Native language: German	0.1370*	0.0016
	(0.0716)	(0.0504)
Household controls	Yes	Yes
Economic preferences	Yes	Yes
Advisors' own investment	Yes	Yes
Attitudes towards the program	Yes	Yes
Peer score	Yes	Yes
Fixed effects	Yes	Yes
Fit statistics		
Observations	5,747	5,747
R^2	0.43006	0.42028

Notes: Standard errors in parentheses are clustered by advisor. Household characteristics include the number of inhabitants, electricity price per kWh, old refrigerator age, estimated savings from refrigerator replacement and the governmental transfer type received. Fixed effects account for ZIP code, year-month, branch-year and co-advisor. p-values ***: 0.01, **: 0.05, *: 0.1.

Economic preferences

We elicit advisors' preferences regarding their willingness to take risks, to give to good causes without expecting anything in return and to forgo something that is beneficial to them today to profit more in the future on ten-point Likert scales (Falk et al., 2022). Estimation results are presented in Table 4.10. We standardize all values by subtracting the mean and dividing by the standard deviation.

Table 4.10: Economic preferences

Dependent Variable:	Voucher request	Refrigerator replacement
Model:	(1)	(2)
Altruism	0.0002	-0.0325**
	(0.0156)	(0.0123)
Risk	0.0053	0.0160
	(0.0168)	(0.0152)
Patience	0.0369**	0.0513***
	(0.0169)	(0.0111)
Household controls	Yes	Yes
Program controls	Yes	Yes
Socio-economic characteristics	Yes	Yes
Advisors' own investment	Yes	Yes
Attitudes towards the program	Yes	Yes
Peer score	Yes	Yes
Fixed effects	Yes	Yes
Fit statistics		
Observations	5,747	5,747
R^2	0.43006	0.42028

Notes: Standard errors in parentheses are clustered by advisor. Household characteristics include the number of inhabitants, electricity price per kWh, old refrigerator age, estimated savings from refrigerator replacement and the governmental transfer type received. Fixed effects account for ZIP code, year-month, branch-year and co-advisor. p-values ***: 0.01, **: 0.05, *: 0.1.

For the voucher request decision of households, we do not find altruism and risk to play a role. However, patience shows to be positively related to households' voucher request decisions. An increase in advisors' reported patience by one standard deviation increases household request rates by 4 percentage points. For refrigerator replacement the effect is slightly higher at 5 percentage points. Moreover, for replacement decisions altruism shows to be a negative determinant: for a one standard deviation increase in reported altruism, replacement rates decrease by 3 percentage points. Risk preferences of the advisor do not play a role in the second-stage decision of households.

Advisors' own investment

We ask advisors about the age of their refrigerator to elicit how long it has been since they replaced their own appliance. We group answers into three categories: replacement less than 11 years ago, replacement 11 years or longer ago and not specified. The cutoff at 11 years is aligned with the age criterion for eligibility to the Refrigerator Replacement Program. Estimation results are shown in Table 4.11.

Table 4.11: Avisors' own replacement

Dependent Variable:	Voucher request	Refrigerator replacement
Model:	(1)	(2)
Ref. cat. Replacement $\leq$ 10 years ago		
Replacement > 10 years ago	-0.3821***	-0.2594***
	(0.1131)	(0.0839)
Replacement timing not reported	-0.0224	0.3008
	(0.2148)	(0.2017)
Household controls	Yes	Yes
Program controls	Yes	Yes
Socio-economic characteristics	Yes	Yes
Economic preferences	Yes	Yes
Attitudes towards the program	Yes	Yes
Peer score	Yes	Yes
Fixed effects	Yes	Yes
Fit statistics		
Observations	5,747	5,747
R^2	0.43006	0.42028

Notes: Standard errors in parentheses are clustered by advisor. Household characteristics include the number of inhabitants, electricity price per kWh, old refrigerator age, estimated savings from refrigerator replacement and the governmental transfer type received. Fixed effects account for ZIP code, year-month, branch-year and co-advisor. p-values ***: 0.01, **: 0.05, *: 0.1.

We find that households advised by advisors that own refrigerators that were replaced recently – less than 11 years ago – are significantly more likely to request a voucher relative to households advised by advisors that own a refrigerator that is older than 10 years, by around 38 percentage points. The effect carries over to the second-stage decision to replace a refrigerator: households that are visited by advisors with newer refrigerators are 26 percentage points more likely to replace their appliance. Advisors that do not report their replacement timing do not differ in neither voucher request nor replacement rates from advisors with recent replacements.

Attitudes toward the program

We elicit advisors' attitudes toward the program. We ask for their agreement with five statements about several aspects of the replacement scheme on a five-point scale ranging from strong disagreement (-2) to strong agreement (2). First, we ask for advisors' general agreement about the program being a good idea. Second, we state that the right households profit from the replacement program. The third and fourth statement declare that the program helps households save electricity and contribute to protecting the climate, respectively. Fifth, we state that for participating households the replacement voucher is worth the effort of requesting and redeeming it.

Table 4.12: Attitudes towards the program

Dependent Variable:	Voucher request	Refrigerator replacement
Model:	(1)	
Program is a good idea	-0.0438	-0.0272
	(0.0548)	(0.0469)
Right households profit	0.0100	-0.0177
	(0.0557)	(0.0395)
Households save energy	-0.0421	-0.0022
	(0.0756)	(0.0599)
Households protect climate	0.0985**	0.1432***
	(0.0484)	(0.0238)
Voucher worth the effort	-0.0661	-0.0967**
	(0.0526)	(0.0442)
Household controls	Yes	Yes
Program controls	Yes	Yes
Socio-economic characteristics	Yes	Yes
Economic preferences	Yes	Yes
Advisors' own replacement	Yes	Yes
Peer score	Yes	Yes
Fixed effects	Yes	Yes
Fit statistics		
Observations	5,747	5,747
R^2	0.43006	0.42028

Notes: Standard errors in parentheses are clustered by advisor. Household characteristics include the number of inhabitants, electricity price per kWh, old refrigerator age, estimated savings from refrigerator replacement and the governmental transfer type received. Fixed effects account for ZIP code, year-month, branch-year and co-advisor. p-values ***: 0.01, **: 0.05, *: 0.1.

Table 4.12 presents the results. Advisors' opinions about the program being a good idea, whether the right households profit from it and whether it helps them to save energy do not significantly affect neither voucher request nor replacement decisions of advised households. However, advisors' opinions regarding enabling households to contribute to climate protection are significantly correlated with households investment choices. In the first-stage decision, households visited by advisors with a by one point higher agreement have by 10 percentage points higher voucher request rates. In the second-stage decision, replacement rates are by 14 percentage points higher. Moreover, opinions regarding the voucher value being worth the administrative effort for households are significant negatively associated with replacement rates. A by one point higher agreement of advisors on the scale reduces replacement rates by 10 percentage points. The first stage decision is not significantly affected; however, the sign of the coefficient is negative as well.

4.5.3 Peer effects on household investment decisions

Following the large literature on peer effects in Photovoltaic adoption, we check whether demographic similarities between households and advisors affect investment decisions. For this purpose, we make use of three demographic characteristics that we observe in both advisors and households (and which we use for testing for random assignment in Section 4.4): age, migration background and being a parent (see Section 4.4 for how we define and measure these characteristics for households and advisors).

Advisors learn about all three characteristics of households before or during administering the audit. Households likely learn about their advisors' age and migration background as these characteristics are often visible, audible or in the case of the migration background may be recognized by the surname. Households do not know whether visiting advisors have children if advisors do not tell them; so this information may remain asymmetric. For the analysis of peer effects, we construct a peer score for each advisor-household combination. We sum up the number of matches between advisors and households in terms of sharing demographic characteristics. Consequently, the peer score can take values between 0 for no shared characteristics to 3 for households and advisors matching in age, migration and parental status.[17] In practice, we do not observe an advisor-household match of all three characteristics. Two household-advisor combinations share two characteristics, 703 share one characteristic.

Table 4.13 shows estimation results for regressing the investment outcomes on the peer score in columns (1) and (5) as well as on indicators for matches based on migration in columns (2) and (6), on age in columns (3) and (7), and on being a parent in columns

[17]Note that we do not observe all advisor-household matches for the migration characteristic, as we only identify this via the households' transfer type. That means that we identify the subsample of asylum-seeking households, but we do not know about the migration background of other households.

Table 4.13: Peer score

Dependent Variable:	Voucher request				Refrigerator replacement			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Peer score	0.0668**				0.0630**			
	(0.0324)				(0.0239)			
Peers: Migration		0.0202				0.1894		
		(0.1900)				(0.1724)		
Peers: Age			0.1409**				0.1231**	
			(0.0534)				(0.0477)	
Peers: Parental status				0.0582				0.0675*
				(0.0482)				(0.0361)
Household controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Program controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-economic characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic preferences	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Advisors' own investment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Attitudes towards the program	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Observations	5,747	5,757	5,740	5,739	5,747	5,757	5,740	5,739
R^2	0.43006	0.42988	0.43061	0.43008	0.42028	0.41944	0.42085	0.42046

Notes: Standard errors in parentheses are clustered by advisor. Household characteristics include the number of inhabitants, electricity price per kWh, old refrigerator age, estimated savings from refrigerator replacement and the governmental transfer type received. Fixed effects account for ZIP code, year-month, branch-year and co-advisor. p-values ***: 0.01, **: 0.05, *: 0.1.

(4) and (8). Aggregating over all characteristics using the peer score, we find a positive peer effect for both the voucher request and the refrigerator replacement decision of households. An additional shared characteristic increases voucher request rates by 7 percentage points and replacement rates by 6 percentage points. Looking at the peer effect of each characteristic separately, we find a significant positive effect on voucher request rates for age at 14 percentage points. The coefficients for migration and children are positive but small and insignificant. In the second-stage decision, the match of older advisors with older households increases replacement rates by 12 percentage points and the match of advisors and households with children increases replacement by 7 percentage points. Sharing the migration background does not significantly affect replacement rates; the coefficient is large positive but imprecisely estimated.

4.6 Discussion

Socio-demographic characteristics. While the contract type of advisors in the program does not impact household replacement rates, the timing when advisors joined the program has a significant impact. Both advisors that have been active in the program for the longest and shortest time, have lower household voucher request and investment rates than advisors who joined the program between 2013 and 2018. As we do not find a significant effect of audit experience on outcomes (see Section 4.4), the result cannot be explained by learning.

We find surprisingly little impact of socio-demographic characteristics on households' investment choices. Advisors with children and middle-aged advisors have significantly higher voucher request rates in the households that they visit. Advisors with at least an *Abitur* degree are less successful than advisors with lower educational attainment. The latter result could relate to a generally lower educational level in the population of advised households. Similar to the detected peer effects, advisors with a level of education whose educational attainment is closer to the average advisee may be more successful in reaching out to households or are more able to emphasize with their clients.

Economic preferences. We find advisors that report greater pro-sociality to be less successful in convincing households to replace their refrigerators. The finding is unexpected as a plausible hypothesis is that more altruistic advisors would increase investment rates via a higher intrinsic motivation and, consequently, increased efforts to help advised households. An alternative explanation could be that altruistic advisors adopt softer, more empathetic styles of advising and communicating, prioritizing households' comfort over persuasion to invest. Another hypothesis is that altruistic advisors attempt to convince households based on pro-social values, which may be less important to the average household than monetary savings. Higher self-reported patience of advisors increases voucher request and investment rates of advised households. The finding is intuitive as patient advisors likely take more time to thoroughly explain the benefits of appliance replacement, answer any open questions regarding the administrative process of redeeming the voucher, and provide otherwise helpful assistance to households.

Advisors' own investment. We find a positive effect of advisors' own investment decisions on households' voucher request and replacement rates. If advisors have replaced their refrigerator more recently, advised households are significantly more likely both to request a voucher and to replace their refrigerator. This result is in line with findings by Kraft-Todd et al. (2018) who find that community organizers are more likely to convince residents to install PV panels if they themselves have installed panels and argue that beliefs are more convincing if underlined by actions rather than by words alone.

Attitudes towards the program. We find some of advisors' attitudes to matter while others do not play a role for household investment choices. The opinion of advisors that the program helps households to contribute to climate protection is positively associated with both voucher request and appliance replacement rates. This could either matter via an increase in the intrinsic motivation of advisors to raise efforts to convince households of the benefits of a replacement or advisors might directly use climate protection as a persuading reason to convince households if those are environmentally conscious. The opinion of advisors that the voucher value is worth the administrative hassle associated with requesting and redeeming it is negatively associated with households' replacement rates. Advisors who hold this opinion may not put enough effort into convincing households of the benefits of replacement as they underestimate the perceived hassle cost for households. Alternatively, they may simply not comprehend the difficulties some households have dealing with the administrative procedures and, consequently, do not provide appropriate administrative assistance.

Peer effects. We find a positive peer effect for households' voucher request and replacement decisions. For voucher request rates, the effect is clearly driven by advisor-household matching in age. For replacement rates, the coefficient is strongest for migration matches, but insignificant, while the coefficients for matches based on age and parental status are smaller but significant.

Our findings are in line with (i) the literature on the presence of peer effects in energy investment and consumption decisions, (ii) the literature on the presence of stronger peer effects for demographically close peers on various economic outcomes, and (iii) the literature on peer effects in the adviser-advisee relationship. (i) Our findings resonate with earlier findings of a positive effect of neighbors' solar panel installations on an individual's installation decision (Bollinger and Gillingham, 2012; Graziano and Gillingham, 2015) and a positive effect of neighbors' water conservation on an individual's water conservation (Bollinger et al., 2018). (ii) Our findings also connect to studies that look at the effects of demographically close peers on other economic outcomes, such as practicing the same religion on contraception decisions (Munshi and Myaux, 2006), speaking the same language in a foreign country on welfare use (Bertrand et al., 2000), belonging to the same ethnic group on the choice of saving vehicle (Mugerman et al., 2014), or being demographically close in terms of age, sex, city of residence and income on the choice of the health care plan (Sorensen, 2006).

In our setting, the relationship between energy advisors and households is different from a horizontal relationship between neighbors or individuals that share demographics and that interact on the same level. Rather, advisors are experts that consult households and whose influence in that sense is more direct and deliberate than between neighbors. To our knowledge, (iii) literature on peer effects in the vertical adviser-advisee relationship is scarce. Potentially most closely linked are findings on the constellation between advisor and advisee on following financial advice that the likelihood for male advisees is affected by the age and gender of the advisor and the likelihood for female advisees is affected by marital and parental status of the advisor (Stolper and Walter, 2019). However, our context may differ from financial advising. The program does not provide commission-based incentives for successful replacement by advised households. In this setting, advisers are intrinsically motivated to convince households to replace their inefficient appliances, which is in the best interest of households.

Summary. We find advisors to matter for households' investment decisions. Advisors are heterogeneous in their performance when it comes to advising households, offering support and convincing them about the benefits of appliance replacement. We find various determinants that can explain part of this heterogeneity. Advisors' own investment in an efficient appliance, their economic preferences and attitudes toward the program all affect households' investment decisions. Interestingly, socio-demographic characteristics do not play a role in households' decision, expect via peer effects in advisor-household matches when sharing common characteristics.

4.7 Conclusion

Energy efficiency assistance programs are a widely used policy tool to support low-income households in times of high energy prices. However, uptake among eligible households remains disappointingly low, limiting the effectiveness of such initiatives. One underexplored avenue for improving program performance lies in optimizing the human factor—selecting and training frontline staff who effectively engage with program beneficiaries. This paper studies determinants of successful frontline staff in the context of an energy efficiency assistance program in Germany. We take a systematic look at the impact of energy advisors on households' decisions to replace their old inefficient refrigerators, identifying determinants of advisors' success.

We find that advisors' economic preferences and attitudes toward the program significantly affect households' investment decisions. Moreover, advisors who recently replaced their own refrigerators are more successful in convincing households to do the same. Surprisingly, most advisor demographic characteristics, such as gender or age, do not matter for household decision-making. However, we document demographic peer effects: similarities between advisors and households —such as a shared migration background, age, or being a parent— affect households' likelihood of replacing their refrigerators.

Our results are in line with prior research on the effects of (demographic) similarities and investment behavior of frontline staff on household decision-making. We make three contributions to literature. First, we systematically analyze the role of energy advisors in assistance programs, identifying key characteristics that drive household investment decisions. Second, we extend the literature on peer effects documenting vertical effects between benevolent advisors and advisees with shared traits. Third, we contribute to the literature on the energy efficiency gap, shifting the focus from the influence of household preferences and biases to external actors' —advisors— effects on household investment behavior.

Our findings offer valuable insights for policy-makers and managers of assistance programs seeking to improve program performance through staffing decisions. While characteristics commonly found on a CV may not be used as a basis for optimizing hiring decisions, the elicitation of preferences and attitudes may be integrated in hiring processes, and training on the job can focus on correcting beliefs that are potentially destructive for advisor-advisee interactions. A further margin of improvement that can be implemented with the staff employed is targeted matching between frontline staff and beneficiaries based on demographic similarities that could lead to improved interactions.

Chapter 5

Affording to pay attention? Energy cost in low-income households' investment decisions

Abstract: Purchase decisions for energy-using durables provide a promising setting to study the role of behavioral biases and financial constraints in economic decision making of low-income consumers. While lifetime energy cost of durables is significant, it is not salient at the time of purchase. Moreover, financially constrained consumers may not be able to afford more efficient durables. This paper studies whether low-income households consider energy cost when purchasing household appliances and the role of financial constraints and inattention to energy efficiency. To infer consideration of energy cost, I compare the demand responsiveness to energy cost and purchase prices. The analysis uses more than 20,000 purchases from lowest-income consumers in Germany and temporal and spatial variation in electricity rates, purchase prices and subsidies. Exogenous changes in subsidies and the design of the EU Energy Label provide variation to study the role of financial constraints and inattention to energy efficiency. Lowest-income consumers do not rationally consider energy cost under conservative assumptions. An important factor is inattention, while financial constraints do not matter. The findings are important for designing effective and equitable energy policy instruments.

Acknowledgments: I thank Erica Myers, Marten Ovaere, Mateus Souza and Jakob Wegmann for valuable comments. This paper also benefited from feedback by participants of the LSE Environmental Week, SURED Conference, EAERE Conference, Mannheim Conference on Energy and the Environment, Annual Conference of the German Economic Association (VfS), as well as at seminars at ZEW Mannheim, University of Calgary, ETH Zurich, University of Gothenburg and ifo Fürth. I thank Max Stellbrink for outstanding research assistance.

5.1 Introduction

Individuals have been observed to discount future payments at much higher than market interest rates. Some studies find discount rates in the range around 100 percent and even higher (e.g., Warner and Pleeter, 2001; Harrison et al., 2002; Ifcher and Zarghamee, 2011). Moreover, consumers sometimes partially or completely neglect non-salient payments in their purchase decisions, such as the shipping fee in online purchases (Hossain and Morgan, 2006) or the VAT when shopping groceries (Chetty et al., 2009). Behavioral biases in economic decision making, such as inflated discounting and neglecting non-salient payments, can lead to market distortions and affect consumer welfare. Low-income consumers are particularly vulnerable to biases, for instance, due to lower financial literacy (e.g., Lusardi and Mitchell, 2014) and financial stress occupying cognitive capacities (e.g., Mani et al., 2013). At the same time, low-income consumers' decisions are more restricted by financial constraints. Heterogeneities along the income distribution in decision making matter as distributional outcomes can be adversely affected. Addressing these issues requires knowledge about the financial and behavioral mechanisms driving the heterogeneities.

The context of purchase decisions for energy-using durables provides a promising setting to study the role of behavioral biases and financial constraints in decisions of low-income consumers. The lifetime energy cost of durables accounts for a substantial share in total costs of purchase and operation, but benefits from investing in higher efficiency occur in the future and are not salient at the time of purchase. At the same time, financially constrained consumers may not be able to afford efficient durables with low energy cost. If low-income consumers do not take into account energy cost in purchase decisions – or do so less than wealthier consumers, this can lead to market distortions and affects the distributional consequences of carbon taxes, and it can reinforce a cycle of poverty if the burden from high energy bills inhibits future investments.

This paper investigates whether households in the lower tail of the income distribution take into account lifetime energy cost when purchasing refrigeration appliances, and the role of two candidate mechanisms that may lower their consideration of energy cost: financial constraints, in the form of the budget available for appliance purchases, and behavioral constraints, in the form of inattention to energy efficiency. First, in a theoretical framework, I derive hypotheses of how financial constraints and inattention to energy efficiency affect the trade-off between the purchase price and discounted lifetime energy cost. Second, I estimate demand responses to changes in energy cost and purchase prices in refrigerator purchases by consumers from the 7 percent lowest-income households in Germany. I infer consideration for energy cost by comparing the relative responsiveness to both cost components, calculating a trade-off ratio. The estimation employs variation in electricity rates over time and across regions and variation in net purchase prices stemming from changes in prices over time and in subsidies over time and across states. The data consists of 20,601 refrigerator purchases by lowest-income households from Germany's largest energy efficiency assistance program.[1] Third, I investigate the role of financial constraints and inattention to energy efficiency for consideration of energy cost, testing the hypotheses derived in the theoretical framework. To study the role of financial constraints, I exploit the effect of exogenous changes in the federal subsidy and subsidy top-up programs by federal states within the assistance program on demand responses. I estimate a comparison of means, two-way fixed effects specifications and Difference-in-Difference designs. To investigate the role of inattention to energy efficiency, I use the revision of the EU Energy Label and associated adjustments of the assistance program's minimum efficiency requirements that made differences in efficiency levels between energy classes more salient to households. I compare the relative responsiveness to both cost components before and after to infer how consideration of energy cost is affected by increased salience of energy efficiency.

I find that households do not fully consider energy cost, even under conservative assumptions for discount rate and appliance lifetime that provide a strict test of under-consideration of energy cost.[2] I find that financial constraints do not matter in this setting. However, an increased salience of appliance energy efficiency after the EU Energy Label revision significantly increases households' consideration of energy cost.

In previous studies, consideration of lifetime energy cost in investment decisions has been found to differ across the income distribution. For the general population, most recent valuation studies have found (close to) full consideration of lifetime energy costs when investing in cars (e.g., Sallee et al., 2016; Allcott and Wozny, 2014; Busse et al., 2013), household appliances (Houde and Myers, 2021) and heating systems (Myers, 2019; Sahari, 2019).[3] Earlier work on implicit discount rates for investments in energy-using durables finds that discount rates that rationalize investments by low-income consumers tend to be much higher than by wealthier consumers (Hausman, 1979; Train, 1985).[4] This finding is in line with more recent empirical evidence that suggests that

[1]The sample mainly consists of households receiving unemployment benefits or a housing allowance. Households in the lowest seven percent of the income distribution in Germany are eligible to register for the program (according to eligibility criteria and own calculations based on numbers from Bundesagentur für Arbeit (Bundesagentur für Arbeit)).

[2]I assume an appliance life of 10 years, expectations for electricity prices to remain constant and a discount rate of 20 percent and provide sensitivity checks.

[3]To my knowledge, Gillingham et al. (2021) and Leard (2018) are the only recent studies that find less than full consideration of lifetime energy cost in the context of car purchases. Most of these studies report the relationship between purchase price and lifetime energy cost as how much consumers value the latter. I use the more neutral terms "trade-off" and "consideration" which do not rule out that consumers value lifetime energy cost but do not purchase according to their preferences, for instance, due to financial constraints.

[4]The concept of implicit discount rates is closely connected to valuation and the consideration of

low-income households are less likely to upgrade energy-using durables with a more efficient model (Schleich, 2019; Ameli and Brandt, 2015; Mills and Schleich, 2012). At the intensive margin of purchase, lower-income consumers tend to value lifetime energy cost slightly less. Houde and Myers (2021) find full valuation of energy cost in the lower and middle tercile of the income distribution while consumers in the upper tercile slightly overvalue energy cost.

At least three rationales could explain why low-, and particularly lowest-income consumers' investment decisions differ from those in the general population. First, these consumers are more liquidity- and budget-constrained and may not be able to afford more energy-efficient durables with lower lifetime energy cost.[5] Second, financial literacy, specifically regarding energy decisions, tends to be lower in low-income consumers (Brent and Ward, 2018; Blasch et al., 2021). Third, poverty may affect economic decision-making via its impact on time and risk preferences. Constant financial worries can absorb cognitive capacities and steer attention of individuals who are scarce in resources such as money or time away from long-term problems (Shah et al., 2012; Mani et al., 2013; Haushofer and Fehr, 2014). Investment decisions in energy-using durables, which cover long time horizons, made by individuals in poverty may be affected via this channel.

In addition to having a harder time getting the investment decision right, there is also more at stake for low-income households. This vulnerable group bears a disproportionate burden from high energy bills, as a larger portion of their income is dedicated to energy expenditures. Compounding this issue, their energy usage tends to be less efficient and energy price elasticity is lower compared to the general population, making it difficult to respond to price hikes (Frondel et al., 2019; Andor et al., 2021). As a result, it is crucial for low-income households to get the investment decision right – choosing the optimal level of energy efficiency – to alleviate the burden from high energy bills. Failure to do so bears the risk of a reinforcing cycle of poverty, where high energy cost keeps adding to monthly energy bills, further reducing the budget for future investments.

My paper contributes to literature in three ways. First, I contribute to the literature on the energy efficiency gap, providing yet lacking evidence on lowest-income consumers. I work with a unique dataset on purchase decisions from the lower tail of the income

operating energy cost. Instead of a parameter that describes the relationship between purchase price and lifetime energy cost, it refers to the discount rate that would rationalize full valuation or a full trade-off. See Schleich et al. (2016) for a comprehensive framework of implicit discount rates.

[5] Credit constraints and the cost of credit do not seem to play a relevant role in explaining the energy efficiency gap in car purchases in the general population in the US (Ankney et al., 2021; Levinson and Sager, 2023). However, credit constraints have been shown to matter for small investments in developing economies (Levine et al., 2018; Berkouwer and Dean, 2022). To my knowledge, there is no evidence on the relevance for low-income consumers in industrialized economies.

distribution in Germany. This group is generally difficult to access for researchers as it is underrepresented in most surveys and response rates are low.[6] I present first evidence on the consideration of lifetime energy cost for lowest-income households in a rich industrialized economy. I document fairly low consideration of energy cost in these households, which is in contrast to full valuation of energy cost measured in the general population in earlier studies for different classes of energy-using durables.

Second, I am the first to empirically study the relevance of financial constraints in affecting the adoption of energy-efficient technologies in low-income households in a rich economy. The role of liquidity constraints has been studied in the context of developing economies like Uganda and Kenya where access to credit is a relevant factor for the adoption of fuel-efficient cook stoves (Levine et al., 2018; Berkouwer and Dean, 2022). In the US general population, credit constraints and the cost of credit do not seem to matter for the consideration of fuel cost in car purchases (Ankney et al., 2021; Levinson and Sager, 2023). I do not find evidence for financial constraints to matter for consideration of energy cost in lowest-income households when purchasing household appliances.

Third, I add new evidence on the role of inattention in investment decisions for energy-using durables. Inattention has been found to matter in some contexts with less than full consideration of energy cost. Information provision on the lifetime energy cost increases willingness-to-pay for efficient fluorescent light bulbs in the US, however, the effect is at least partially driven by updated beliefs rather than inattention (Allcott and Taubinsky, 2015). Contrarily, while a less informative nudge on percentage energy savings increases willingness-to-pay for LED light bulbs, a more informative nudge on the absolute value of energy cost savings decreases willingness-to-pay in an EU population (Rodemeier and Löschel, 2022). In Kenya, inattention appears not relevant in explaining the energy efficiency gap in the adoption of cook stoves (Berkouwer and Dean, 2022). Self-reported inattention in the US general population is heterogeneous and correlates with lower valuation of fuel cost in car purchases (Leard, 2018). Closest related to my study, Houde (2018b) finds that the degree of inattention to lifetime energy cost for purchases of household appliances in the US is heterogeneous, only a share of consumers pays attention, predominantly in the higher-income group. In this setting, coarse binary certification of the Energy Star labels improves attention to efficiency for some consumers, but crowds out attention to more precise information on energy cost for others. I document that inattention to energy efficiency matters in the lower tail of the income distribution and that in the presence of the coarse classification of the EU Energy Label, low-income consumers are attentive to differences in efficiency

[6]Studies in the energy context have reported high drop out rates and low take-up of investment subsidies in the low-income group (Allcott, 2011; Fowlie et al., 2018; Chlond et al., 2025).

between energy classes. However, the data suggests that they are not attentive to more accurate information in energy consumption within class.

Full consideration of lifetime energy cost is crucial for the functioning of Pigouvian price instruments. Carbon taxes only fully internalize the social cost of carbon if households rationally consider energy cost, which carbon taxes are part of, in their purchase decision.[7] Under the assumption that all income groups consider energy cost in their investment decision for energy-using durables, Levinson (2019) finds that carbon taxes are also preferred from a distributional perspective, while energy efficiency standards are more regressive. When heterogeneities in the trade-off are present across the income distribution, carbon taxes imply relative more regressive distributional consequences that increase the burden on low-income households (Houde and Myers, 2019; Houde, 2018a). If consideration of energy cost is low in the lower tail of the income distribution, this group carries a double burden from higher energy and carbon cost. In light of my findings, minimum efficiency performance standards may be the preferred instrument for the affected group, while for the general population that is attentive to energy cost Pigouvian taxes are efficient. To mitigate adverse distributional consequences, policymakers need to address the barriers to energy-efficient investment at the lower end of the income distribution, developing suitable instruments that address financial constraints, inattention, and other factors.

I proceed as follows: In Section 5.2, I outline a theoretical framework of the purchase decision for energy-using durables. Section 5.3 provides the background on the source of my data and gives an overview of the dataset my analysis is based on. In Section 5.4, I provide descriptive evidence for investment choices by energy efficiency level in my sample. Section 5.5 discusses the empirical strategy for the demand response estimation and presents its results. Section 5.6 deals with the empirical strategy and results for the effect of financial constraints on consideration of energy cost, and Section 5.7 does the same for the salience treatment via the revision of the EU Energy Label. Section 5.8 concludes.

5.2 Theoretical framework

The purchase decision for energy-using durables by utility-maximizing consumers can be described by a decision problem in which the consumer chooses between appliance models with different levels of energy efficiency. Models with lower (higher) energy efficiency imply higher (lower) lifetime energy cost and a lower (higher) purchase price. The framework presented here is an adaptation and extension of the model by Allcott

[7]Since 2013, the EU ETS indirectly taxes the carbon content in consumer electricity prices via a carbon price on the generation of electricity from fossil fuels.

and Greenstone (2012). After showing the decision problem without financial constraints and inattention to energy cost as base case, I introduce both factors into the framework and derive hypotheses about how they affect consumers' purchase decisions.

The consumer chooses between appliance models on the market and pays for the initial investment price and the energy cost incurring over the durable lifetime. Appliance models in her choice set, denoted by model index m, are continuously distributed in their energy efficiency level over $m \in [0,1]$. Discounted lifetime energy cost e of model m is given by $e_m = e_0 - \Delta e * m$, where e_0 is the lifetime energy cost of the least efficient model $m = 0$ in the choice set, and e_1 the energy cost of the model with the highest efficiency on the market $m = 1$. $\Delta e = e_1 - e_0$ represents the upper bound of reduction in lifetime energy cost relative to the baseline. $\Delta e * m = e_0 - e_m$ then gives the reduction in energy cost as efficiency increases in m. The choice set available on the market is given by purchase price p and lifetime energy cost e for model m, providing curve $e_m(p_m)$. For the purpose of this framework, $e'_m < 0$ and $e''_m < 0$ (see Figure 5.1 for illustration).

In the basic case, the consumer's willingness-to-pay for model m compared to the baseline model $m = 0$ is $WTP_m \leq e_0 - e_m = \Delta e * m$, less or equal to the savings in lifetime energy cost relative to the baseline model. The baseline model m_0 can also be interpreted as the choice of remaining in status quo – the consumer choosing to keep her old appliance and to not purchase a more efficient model. A rational consumer will choose model m if the energy cost savings WTP_m outweigh additional investment cost $\Delta p_m = p_m - p_0$, the difference in the purchase price p between model m and the baseline $m = 0$. The purchase price p is increasing in m; models with higher efficiency are more expensive. The decision criterion is given by:

$$WTP_m \geq p_m - p_0 \quad \Leftrightarrow \quad \Delta e * m \geq \Delta p_m$$

m^* is the threshold efficiency level m where energy savings are equal to the efficiency premium in the purchase price:

$$m^* \geq \frac{\Delta p_m}{\Delta e}$$

m^* is the point at which $e_m(p_m)$ has a slope equal to -1: at this point, an increase of one unit in the purchase price p_{m^*} can be rationalized by an increase of one unit in savings in lifetime energy cost $\Delta e * m^*$. If the consumer takes a rational purchase decision in absence of financial constraints and inattention, she trades off an increase of €1 in the purchase price against a decrease of €1 or more in savings in discounted lifetime energy cost, so that purchasing a model $m \geq m^*$ is rational. Figure 5.1 provides a graphical illustration. The solid red curve plots the distribution of appliance models $m \in [0,1]$

available in the choice set as a combination of purchase price and lifetime energy cost, $(p_m,\ e_m)$. The point $(p^*_m,\ e^*_m)$ is located where the curve's slope is -1 (dashed black line).

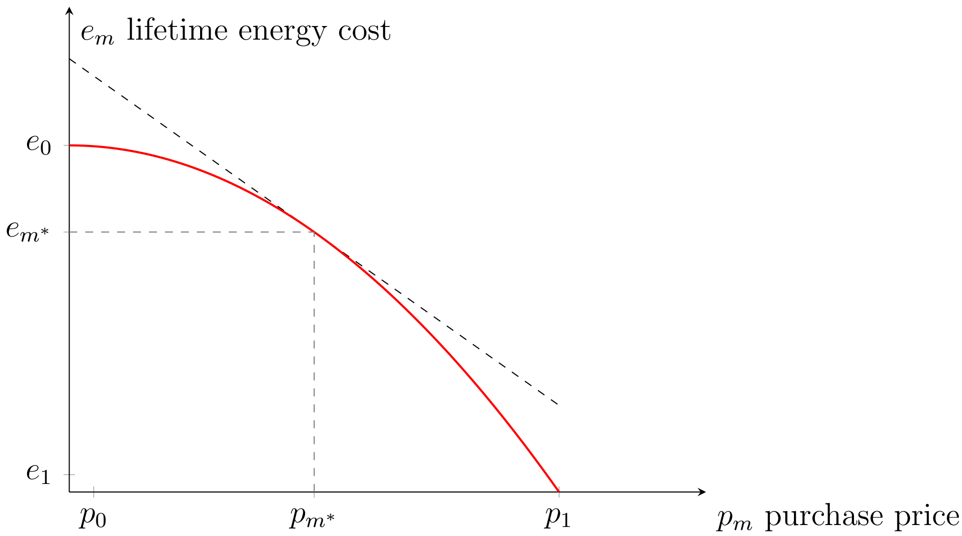


Figure 5.1: Optimal efficiency threshold m^*

Notes: This figure illustrates the location of the optimal efficiency threshold m^*. The purchase price p_m on the x axis increases in m from 0 to 1. The lifetime energy cost on the y axis decreases in m from 1 to 0. The red curve plots the distribution of appliance models $m \in [0, 1]$ in the choice set. The optimal efficiency threshold in the absence of financial constraints and inattention is the point on the curve where the slope equals -1, $(p^*_m,\ e^*_m)$.

The consumer's decision is constrained by budget if her available income after deducting regular expenses is less than p_m and she does not have access to the credit market. The consumer can afford model m if $p_m \leq Y - exp$, Y denoting available income and exp regular expenses. If the condition is not met for higher m, she cannot afford model $m \geq m^*$ over the optimal efficiency threshold and is forced to purchase $m' < m^*$, which satisfies $p_{m'} - p_0 = \Delta p_{m'} = Y - exp$. Lifetime energy cost $e_{m'}$ will be higher and efficiency level m' will be lower compared to the optimal model $m \geq m^*$ chosen when unconstrained by budget or with access to credit.

$$m' < m^* \quad \Leftrightarrow \quad \frac{\Delta p_{m'}}{\Delta e} < \frac{\Delta p_{m^*}}{\Delta e}$$

Moving further to the left in the choice set on $e_m(p_m)$ to m' will give a slope closer to 0 (see Figure A5.1 in the Appendix). For $m < m^*$, the consumer trades off an increase of €1 in the purchase price against a decrease of less than €1 in savings in discounted lifetime energy cost.

If a subsidy S is offered to reduce the initial purchase price, the consumer can afford a model of higher efficiency m^S, where $m^S > m'$:

$$m^S = \frac{\Delta p_m - S}{\Delta e}$$

In this case, the consumer's trade-off moves closer to €1 in the purchase price versus €1 in energy cost savings.

If the consumer is inattentive to lifetime energy cost, she does not take into account the full energy cost savings from efficiency gains when she makes her purchase decision. Let γ represent the degree of inattention, with $0 < \gamma < 1$. This reduces the perceived willingness-to-pay and the decision criterion is $\gamma \, \Delta e * m \geq \Delta p_m$. The efficiency level chosen under inattention m^γ is then:

$$m^* > m^\gamma = \frac{\Delta p_m}{\gamma \Delta e}$$

In the presence of inattention, as for a binding budget constraint, the trade-off is €1 in the purchase price versus less than €1 in energy savings.

Introducing a subsidy can correct for the diversion from the optimal efficiency threshold introduced by both the binding budget constraint and presence of inattention. This connects to the concept of the Internality Dividend from Externality Taxes (Allcott et al., 2014), suggesting that behavioral biases in investment decisions for energy-using durables can be corrected using Pigouvian subsidies.

5.3 Data

5.3.1 Program description

The empirical analysis uses data on purchase decisions for refrigeration appliances made by low-income households between 2013 and 2023. These decisions were made as part of Germany's largest energy efficiency assistance program, *Stromspar-Check* (translates to Electricity Saving-Check).

Stromspar-Check (SSC) aims to help households receiving government assistance reduce their energy expenses. Funded by the Federal Ministry of Economic Affairs and Climate, the program offers two main services: first, a home energy audit offered to all households who register, and, second, a refrigerator replacement program for those registered households that own an old inefficient refrigeration appliance. The audit consists of two visits to the household. During the first visit, energy advisors record total electricity consumption, the marginal rate paid per kWh, and compile an inventory of all electric devices in the household. They collect data on all refrigeration appliances concerning energy consumption, age, volume and type. The data is entered into the program database, which determines eligibility for the replacement program.[8] Eligible households receive a voucher during the second visit of the audit, providing a subsidy of 100 Euro for refrigerator purchases upon redemption (up to March 2019, the subsidy was 150 Euro). Several state governments provide complementary subsidies to increase the total amount.[9]

To redeem the replacement voucher, households must submit a purchase receipt for the new appliance and a recycling certificate for the proper disposal of the old appliance. Before March 2021, the program required new appliances to comply with the A+++ standard, the highest efficiency class on the scale of the EU Energy Label. With the introduction of the new scale on the label, the classes were revised and Stromspar-Check adjusted its minimum efficiency standard to class D on the new scale.[10]

Recruitment of participating households happens predominantly via social assistance agencies. That means that a broad base of households among government benefit recipients is invited to join the program. However, households that are more attentive to energy cost may be more likely to select themselves for participation. Consequently, the analysis would be based on an attentive subsample of households and I would estimate an upper bound of attention to energy cost for this group – a stricter test for the hypothesis that lowest-income households do not rationally consider energy cost.

5.3.2 Dataset on household investment decisions

My data are comprised of the universe of investments made as part of Stromspar-Check's refrigerator replacement program from October 2013 to February 2023.[11] The prepared dataset counts 20,601 observations at the household level, observing each household once with their purchase decision. The data encompasses a range of household demographic as well as energy-related information, audit records, and details on

[8] Eligibility is based on two criteria: (1) The appliance is older than 10 years, and (2) replacing it with an energy-efficient model of the same type and similar volume would save the household at least 200 kWh annually.

[9] Hamburg has offered a top-up subsidy of 100 Euro since September 2010, Saxony-Anhalt of 75 Euro since May 2020, and Berlin of 50 Euro since December 2020. Since July 2016, North Rhine-Westphalia has offered a graded top-up subsidy according to household size: single households receive 50 Euro, 2-person households 100 Euro, and 3-person households 150 Euro. Households with 4 or more persons receive 200 Euro. This same graded scheme was adopted by Berlin in October 2022. A different grading system was introduced by Saarland in May 2020 and by Hamburg in April 2023, single households receiving 50 Euro, 2-person households 100 Euro, 3- to 4-person households 150 Euro, and 5 or more persons receiving 200 Euro on top.

[10] This means that the stringency of the program's minimum efficiency requirements was reduced as I document in Section 5.4.2.

[11] Chlond et al. (2025) use data from the same program to study how changes in the subsidy value as well as changes in the enrolment mode and voucher deadline in the replacement scheme affect the investment rate of participating households.

both the new purchased and old replaced appliances. Household information includes the type of government support received, the household size and age group categories of members, total electricity consumption, the electricity rate per kWh and the NUTS3 region of residence.[12] Audit records contain the dates of the first and second visit, the voucher issuance and redemption dates, the state and the local site administering the audit. Data on the purchase of the new appliance contain the purchase price, type, manufacturer, model, volume and kWh consumption as reported by the manufacturer. Information on the old replaced appliance includes its age, thermostat setting during the audit, and the total number of thermostat levels.

Households in the dataset can be broadly ranked into three groups by their income. The first group, with the lowest income rank, includes all households receiving unemployment insurance ("Arbeitslosengeld II"), benefits for pensioners with pensions below the minimum subsistence level ("Grundsicherung"), benefits to secure the minimum subsistence level where no other transfers apply ("Hilfe zum Lebensunterhalt"), and benefits for asylum seekers ("Asylbewerberleistungsgesetz"). Conditional on the number and age of persons in the household, these transfers are roughly equivalent in the amount of income they provide to the households. In addition, households receiving Arbeitslosengeld II can earn a small additional income as long as, after deductions, total income stays below the minimum income for receiving a housing allowance ("Wohngeld"). The second group, with a relatively higher income, consists of households that earn a low income and receive a housing allowance because they cannot afford their rent ("Wohngeld"), or child supplements because their salary is not high enough to raise their children ("Kinderzuschlag"). The third group, with the relatively highest income rank, comprises households that do not receive any government support but have incomes below the level for attachment. The second and third groups are more heterogeneous in income levels than the first group. Using information on the number of households receiving governmental benefits by Germany's Federal Employment Agency, I calculate that households in my sample are among the seven percent lowest-income households in Germany.

The raw data from the program database includes 409,927 observations after removing duplicates. I keep only observations of households that participate in the replacement program: households that (i) fulfill the criteria for eligibility and (ii) replace their refrigerator and claim the subsidy. Eligibility criteria are participation in the Stromspar-Check energy audit, owning a refrigerator older than 10 years and saving at least 200 kWh when replacing the old appliance with an efficient model of equivalent size and volume. I drop observations with implausible and missing values of for the

[12]Based on the EU NUTS classification, Germany is made up of 400 NUTS3 regions, where each region includes between 150,000 and 800,000 inhabitants.

Table 5.1: Descriptive statistics on investment decisions

Statistic	N	Mean	Std. Dev.	Min	Median	Max
No. persons	20,601	2.747	1.748	1	2	14
Electricity rate per kWh (€)	20,601	0.28	0.027	0.12	0.28	0.90
Age of old appliance (years)	20,601	17.97	4.93	11	17	45
Cons. of new appliance (kWh)	20,601	141.52	34.30	51	149	317
Total subsidy (€)	20,601	170.62	62.68	100	150	350
Purchase price (€)	20,601	489.51	186.05	70.41	474.00	2,439.00

Notes: The table presents summary statistics for the final sample from the Stromspar-Check database.

analysis relevant variables. Additionally, I exclude purchases of appliance types "chest freezer" and "upright freezer".[13] This leaves me with a final sample of 20,601 observations. To correct for typographical errors in appliance model information, I align energy consumption of new appliances with the mode of consumption by model and year. Moreover, I code observations with implausible volume, price or consumption as missing. To calculate the total subsidy amount each household receives, I sum the federal subsidy and any state-specific top-up subsidies available at the date of voucher issuance.[14] Table 5.1 provides descriptive statistics for the sample and variables used in the main analysis.

5.3.3 Aggregation of investment decisions

Using the final sample of 20,601 purchase decisions, I construct the regional market-level dataset used for the demand response estimation. Investment decisions are aggregated at the model-quarter-year-NUTS3 region level. I create a dataset with all possible combinations of model, quarter-year and NUTS3 region and restrict the set of combinations to the period between the first and the last quarter the model was sold in any region in my data – the assumption being that this is the period during which the respective model was available on the national market and therefore part of the choice set of households. I also limit the time series for each model region to range between the first and the last quarter a household was observed in a region as otherwise information on electricity prices and purchases is not available.

The demand response estimation requires the amount of appliances sold by model, av-

[13]I exclude freezers from the sample as energy consumption is more dependent on usage behavior, e.g., regular defrosting or not using the appliance year-round and unplugging it, while refrigerators are typically plugged in 24/7 year-round.

[14]In states and periods where a top-up subsidy is available to participants, it is integrated into the administrative process of claiming the federal voucher. Households incur no additional transaction cost to claim the top-up subsidy.

Table 5.2: Descriptive statistics on aggregated investment decisions by model-quarter-year-region

Statistic	N	Mean	Std. Dev.	Min	Median	Max
Models sold	1,183,132	0.016	0.189	0	0	32
Net purchase price	1,183,132	363.4	196.15	50.0	359.5	2,189.0
Annual energy cost	1,183,132	39.60	11.26	7.41	42.12	121.03
NUTS3 regions	265					
Quarter-year	38					
Model	1,099					

Notes: The table presents summary statistics for the aggregated dataset of investment decisions at the model-quarter-year-region level.

erage electricity rates and average purchase prices. Demand for each appliance model is given by the amount of appliances sold in a region-quarter-year. Aggregation of purchase prices is done at the model-region-quarter-year. For missing values in cells with no purchase observed, I impute with the model-quarter-year price at the national level. Remaining gaps are filled by imputing averages within each model-region time series. To determine the net purchase price, I subtract the federal- and state-level subsidies from the purchase price. For the state-level top-up programs that are graded according to the household size, I use the average number of inhabitants by quarter-year-region.[15] Due to a minimum own contribution of 50 Euro for purchasing an appliance in the program, I set prices that drop below 50 Euro after subtracting subsidies to 50 Euro. For aggregation of the electricity rate, I average over region-quarter-year, imputing missing values for quarter-year-regions for which no household is observed with the averages within each regional time series. I calculate annual energy cost multiplying model-specific annual energy consumption as reported by the manufacturer with the average electricity rate. Finally, I drop observations with missing values for any variables used in the analysis. This approach leaves me with a dataset of 1,183,132 observations. Table 5.2 shows summary statistics of the main variables used in the analysis and a count of distinct regions, quarter-years and models in the data.

5.3.4 Dataset on appliance characteristics

I complement the household investment data with data on the EU refrigerator market supply from the EPREL (European Product Registry for Energy Labelling) Database, established under EU Delegated Regulation 2016/2019 (European Commission, 2019).

[15] For observations with missing information on inhabitants, I use the yearly average of inhabitants in the region and impute any remaining missing values by averaging over each region's time series.

EPREL provides detailed information on household appliances placed on the market in any EU country as of August 1, 2017. The database differentiates between models introduced before March 2021 under EU Regulation 1060/2010 (European Commission, 2010) and those introduced after, under Regulation 2016/2019. There is only a small overlap of models that are registered under both regulations which indicates that manufacturers replaced a large part of their product line at the onset of the new regulation.

After downloading the information on household refrigeration appliances[16] and removing duplicate records, the dataset on models falling under Regulation 2016/2019 contains 37,984 appliances, and the dataset on models falling under Regulation 1060/2010 contains 28,251 appliances. Variables provided include the model name, the manufacturer, start and end date on the market, cooling and freezing capacity, and annual energy consumption. For models registered under Regulation 2016/2019 the data also includes the Energy Efficiency Index (EEI) and the energy class.[17] For earlier records under Regulation 1060/2010, I have to rely on calculation of the 2010 EEI with the information provided in EPREL following the formula in Regulation 1060/2010. The energy class for a model can be determined using the EEI thresholds as provided in the regulation. The formula for calculating the EEI as well as the associated energy class thresholds were changed with Regulation 2016/2019 so that both the EEI and energy classes in data collected under Regulations 2010 and 2019 are not directly comparable. To have a common measure of comparison, I also calculate the EEI in its 2010 version for models registered under Regulation 2016/2019. For a few models with specific features which were introduced as categories with the new directive as well as for model records with missing information, calculation of the 2010 EEI is not feasible.[18]

Only the choice set of models available on the German market is relevant for households in my sample. However, information on the placement country in EPREL is missing for most models. To identify which models included in EPREL are sold on the German market, I use the list of all manufacturer brands households in the SSC data purchased from. I assume if a household purchased a model by a particular manufacturer brand,

[16] The database is accessible online: https://eprel.ec.europa.eu/screen/product/refrigeratingappliances2019. A personal API key can be requested to crawl the data.

[17] Manufacturers are responsible for calculating the EEI using a formula provided in Regulation 2016/2019. The value of the EEI determines the energy class the model is assigned.

[18] For the calculation of the EEI, the formula changed from Regulation 1060/2010 to 2016/2019. In the 2010 Regulation, the formula uses the type of appliance, annual energy consumption in kWh, the storage volume of all compartments in liters and the nominal temperature of each compartment accompanied by volume correction factors for specific features. In the 2019 Regulation, the formula from 2010 is extended to adjusting annual energy consumption in kWh with a load factor and a factor for the extent of noise pollution. Moreover, consideration of different features and different compartment types in the formula is more comprehensive. For models that have compartment types "Pantry" and "** Section", the 2010 EEI and energy class cannot be determined as these categories do not exist in the old calculation.

the full product line of that brand contained in EPREL is sold on the German market. I drop one observation where a manufacturer reported an EEI of 1 under the 2016/2019 Regulation. The procedure leaves me with 15,343 models under Regulation 1060/2010 and 15,915 models under Regulation 2016/2019.

Table 5.3 provides descriptive statistics for the EPREL data under both regulations. On average, models registered under Regulation 2016/2019 are larger in volume and consume slightly more energy. The average 2010 EEI for models under 1060/2010 is slightly higher than the 2010 EEI for models under 2016/2019, indicating energy efficiency as measured by the 2010 EEI appears to have improved marginally.[19] Energy classes under 1060/2010 range from A+++ (highest efficiency/lowest EEI) to G, with 10 percent of models in A+++, 40 percent in A++, and 45 percent in A+. Classes A-G cover only 5 percent of models available on the market. Under 2016/2019, energy classes range from A (highest efficiency/lowest EEI) to G, with around 1 percent of models in classes A and B, 4 percent in C and 9 percent in D. Classes E and F are most populated with 37 percent and 44 percent respectively.

I combine the EPREL data with the program dataset based on the model identifier. Out of the 1,499 models recorded in the SSC database, 30.1 percent can be linked to their EPREL entries, the rest of the model names cannot be connected to one specific model without doubt. That means that out of the 20,601 purchase decisions from the program database, 5,649 can be connected to EPREL models under Regulation 1060/2010 and 2,842 can be connected to EPREL models under Regulation 2016/2019, in total covering 41.2 percent of the total number of purchase decisions in my sample.

[19] Figure A5.2 in the Appendix illustrates the relationship between the 2010 EEI and the 2019 EEI for the EPREL sample unde the 2019 Regulation.

Table 5.3: Descriptive statistics on EPREL data

Statistic	N	Mean	Std. Dev.	Min	Median	Max
Models under 1060/2010						
Net cooling + freezing volume in l	9,970	320.403	117.384	33	312	836
Annual energy consumption in kWh	15,343	221.076	85.383	12	219	676
EEI (1060/2010)	15,293	34.173	8.652	4	33	259
Energy class:						
A+++	2,969	10.0%				
A++	11,841	40.1%				
A+	13,246	44.8%				
A-G	1,506	5.1%				
Models under 2016/2019						
Net cooling + freezing volume in l	11,991	334.753	136.289	28	324	836
Annual energy consumption in kWh	15,915	222.473	81.491	12	218	803
EEI (2016/2019)	15,915	104.925	26.760	21	100	999
EEI (1060/2010)	11,481	33.100	8.313	10	32	93
Energy class:						
A	70	0.4%				
B	122	0.8%				
C	913	5.7%				
D	2,322	14.6%				
E	6,433	40.4%				
F	5,525	34.7%				
G	513	3.2%				

Notes: The table presents summary statistics for data from the EPREL database. The upper panel reports statistics for models registered in the database under Regulation 1060/2010, the lower panel does the same for Regulation 2016/2019. For energy classes, the percentage share of models that belong in each is reported.

5.4 The EU Energy Label in purchase decisions

In this section, I employ the data from the EPREL Database to achieve two objectives: first, identify the energy efficiency segment from which low-income households in my sample purchase, and second, examine the influence of changes in the EU Energy Label on the segment available for purchase to households within the assistance program.

5.4.1 Energy efficiency

To understand how households make choices regarding energy class and efficiency level, I use the sample of appliance purchases by households that could be merged with models in the EPREL Database. Notably, my sample is not representative of the purchasing decisions of lowest-income households since the SSC program voucher limits their choices to appliances in energy classes A to D under the 2019 regulation and to A+++ under the 2010 regulation. This means that I focus on the upper efficiency segment of the market, examining purchase decisions conditional on the program's minimum efficiency requirements.

Figure 5.2 plots the density distribution of the EEI of households' appliance purchases matched with the EPREL 2019 data. The histogram shows the number of purchases by EEI and marks the energy class thresholds in red dashed lines. I find quite pronounced bunching of investments at the lower thresholds of energy classes C and D. The only appliances purchased from classes A and B are also stacked at the respective lower thresholds.[20] For the purchases matched to the EPREL 2010 dataset, bunching at the lower end of the A+++ class is a bit less pronounced but still clearly visible (see Figure A5.3 in the Appendix). The investments appear to be concentrated in a small segment towards the lower end of each energy class just above each respective threshold.

Next, I check whether the observed pattern is solely driven by households sorting at the lower end of each energy class or whether supply in the market is also bunching at the lower ends. I inspect the density distribution of the EEI of models offered on the EU market as documented in EPREL. Systematic bunching at the lower end of each energy class is visible for models placed on the market under both the 2010 and 2019 regulation (see Figure 5.3 for the 2019 EPREL data and Figure A5.4 in the Appendix for the 2010 EPREL data). The finding has been documented before and is as expected since the emphasis of the EU Energy Label on the energy class incentivizes manufacturers to place models on the market that locate just above the threshold to the lower energy class.[21]

[20] Apparently, a few households choose a model that falls in energy class E and does not fulfill the program requirements. In these cases, households appear to have been granted an exception.

[21] In the context of the EU Energy Label, Goeschl (2019) documents manufacturer bunching for

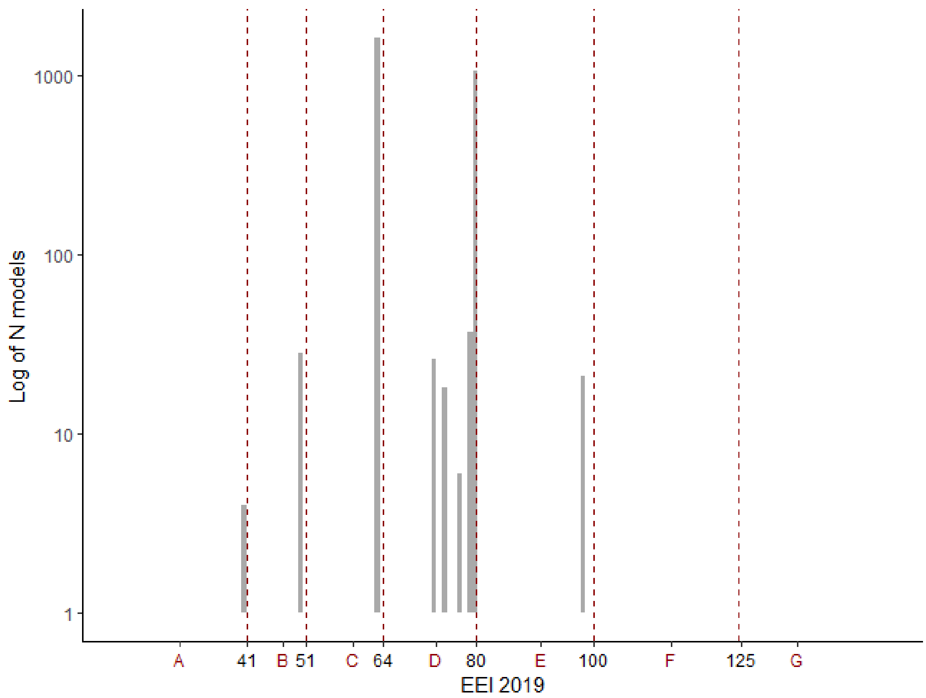


Figure 5.2: Density distribution of household investment by 2019 EEI
Notes: This figure shows the logarithmized distribution of appliance purchases by Energy Efficiency Index (EEI) that could be matched to EPREL 2019 appliances. A lower EEI indicates higher energy efficiency. The red dashed lines mark the EEI thresholds of each energy class of the EU Energy Label.

I examine whether bunching of household demand is proportional – not higher or lower than expected – to the bunching of manufacturers' supply. I compare the proportion of purchases that falls just above each threshold to the lower class with the proportion of models available on the market at the respective threshold. Assuming that, in the absence of EU Energy labelling, households would select continuously across the EEI distribution, the comparison is informative about households' regard of the energy label in their purchase decision. If households disregarded the label in their purchase decision, their choices should be continuously affected by differences in energy consumption across models. However, no discontinuities in their choices should be visible around the thresholds of energy classes. I assume that the choice set of households comprises all models that I observe in the EPREL database.[22]

To check for bunching, I only consider the period in a bandwidth of 2 years around the revision of the EU Energy Label as this is the period of interest in the analyis of the EU Energy Label revision (see Section 5.7). I conduct a two-sided proportion test to check whether the proportion of households' purchases at the lower end of each

refrigeration appliances at the lower end of energy classes and Kesselring (2023) shows both manufacturer and sales bunching for washing machines. For the US Energy Star Label, Houde (2018a) documents manufacturer bunching of refrigeration appliances at the certification threshold.

[22]This assumption concerns (1) how I define appliances sold in Germany (see Section 5.3.4) and (2) whether households observe all appliances on the market. The latter is plausible for online purchases. For on-site retailer stores, the selection of models presented may vary.

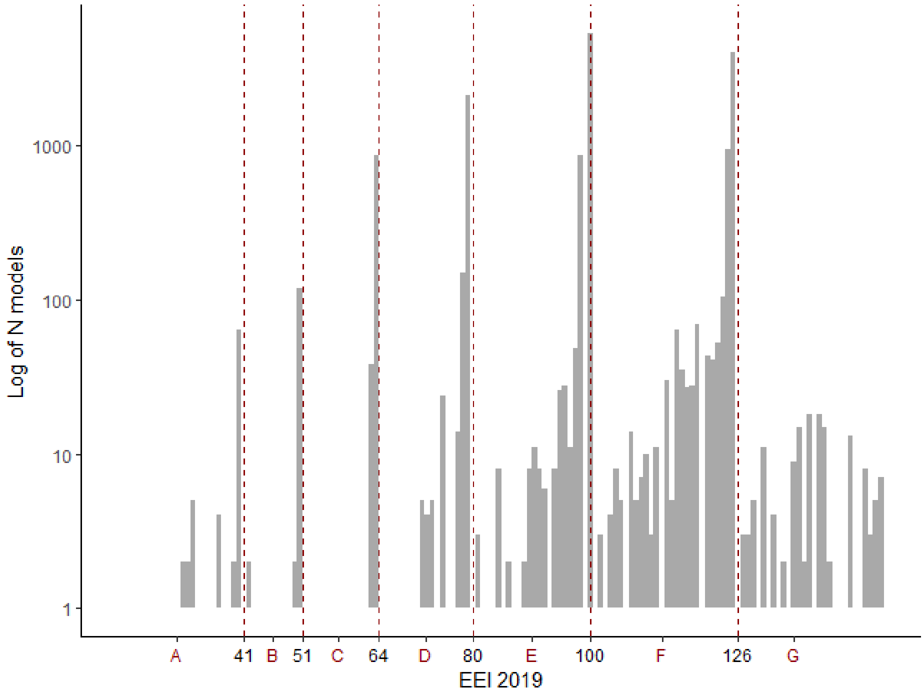


Figure 5.3: Density distribution of models on the market by 2019 EEI

Notes: This figure shows the logarithmized distribution of the models supplied on the market by Energy Efficiency Index (EEI) that were registered under the 1060/2010 Regulation. A lower EEI indicates higher energy efficiency. The red dashed lines mark the thresholds of each energy class of the EU Energy Label. Manufacturers offer more models that are located just above the threshold to the lower energy classes.

Table 5.4: Two-sided proportion test for bunching at energy class thresholds

	Household investment			Market supply			
Energy class	N at lower end	N total	% at lower end	N at lower end	N total	% at lower end	Diff.
Models under 1060/2010							
A+++	490	478	97.6	70	46	65.7	31.9***
Models under 2016/2019							
A	2	2	100	4	4	100	0
B	27	27	100	25	25	100	0
C	1,361	1,361	100	31	31	100	0
D	272	298	91.3	100	128	78.1	13.2***

Notes: The test is conducted for the period 1 March 2019 to 28 February 2023 – two years before and after the rescaling of the EU Energy Label. The lower-efficiency end of each energy class is defined as the interval $(EEI_{low\,threshold} - 1, EEI_{low\,threshold}]$. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

energy class is significantly different from the proportion of model supply in EPREL that is positioned at the lower end of the respective energy class. Table 5.4 presents the results. For models under Regulation 1060/2010, 98 percent of all household purchases are located at the lower end of energy class A+++, while only 66 percent of the market supply is located at the lower end. For models under Regulation 2016/2019, all models offered in classes A to C are located at the lower end. Hence, there is no variation to differentiate between consumer and manufacturer bunching as manufacturer bunching is complete and perfectly overlaps with household bunching. households need to pick a model at the lower end if they went to buy the same energy class. For D, 91 percent of households' purchases fall to the lower end, while only 78 percent of models offered in D are located there. For both A+++ and D, the difference in shares is statistically significant ($p<0.01$), suggesting that households bunch at the lower end beyond the proportion that would be expected according to market supply.

Over-proportional bunching in household purchases suggests that households are attentive to the energy classes shown on the EU Energy Label. Households tend to choose appliances with lower energy efficiency (as proxied by the EEI) within an energy class. However, many buy models in classes above the minimum efficiency requirement under the 2019 regulation, i.e., in classes A to C, despite higher prices of more efficient models. [23] This indicates that households value energy efficiency as provided by the energy class on the label but disregard the more accurate information on annual energy consumption less prominently presented on the label. This pattern is in line with literature on the EU Energy Label that documents consumers tend to value energy classes beyond the implied savings in energy cost (Andor et al., 2020; d'Adda et al., 2022).

5.4.2 Revision of the EU Energy Label

I also utilize the EPREL data to explore how the change in the EU Energy Label in March 2021 altered the choice set available to households in my sample. The rescaling became necessary due to continuous improvements in the energy efficiency of refrigeration appliances. Initially, the European Commission had extended the scale from A to G as set in 1994,[24] adding classes A+ and A++ in 2003 (European Commission, 2003), and A+++ in 2010 (European Commission, 2010), to incentivize manufacturers to continue enhancing the energy efficiency of their products. By 2020, a substantial share of appliances on the market was grouped in the highest class, A+++. To further encourage efficiency improvements and provide consumers with more accurate information, the Energy Label was rescaled, eliminating classes A+++ to A+ for refrigeration appliances.[25]

[23] Using the sample of households' purchases matched to EPREL characteristics and aggregating purchase prices by model, I find a significant relationship between purchase price and EEI. Controlling for manufacturing brand and volume, an increase in the EEI (indicating lower energy efficiency) is associated with a reduction in the purchase price by 2.6 Euro ($p = 0.054$).

[24] The EU Energy Label for household appliances, which provides information on energy consumption and other standard product details, was introduced with EU Directive 92/75/EC. Specific requirements for refrigerators and freezers were established in EU Directive 94/2/EC.

[25] See EU Regulation 2017/1369 and EC Delegated Regulation 2019/2016 for details.

In the SSC assistance program, the EU Energy Label requirement for new appliances was to comply with the A+++ standard before March 2021. With the rescaling, the program had to update its requirements: from March 2021 onward, households needed to purchase an appliance of energy class D or higher to qualify for the voucher subsidy. The formula for calculation of the EEI, which serves to determine the energy class for each model, changed with introduction of the new label scale.[26] As a result, there is no direct mapping between the old and new classes, the ranking of models within the distribution shifted based on included features, and the spread between models increased.

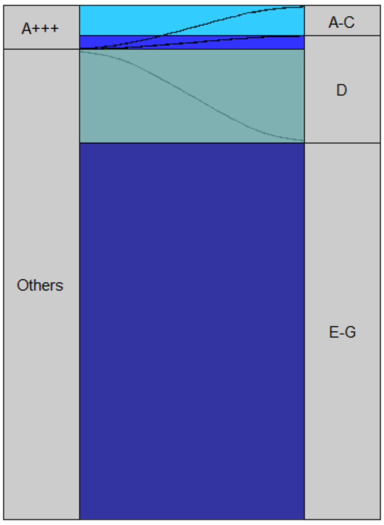


Figure 5.4: Mapping of 2010 energy classes into 2019 energy classes for SSC compliant models
Notes: This figure shows how energy classes under Regulation 1060/2010 map into energy classes under Regulation 2016/2019 for the sample of models on the market March 2021-February 2023 under the latter regulation.

For interpretation of the effect of the label rescaling on households' consideration of lifetime energy cost, I investigate whether the program's requirements made participating households' choice sets stricter or more lenient in terms of energy efficiency. I use models available on the market between March 2021 and February 2023 as documented in the EPREL 2019 data, for which both the 2010 and 2019 energy class and EEI are available, to check (i) how old energy classes (under 1060/2010) of program-compliant models map into new classes (under 2016/2019), and (ii) how the 2010 EEI of compliant models compares before and after the rescaling. To examine the first question, Figure 5.4 and Table 5.5 illustrate how models in A+++, the only class complying with assistance program requirements before rescaling versus others below, are classified in the new program-compliant energy classes A to D versus E to G after rescaling.

[26]The formula for calculating the EEI was revised. See Section 5.3.4 for more information.

The largest share of models sorted into A+++ falls in C after the revision. A few models each are sorted into A and B, and a few more go to D. None of the models falls in classes E to G, i.e., none of the models is compliant before but not after the rescaling. Conversely, the largest proportion of models that were non-compliant before rescaling (below A+++) remains in a class below D. About 20 percent are classified as D. In summary, all models that would have been sorted as compliant in A+++, are also compliant under the new classification. The choice set of households under the new requirements additionally incorporates models from classes A++ and below which become compliant under the new rules. In conclusion, the choice set of households proportionally grew with the adjustment to the new scale.

Table 5.5: Mapping of 2019 energy classes to 2010 energy classes

	A	B	C	D	E-G
A+++	2	4	29	16	0
Others	0	0	0	108	440

Notes: This table shows how energy classes under Regulation 1060/2010 map into energy classes under Regulation 2016/2019 for models on the market March 2021-February 2023 under the latter regulation.

To examine the second question, how energy efficiency of models on the market compares before versus after the label revision, I use the samples of models registered under both regulations 2010 and 2019 that are on the market during the two years before (70 models) and after the revision came into force (166 models) and compare the level of energy efficiency of program-compliant models using the 2010 EEI as a common measure of energy efficiency level. Figure 5.5 shows the distribution of the 2010 EEI for both samples. For models on the market before the label revision (2010 Directive), no model has an EEI below 21 as the cut-off for A+++ models is at 22. The largest mass of the distribution is directly above the cut-off as seen in the bunching analysis. For models on the market after the revision, the distribution spreads over EEI values from 12 to 31. While there are a few more observations above the 75th percentile of the sample under the 2010 regulation, more than 50 percent of the distribution is close to the A+++ cut-off. The revision of the label not only added additional models to the choice set, it also lowered the minimum efficiency cut-off, opening the option for households to buy lower-efficiency models. I conclude that the efficiency requirements of the program did not get stricter after the revision. On the contrary, it allowed participants to choose from additional models and purchase less efficient appliances.

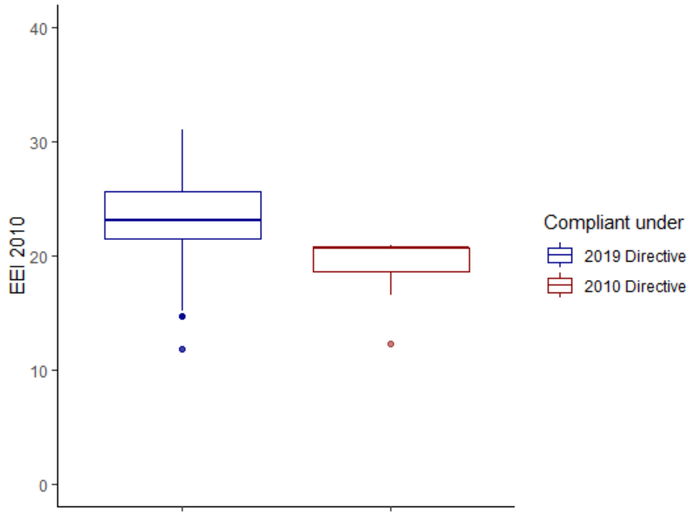


Figure 5.5: 2010 EEI of models under 2010 and 2019 Directive
Notes: This figure shows the distribution of the Energy Efficiency Index under the 2010 Regulation for models registered under both Directives 1060/2010 and 2016/2019 that are compliant with assistance program requirements (energy classes A+++ and A to D, respectively).

5.5 Consideration of energy cost

5.5.1 Empirical strategy

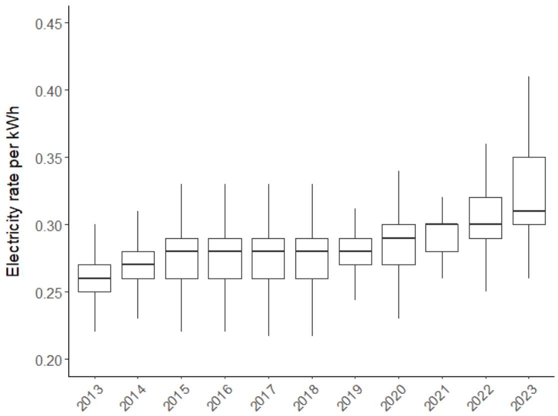
The analysis investigates how households trade off a change in the purchase price versus a change in lifetime energy cost, which reveals to what extent households consider energy cost in their investment decisions. If households rationally consider energy cost, they are expected to trade off an additional Euro to be paid in the purchase price against an additional Euro in discounted lifetime energy cost. A rational household should be indifferent between an appliance with a lower purchase price and higher energy cost, and an appliance with a higher purchase price and lower energy cost, as long as the net present value of total cost is the same, other features do not differ between the two appliances and abstracting from risk preferences.[27]

To study the trade-off between purchase price and energy cost, I employ demand response estimation to infer the relative demand responsiveness to changes in the net purchase price and the energy cost at the regional market level. If households fully consider energy cost, the strength of demand responses to both cost components should be equivalent.[28] Estimation relies on within-model variation in electricity rates and net

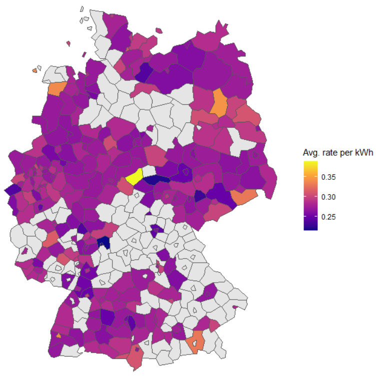
[27]Most valuation studies do not consider risk preferences, even for larger purchases, such as cars or heating systems. Compared to these studies, the context of refrigerators provides a relatively risk-free setting: the technology is established and households know the product class well. Moreover, purchase price and energy cost is low relative to a car or heating system.

[28]The approach follows Houde and Myers (2021) who use demand response estimation to determine valuation of energy cost for household appliances in the US general population.

purchase prices over time and across space to identify households' demand responses to both cost components. This section discusses, first, variation in electricity rates and purchase prices, second, estimation of the demand responses and threats to identification, and, third, assumptions regarding lifetime energy cost and the calculation of the trade-off ratio that compares demand responses to purchase prices and energy cost.



(a) By year



(b) By NUTS3 region

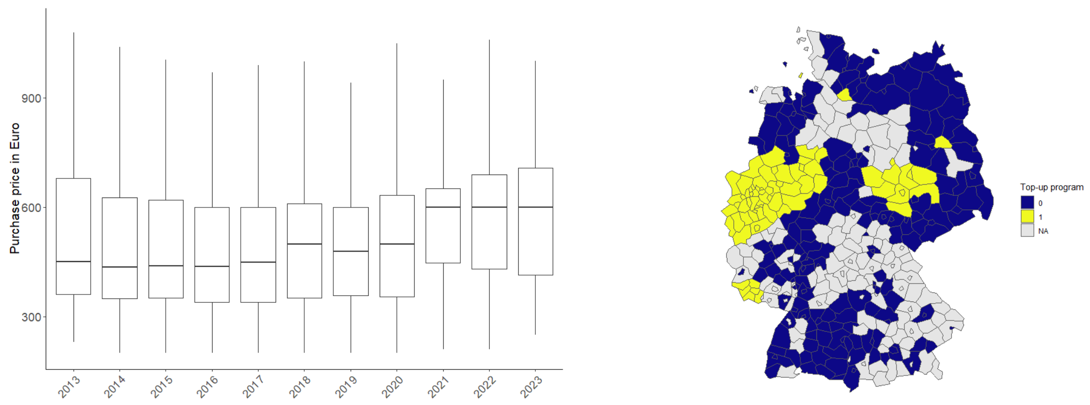
Figure 5.6: Variation in electricity rates

Notes: This figure shows the development of average electricity rates per year between 2013 and 2023 in Panel (a) and average electricity rates by NUTS3 region in Panel (b).

Variation in electricity rates. The estimation relies on variation in electricity rates over time and across space to identify households' consideration of energy costs. In the aggregated data, electricity rates vary over time and across region, as illustrated in Figure 5.6. In Germany, electricity prices have continuously increased over the past 20 years, rising from on average 0.29 Euro in 2014 to 0.46 Euro in 2023, caused predominantly by increases in taxes, charges and levies (BDEW, 2024). This trend is also evident in my data (see Panel (a)).[29] Across regions, variation in electricity rates is driven by differences in network, procurement and distribution charges (BDEW, 2024), where network charges account for the largest share of the regional price component. The sharp increase in rates over time is largely due to a steep rise in these regional components. In 2014, regional components accounted for around 0.16 Euro on average, while national components made up about 0.13 Euro. In 2023, regional components had risen to 0.35 Euro, while national components had decreased to 0.11 Euro of the average rate price. Panel (b) illustrates spatial variance in my data at the NUTS3 level.

Variation in purchase prices and subsidies. Variation in net purchase prices stems from variation in retailer purchase prices and variation in subsidies that house-

[29]Note, however, that rates in my data reflect the marginal price households pay for each additional kWh, whereas BDEW calculations represent the average rate per kWh consumed, accounting for monthly service fees.



(a) By year

(b) By NUTS3 region

Figure 5.7: Variation in purchase prices and program subsidies

Notes: This figure shows the development of average purchase prices per year between 2013 and 2023 in Panel (a) and the existence of top-up subsidy schemes across states in Panel (b).

holds receive as part of the refrigerator replacement scheme. Average purchase prices have slightly increased over the observation period, as illustrated in Panel (a) of Figure 5.7. On average, prices have increased most from 2017 to 2018 and from 2020 to 2021. Additional variation in net purchase prices is induced by changes to the program subsidy over time and across states. The federal subsidy for all participating households changed by 50 Euro in April 2019. Moreover, the states Hamburg, North Rhine-Westphalia, Saxony-Anhalt, Berlin and Saarland operated top-up subsidy schemes during the observation period, with subsidy top-ups ranging between 50 and 200 Euro.[30]

Estimation. The demand response estimation uses data aggregated at the model-region-quarter-year level.[31] The estimating equation is derived from a discrete choice model, modeling utility of consumer i for model m in region r and time period t, which, once aggregated at the region-period level, describes the demand for a model m in that cell. Demand is determined by purchase price, energy cost and region- and time-specific preferences. The number of appliances sold of each model is the product of choice probability at the individual consumer level and region-period market size. Using Poisson GLM, I estimate the following equation:

$$\begin{aligned} Nsales_{mrq} =& \alpha EC_{mrq} + \theta P^{net}_{mrq} + \gamma_m + \delta_q + \zeta_r + \eta_{ys} \\ & + Brand_m \times \iota_y + Brand_m \times \kappa_s + Volume_m \times \kappa_s + \varepsilon_{mrq} \end{aligned} \tag{5.1}$$

I regress the number of appliance sales $Nsales$ by model m, region r, and quarter-year q on the annual energy cost EC and the net purchase price P^{net}. Model fixed effects

[30]See Section 5.3 for a detailed account of how subsidies have changed over the observation period.
[31]The process of data preparation is described in Section 5.3.

γ control for time- and region-invariant preferences common to all consumers. Hence, the estimation exploits only within-model variation in energy cost based on differences in electricity rates across regions and over time and in net purchase prices based on differences over time and variation in subsidy levels. I use a comprehensive set of fixed effects to control for potential correlated demand shocks based on the approach by Houde and Myers (2021). δ and ζ capture differences in preferences across regions and over quarter-years. η captures region-specific trends at the state-year level sy, and ι and κ differences by state and year by manufacturing brand and differences across states by above- or below median volume of appliances, controlling for state-specific differences in preferences for refrigerator attributes. Standard errors are clustered at the model level.

Threats to identification. Causal identification of the demand responses relies on the assumption that purchase prices and electricity rates are exogenous to demand. Threats to this assumption come from omitted variable bias in the form of correlated demand shocks and measurement error in energy cost. Demand shocks are problematic if they are correlated with electricity rates or purchase prices. For instance, preferences over time and across space may differ and these differences could simultaneously affect electricity rates and demand for specific appliances. To mitigate this concern, I use a comprehensive set of fixed effects to control for demand shocks. Moreover, the largest part of variation in electricity rates stems from differences across regions which in turn is driven by differences in network charges. These are plausibly exogenous to consumer preferences. In an extension, I plan to use network charges as instrument for energy cost to isolate exogenous variation in electricity rates. Potential endogeneity could also be introduced by measurement error in energy cost. If households do not calculate with the exact energy consumption of their appliances but use a heuristic of sorting appliances into high- versus low-efficiency models, it could bias the estimation. In a further extension, I plan to instrument energy cost with a grouping estimator as used in Allcott and Wozny (2014) and Houde and Myers (2021) to address the concern.

For purchase prices, identification of the demand response relies on the assumption that national pricing strategies by retailers are exogenous to regional market conditions. Regional demand shocks are problematic if they are correlated with the level of purchase prices. This could be the case if retailers would set prices for appliances at the regional level. The assumption of national price-setting is supported by two key factors: First, the widespread availability of online purchasing that reduces the potential for regional price discrimination. As many consumers buy appliances online - this may in particular be the case for low-income households who are less likely to own a car and do not have to worry about transporting the appliance from the store if it is delivered - region-specific pricing should be less important. Second, marketing policies like the *Preisversprechen* suggest uniform prices across retailers online and in-store: The two largest electronics retail chains in Germany (*MediaMarkt* and *Saturn*) offer customers who find a product for a cheaper price at a competitor, including online platforms such as Amazon, an equivalent discount. This approach suggests that pricing strategies tend to be uniform and set at the national level.[32]

Lifetime energy cost. To analyze households' consideration of discounted lifetime energy cost, I make assumptions regarding the lifetime of appliances, the discount ratio that households face and expectations about future energy prices that households hold. Using these assumptions, lifetime energy cost is calculated summing appliance-specific annual energy cost over the lifetime using exponential discounting, and can be reformulated using the summation formula for geometric series:

$$\begin{aligned} LEC =& \sum_{l=0}^{L} EC * (\frac{1}{1+r})^l \\ =& EC * \rho(\frac{1-\rho^L}{1-\rho}) \end{aligned} \tag{5.2}$$

where LEC represents discounted lifetime energy cost. Annual energy cost EC is discounted at rate r to the year of purchase $t = 0$ and summed over the expected appliance lifetime L. I choose conservative assumptions for these parameters which provide me with a lower bound of households' expected lifetime energy cost. This approach biases the analysis against finding less than full consideration of energy cost and is a strict test of the hypothesis that low-income consumers under-consider energy cost.

A consumer survey on refrigerator age in European households estimates appliance lifetimes ranging between 10 and 17 years (Faberi et al., 2007).[33] Refrigerators in my sample tend to be older – the highest reported age is 45 years – as expected in this income segment. Yet, the SSC assistance program recommends replacing appliances older than 10 years. Therefore, I assume a uniform lifetime of 10 years as a conservative estimate and provide a sensitivity analysis for a product lifetime of up to 45 years.

Making assumptions about the discount rate is challenging as individual market discount rate data usually is not available and estimates for discount rates in the literature vary widely. A strand of literature related to the valuation of lifetime energy cost estimates discount rates that rationalize technology choices in observed investment decisions. These implicit discount rates factor in all aspects that affect consideration of

[32] Houde and Myers (2021) show that a large electronics retail chain in the US follows a national pricing scheme.

[33] The survey was conducted on behalf of the European Commission in preparation for EU Directive 1060/2010. 10 percent of the sample consisted of German households, with responses being fairly homogeneous across participating member states.

lifetime energy cost, such as preferences, irrational behavior or external barriers (Schleich et al., 2016). Implicit discount rates have been shown to vary with income as well as other characteristics, such as employment status and age (Hausman, 1979; Train, 1985; Harrison et al., 2002). The variation across income likely stems from credit constraints and economic preferences that vary with income and affect implicit discount rates (Schleich et al., 2016). However, implicit discount rates usually do not reflect the market discount rates that households face.

For rational households, the interest rate that they can borrow at should inform their discount rate for purchase decisions. Consumer credit rates for the purchase of household appliances at major German retailers are the relevant market interest rates in this context. These rates are in the range between 10 to 20 percent as of May 2024.[34] Since interest rates were lower over the largest part of the observation period, these ranges provide an upper bound of consumer credit rates. Eligibility to installment plans sometimes requires a permanent employment contract or permanent retirement payments as well as a positive credit rating at Schufa or Infoscore, the largest German credit investigation companies, requirements, which may exclude recipients of unemployment benefits (*Bürgergeld, ALGII*). Households on basic retirement benefits or a housing allowance may be eligible depending on their credit history. For small investment amounts some households may also be able to obtain interest-free loans from family or friends. Hence, individual discount rates may vary greatly. I assume a discount rate of 20 percent for calculation of the baseline lifetime energy cost, following the upper bound of consumer credit rates over the observation period. I provide robustness checks at discount rates of up to 100 percent.

Using current electricity rates for calculation of annual energy cost carries the assumption of a no-change forecast regarding future electricity price expectations. This approach is common in the valuation literature and has been shown to align with consumers' expectations about gasoline prices (Anderson et al., 2011). In the context of the German electricity retail sector, the no-change assumption is rather conservative since consumer electricity rates have been increasing since the early 2000s. Widespread expectations for electricity prices to significantly decline in the medium term are fairly implausible, even for uninformed consumers.

My analysis does not account for uncertainty in lifetime energy cost and related risk preferences. Previous studies on the valuation of lifetime energy cost for houses, cars, and household appliances have typically not considered risk and uncertainty as factors. I implicitly address uncertainty about the different components of lifetime energy

[34]Several large retailers offer installment plans for online purchases, including Ikea (0 to 8.99 percent APR over 3 to 32 months for purchases of 800 Euro), Otto Group (18.70 to 18.95 percent APR over 3 to 48 months for purchases of 800 Euro), and Media-Saturn Holding (11.9 percent APR over 6 to 60 months for purchases of 800 Euro).

energy cost by providing sensitivity checks that vary assumptions regarding appliance lifetime and discount rate.

Trade-off ratio. The final step is to compare the relative strength of demand responses to changes in the purchase price and in energy cost. Following Allcott (2013), I calculate the ratio between coefficients for annual energy cost α and purchase price θ. For both cost components to be comparable, I scale θ using the factor for the summation of annual energy cost over lifetime L using discount rate r as calculated in Equation 5.2. If consumers weigh discounted lifetime energy cost equally as the purchase price, the demand responses to energy cost and scaled purchase price should be of comparable size, so that $\alpha = \theta\rho(\frac{1-\rho^L}{1-\rho})$. Taking the ratio between the left- and right-hand side provides me with the trade-off ratio

$$m = \frac{\alpha}{\theta\rho(\frac{1-\rho^L}{1-\rho})} \tag{5.3}$$

If $m = 1$, households trade off one Euro in the purchase price against one Euro in lifetime energy cost and therefore fully consider energy cost. If $m < 1$, households do not fully consider energy cost, and if $m = 0$, households do not consider energy cost at all in their purchase decisions. I use the Delta method to approximate the variance of m with a normal distribution to deduct standard errors for inference.

5.5.2 Results

Table 5.6 reports the results of the estimation. The coefficient on the demand response to annual energy costs is at 0.0053 insignificant, an imprecisely estimated zero result. The estimate suggests that households may not strongly adjust their demand to changes in energy cost. The coefficient on the net purchase price is at -0.0011 significant negative. As expected, higher purchase prices reduce demand. On average and ceteris paribus, an increase in the purchase price of a model decreases its demand at the regional level by 0.001 units sold.

Using both coefficients, m is calculated as detailed in Equation 5.3. Applying assumptions for lifetime and discount rate as discussed above yields a ratio of -1.1 that is insignificant from zero. The relevant test for the hypothesis that low-income households do not rationally consider energy cost is whether $m < 1$. In a one-sided z-test, m is found to be significantly less than 1 at the 5% level (H0: $m = 1$, H1: $m < 1$, p = 0.027). Energy cost is not fully considered in purchase decisions.

Sensitivity. I run sensitivity checks regarding the assumptions for discount rate and appliance lifetime. For different values of the discount rate, Figure 5.8 shows the trade-off ratio and the upper bounds of the 90 and 95% confidence intervals for the one-sided

Table 5.6: Main results for the demand response estimation

Dependent Variable:	N models sold
Model:	(1)
Annual energy cost	0.0053
	(0.0044)
Net purchase price	-0.0011***
	(0.0004)
Estimator	GLM Poisson
Fixed effects	Yes
Observations	1,178,105
Trade-off ratio	-1.138
$\frac{\alpha}{\theta\rho(\frac{1-\rho^L}{1-\rho})}$	(1.1088)
p-value H0: $m = 1$	0.027

Notes: Fixed effects include model, region, quarter-year, state-year, volume-state and manufacturing brand-state fixed effects. Standard errors are clustered at the model level. The estimation drops 5,027 observations due to fixed effects cells that contain only 0 outcomes. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

z-test (H0: $m = 1$, H1: $m < 1$). At the 5% significance level (light-red dashed line), m is significantly less than 1 for discount rates below 33 percent. At 10% significance level (dark-red dashed line), the threshold is at a discount rate of 84 percent. For the appliance lifetime, Figure 5.9 shows the equivalent statistics. At 5% and 10% significance level, m is significantly less than 1 above a respective lifetime of 4 and 2 years.

In another robustness check, I calculate m using the assumptions used in Houde and Myers (2021) who study consumers' consideration of energy cost for purchases of refrigeration appliances in the general population in the US. For their main result, they assume a discount rate of 5 percent and an appliance lifetime of 12 years. Using these parameters, m is at -0.538 and significant from 1 at the 1% level (p-value = 0.002).

Discussion. The result implies that households do not fully consider lifetime energy cost under plausible assumptions regarding the lifetime of appliances and discount rates. The finding of low consideration aligns with theoretical considerations for factors that may reduce consideration of energy cost in lowest-income households, such as financial constraints and behavioral biases. In the following sections, I empirically test

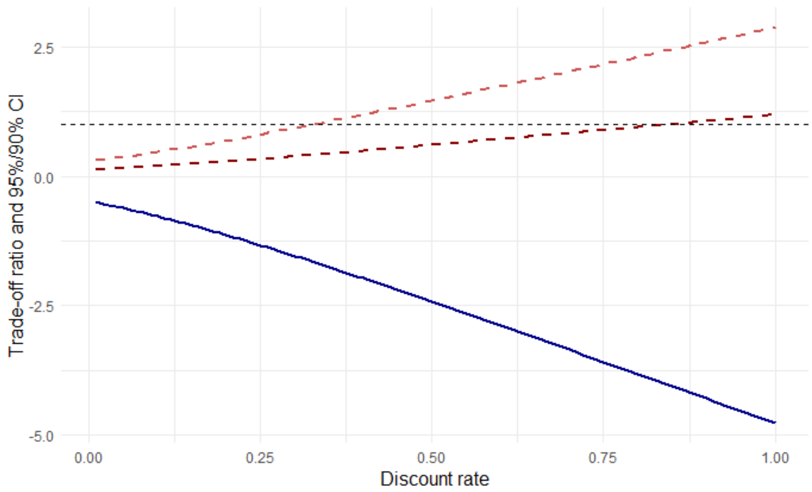


Figure 5.8: Sensitivity of the trade-off ratio to variation in the discount rate
Notes: This figure shows the trade-off ratio as a function of the discount rate in blue, fixing the appliance lifetime at 10 years. The upper bound of the 95% and 90% confidence intervals for a one-sided z-test (H0: ratio = 1, H1: ratio < 1) is drawn in light- and dark-red.

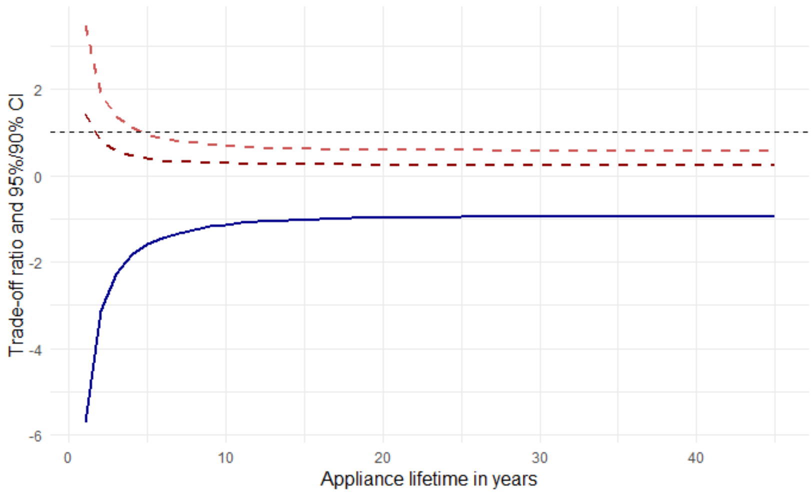


Figure 5.9: Sensitivity of the trade-off ratio to variation in the appliance lifetime
Notes: This figure shows the trade-off ratio as a function of the appliance lifetime in blue, fixing the discount rate at 20 percent. The upper bound of the 95% and 90% confidence intervals for a one-sided z-test (H0: ratio = 1, H1: ratio < 1) is drawn in light- and dark-red.

the role of both factors discussed in the theoretical framework in Section 5.2: the effect of an increased budget for the appliance purchase and the effect of increased salience of energy efficiency on consideration of energy cost in respective Sections 5.6 and 5.7.

5.6 Financial constraints

Liquidity and borrowing constraints may theoretically explain part of the energy efficiency gap (Allcott and Greenstone, 2012). Higher upfront investments in more efficient durables are optimal if the higher investment cost is offset by future energy savings over the durable's lifetime. If financially constrained consumers cannot afford the higher investment, they may be forced to buy a less efficient appliance below the optimal efficiency level. Empirically, the proposition has been tested for the valuation of fuel cost in car purchases within the US general population where credit constraints and the cost of credit are not relevant (Ankney et al., 2021; Levinson and Sager, 2023). However, in developing economies like Uganda and Kenya, access to credit matters for the adoption of fuel-efficient cook stoves (Levine et al., 2018; Berkouwer and Dean, 2022).

For refrigerators on the German market, the premium for more efficient appliances is sizable. Using the household-level dataset of purchase decisions (see Section 5.3.2), I check the relationship between energy consumption and purchase price and calculate the difference for appliances in the same volume bracket for both variables. For these calculations, I aggregate the data at the model level. The theorized relationship between kWh consumption and purchase price is present in my data. Using the full sample of models, I find the purchase price to drop by 0.51 Euro for an increase in energy consumption by one kWh controlling for brand, type, volume and median purchase year. Looking at the same relationship for different volume brackets, the relationship seems stronger for smaller appliances than for larger appliances (see Table A5.1 in the Appendix). Within a volume bracket, the difference between the lowest and highest-efficiency level of models is on average at 37.2 kWh and the average price difference is at 388.5 Euro. The efficiency premium appears sizable. Financial constraints may be a relevant factor driving under-consideration of energy cost.

5.6.1 Empirical strategy

To explore the relevance of financial constraints for consideration of energy cost, I examine how an increase in the budget for appliance purchases affects the trade-off ratio m. I use exogenous variation in the form of changes in the level of program subsidies available to households. These changes vary the budget available for purchases, but are not related to household characteristics and plausibly exogenous to households' decision-making as they are decided by the federal sponsor of the program and the respective state governments. During the period of observation, the federal subsidy was reduced by 50 Euro in April 2019 and top-up state subsidy programs were introduced in North Rhine-Westphalia (NRW) in July 2016, in Saxony-Anhalt in May 2020, and in Berlin in December 2020.[35]

I exploit this variation in three different types of research designs using (i) a comparison of means to study the reduction of the federal subsidy, (ii) a staggered two-way fixed-

[35] Additional top-up programs were introduced in Hamburg and Saarland. However, Hamburg introduced the top-up subsidy already with the start of the program and local branches in Saarland provide only few data points, so that these programs cannot be used for analysis.

effects design to study the introduction of the three top-up programs jointly, and (iii) separate Difference-in-Difference designs to analyze the effect of each top-up program separately. The focus of this section is on (i) as the change in the federal subsidy provides most power for the estimation.[36] Designs in (ii) and (iii) provide robustness checks.

Identifying assumptions. Identifying assumptions for the comparison of means to provide an unbiased estimate of the shift in demand responses due to an increased budget for purchase are, first, random assignment of households to the pre- and post-treatment period and, second, comparability of investment decisions pre- and post-treatment. The first assumption deals with the exogeneity of the subsidy reduction: the change must be exogenous in the sense that participating households should not be able to manipulate whether they purchase a new refrigerator claiming a subsidy of 100 or 150 Euro. Households would only have an incentive to manipulate timing if they knew in advance that the subsidy will change. However, the subsidy reduction was dictated by the sponsor of the program and program officials themselves only learned about this change on short notice before the start of the new funding period. Moreover, households typically join a wait list before entering a program, with wait times varying by season and program location, making precise manipulation unlikely. Nevertheless, I use the household-level data (see Section 5.3.2) to check for empirical evidence of manipulation. The histogram of voucher issuances in the study period shows strong seasonal variation. Moreover, the weekly numbers of vouchers issued are at a higher level before the voucher reduction than after (see Figure A5.5 in the Appendix). This can be explained by changes to the administrative structures of the program. To empirically demonstrate that there is no evidence for sorting into treatment, I test for discontinuities in household characteristics around the treatment date, allowing for different degrees of flexibility in RDD-style regressions (see Table A5.2 in the Appendix). I do not find evidence for discontinuities in household characteristics.

The second assumption deals with the comparability of purchase decisions before versus after treatment. I test the balance of household characteristics in a bandwidth of two years before and after treatment. The age of old refrigerators and total energy consumption do not change. The number of persons drops by 0.1 and estimated savings from replacement increase by 9.4 kWh. Albeit statistically significant, these changes are economically small.

Estimation. The estimating equation to compare the trade-off for households with a

[36]See Chlond et al. (2025) for an analysis of how the reduction in the subsidy affected investment rates in the program.

higher versus a lower federal subsidy is as follows:

$$
\begin{aligned}
Nsales_{mrq} =&\alpha_1 EC_{mrq} + \theta_1 P^{net}_{mrq} + \alpha_2 EC_{mrq} \times Subsidy^{+50}_q + \theta_2 P^{net}_{mrq} \times Subsidy^{+50}_q \\
&+ \gamma_m + \delta_q + \zeta_r + \eta_{ys} + Brand_m \times \iota_y + Brand_m \times \kappa_s + Volume_m \times \kappa_s + \varepsilon_{mrq}
\end{aligned}
\tag{5.4}
$$

The model extends the specification in Equation 5.1 adding respective interactions between annual energy cost EC and the net purchase price P^{net} with treatment indicator $Subsidy^{+50}$, indicating the treatment period with higher subsidy before April 2019.[37] Fixed effects include the same set as in the basic specification: model, region, quarter-year, year-state, manufacturing brand- and volume-year and brand-state. Standard errors are clustered by model. In addition to the estimation on the full sample in a bandwidth of two years around the date of the subsidy reduction (Q2 2017-Q1 2021), I also estimate a second specification that leaves out observations from Berlin and Saarland as these states introduced a top-up program during this period, and a third specification that reduces the bandwidth to one year around the date of the subsidy reduction (Q2 2018-Q1 2020).

Trade-off ratio. The relevant test for whether the trade-off changes due to the increase in the budget is whether m shifts upwards. To check this, I calculate the ratio with the by 50 Euro higher subsidy and subtract the baseline ratio at the lower subsidy level, the difference representing the change in m:

$$
\Delta m^{+50} = \frac{\alpha_1 + \alpha_2}{(\theta_1 + \theta_2)\rho(\frac{1-\rho^L}{1-\rho})} - \frac{\alpha_1}{\theta_1 \rho(\frac{1-\rho^L}{1-\rho})} \tag{5.5}
$$

I calculate standard errors using the Delta method to test whether statistically $\Delta m^{+50} > 0$ holds.

Robustness checks. As robustness checks, I exploit variation in the budget for appliance purchases stemming from state-level top-up subsidies in NRW, Saxony-Anhalt and Berlin in an event study setting and in separate Difference-in-Difference estimations. I estimate a staggered two-way fixed effects model jointly for all three top-up programs, using observations in a bandwidth of two years around each top-up program implementation, and in a bandwidth of one year. I estimate the following equation, with the treatment dummy *Topup* indicating post-treatment periods for the treated

[37] The variable is coded anti-chronologically: As the subsidy was reduced on the treatment date, the treatment period with a higher subsidy before the reduction is coded 1. This approach facilitates the comparison with the results from the robustness checks where treatment in all cases increases the subsidy.

regions:

$$
\begin{aligned}
Nsales_{mrq} =& \alpha_1 EC_{mrq} + \theta_1 P^{net}_{mrq} + \alpha_2 EC_{mrq} \times Topup_{sq} + \theta_2 P^{net}_{mrq} \times Topup_{sq} \\
&+ \gamma_m + \delta_q + \zeta_r + \eta_{ys} + Brand_m \times \iota_y \\
&+ Brand_m \times \kappa_s + Volume_m \times \kappa_s + \varepsilon_{mrq}
\end{aligned} \tag{5.6}
$$

The procedure to calculate the trade-off ratio is the same as in Equation 5.5.

For the separate Difference-in-Difference estimations for top-up programs in NRW, Berlin and Saxony-Anhalt, the estimating equation is as follows:

$$
\begin{aligned}
Nsales_{mrq} =& \alpha_1 EC_{mrq} + \theta_1 P^{net}_{mrq} + \alpha_2 EC_{mrq} \times post_q + \theta_2 P^{net}_{mrq} \times post_q \\
&+ \alpha_3 EC_{mrq} \times treat_s + \theta_3 P^{net}_{mrq} \times treat_s \\
&+ \alpha_4 EC_{mrq} \times post_q \times treat_s + \theta_4 P^{net}_{mrq} \times post_q \times treat_s \\
&+ \gamma_m + \delta_q + \zeta_r + \eta_{ys} + Brand_m \times \iota_y \\
&+ Brand_m \times \kappa_s + Volume_m \times \kappa_s + \varepsilon_{mrq}
\end{aligned} \tag{5.7}
$$

post indicates the onset of the treatment period with top-up subsidys and *treat* indicates belonging to the treated group in the state with top-up program. Calculation of the change in the trade-off ratio caused by the top-up subsidy to test $\Delta m^{Topup} > 0$ is as follows:

$$
\begin{aligned}
\Delta m^{Topup} =& \left(\frac{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4}{(\theta_1 + \theta_2 + \theta_3 + \theta_4)\rho(\frac{1-\rho^L}{1-\rho})} - \frac{\alpha_1 + \alpha_3}{(\theta_1 + \theta_3)\rho(\frac{1-\rho^L}{1-\rho})}\right) \\
&- \left(\frac{\alpha_1 + \alpha_2}{(\theta_1 + \theta_2)\rho(\frac{1-\rho^L}{1-\rho})} - \frac{\alpha_1}{\theta_1 \rho(\frac{1-\rho^L}{1-\rho})}\right)
\end{aligned} \tag{5.8}
$$

In all robustness check estimations I leave out the quarter of implementation for Berlin (Q4 2019) and Saxony-Anhalt (Q1 2020) as the programs do not start at the beginning of a quarter.

5.6.2 Results

Figure 5.10 provides a pewview of the regression results, graphically illustrating changes in demand responses to energy cost and purchase price over the study period. Before the reduction of the federal subsidy by 50 Euro, demand responses to both cost components do not significantly vary. Moreover, after the subsidy reduction, demand responses to annual energy cost (red coefficients, left-hand y-axis) do not shift upward as hypothesized. Demand responses to changes in the purchase price (blue coefficients, right-hand y-axis) do not either change significantly.

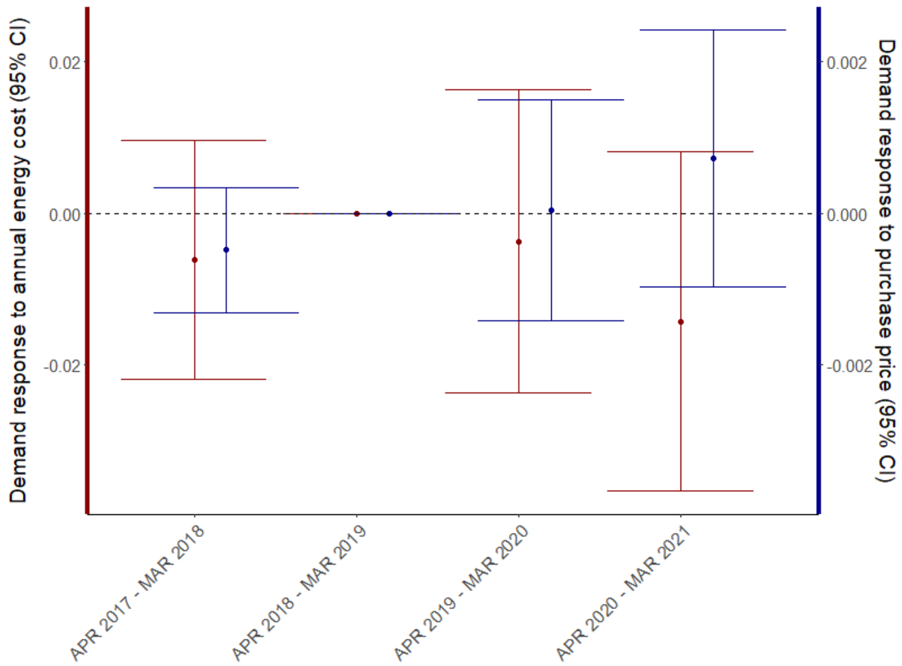


Figure 5.10: Financial constraints: Change in demand response over study period

Notes: This figure is created from a regression of Equation 5.4 that includes a categorical variable with four bins (for each year between Q2 2017 and Q1 2021) in place of the $Subsidy^{+50}$ indicator. The scale for demand responses to annual energy cost is provided on the left-hand x-axis and corresponding coefficients are drawn in red. The right-hand x-axis and blue coefficients show demand responses to the purchase price. The reference year (Q2 2018-Q1 2019) is the period before the implementation of the subsidy reduction.

The results from the estimations presented in Table 5.7 confirm the graphical results. In all three specifications, demand responses to both annual energy cost and the net purchase price do not significantly change due to a higher subsidy. Calculating the change in the trade-off ratio after the subsidy increase using the formula provided in Equation 5.5, I find Δm to be around zero in column (1) using the full sample in a bandwidth of two years around the subsidy change. The change is insignificant positive in column (2) when excluding observations from Berlin and Saarland. In column (3) when only considering observations in a bandwidth of one year, the change is insignificant negative. These results all point to the conclusion that financial constraints do not matter for the consideration of energy cost in this setting.

Robustness checks. Further variation in the budget for the appliance purchase induced by top-up subsidy programs at the state level provides me with various settings to conduct robustness checks. Estimating a staggered two-way fixed effects model (see Equation 5.6) and Difference-in-Difference specifications (see Equation 5.7) for top-up programs in NRW, Saxony-Anhalt and Berlin, I find results that altogether oppose my hypothesis. None of the resulting changes in trade-off ratios is significant positive. Indeed, all five estimations yield negative changes in m (see Tables A5.4 and A5.5 in the

Table 5.7: Main results for the demand response estimation: Financial constraints

Dependent Variable:	N models sold		
Model:	(1)	(2)	(3)
Annual energy cost	0.0169**	0.0211**	0.0127
	(0.0082)	(0.0082)	(0.0133)
Net purchase price	-0.0010	-0.0010	-0.0017
	(0.0008)	(0.0008)	(0.0011)
Annual energy cost × Subsidy^{+50}	0.0065	0.0045	0.0031
	(0.0093)	(0.0096)	(0.0112)
Net purchase price × Subsidy^{+50}	-0.0004	-0.0005	5.85×10^{-6}
	(0.0007)	(0.0007)	(0.0009)
Estimator	GLM Poisson	GLM Poisson	GLM Poisson
Fixed effects	Yes	Yes	Yes
Observations	628,432	568,903	279,365
Change in trade-off ratio:	0.07	0.81	-0.46
$\frac{\alpha_1+\alpha_2}{(\theta_1+\theta_2)\rho(\frac{1-\rho^L}{1-\rho})} - \frac{\alpha_1}{\theta_1\rho(\frac{1-\rho^L}{1-\rho})}$	(2.15)	(2.78)	(1.19)
p-value H0: $m = 0$	0.486	0.385	0.649

Notes: The sample used for column (1) is the sample Q2 2017 to Q1 2021, for column (2) observations from Berlin and Saarland are excluded, and for column(3) the bandwidth is restricted to Q2 2018 to Q1 2020. Fixed effects include model, region, quarter-year, volume-state, manufacturing brand-state and manufacturing brand-year fixed effects. In column (1) 31,753 observations were dropped due to only zero outcomes in fixed effects cells, in column (2) 44,177 observations, and in column (3) 53,555 observations, respectively. Standard errors are clustered at the model level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Appendix). These robustness checks confirm the main result: an increase in the budget for appliance purchases of between 50 and 200 Euro does not change consideration for energy cost.

Discussion. To conclude that financial constraints are not binding in this setting, I need to rule out that the variation in subsidy amounts observed is simply too low to significantly increase the budget and enable households to purchase a higher-efficiency refrigerator. On average, observed households purchase appliances for 490 Euro. An increase in the budget by 50 to 200€ due to higher subsidies provides a sizable fraction of total prices. If households are budget-constrained, I would expect to observe a positive tendency in trade-off ratio changes, which no empirical evidence is found for. Chlond et al. (2025) show in the same setting that a change in the federal subsidy by 50 Euro significantly affects households' purchase decisions at the extensive margin, increasing purchase rates by five to seven percentage points. Financial constraints appear to matter at the extensive margin, however they do not seem binding at the intensive margin. Under-consideration of energy cost is not explained by monetary barriers.

5.7 EU Energy Label revision

Inattention has been put forward as another explanation why consumers do not consider lifetime energy cost in their purchase decisions when other features are more salient at the time of purchase (Allcott and Greenstone, 2012). Inattention to energy cost can take various forms. Inattention is rational in purchase decisions where the cost of information acquisition exceeds expected energy savings (Sallee, 2014). Coarse certification, such as energy efficiency labeling, provides simple but incomplete information which can discourage more informed decisions (Houde, 2018b). Empirically, inattention has been found to matter in some contexts, but not in others. Leard (2018) finds that in the US self-reported inattention to fuel cost is heterogeneous and correlated with lower valuation of fuel cost in stated-choice car purchase decisions, and the evidence is consistent with rational inattention. Allcott and Taubinsky (2015) conduct a field experiment to measure the impact of information provision regarding lifetime energy cost on the willingness-to-pay for efficient fluorescent light bulbs relative to incandescent ones. WTP increases, though at least part of the effect is due to updated beliefs rather than inattention. Berkouwer and Dean (2022) employ an information intervention designed to highlight energy savings, but find inattention of no relevance in explaining the energy efficiency gap in the adoption of cook stoves in Kenya.

This analysis exploits the revision of the EU Energy Label on 1 March 2021 and the resulting change in salience of energy efficiency, closest related to the aspect of coarse certification (Houde, 2018b; Andor et al., 2020; d'Adda et al., 2022). A revised version of the EU Energy Label for refrigeration appliances was implemented under Directive 2019/2016 (European Commission, 2019) and the new label came into force on 1 March 2021. The old and new label are presented side-by-side in Figure 5.11. While the old scale ranged from A+++ to D, the new scale ranges from A to G and appliances which before had been located in energy class A+++ were sorted into classes A to D. With the revision of the label, the formula for the Energy Efficiency Index that determines the energy class of a model was revised. The revision not only shifted appliances that before ranked in the highest class down to the middle range in the new scale. The change in the formula also increased the spread in efficiency differences between models. These changes were made to foster incentives for technological innovation and to improve consumer information.

In accordance with the revision of the label, the SSC assistance program adjusted its

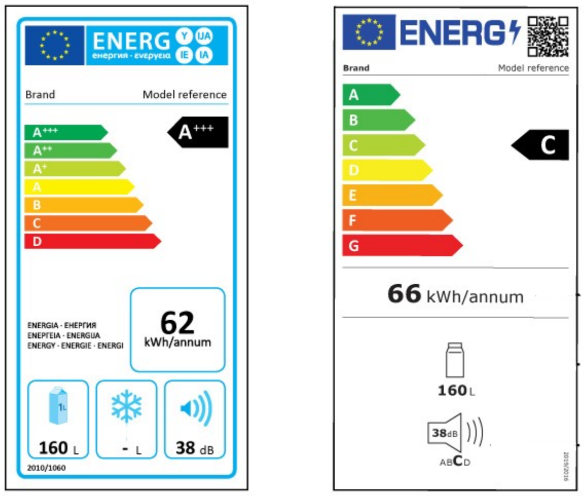


Figure 5.11: EU Energy Label before and after revision

Notes: This figure shows the EU Energy Label in its old version amended in Directive 1060/2010 (European Commission, 2010) on the left-hand side and in its new version established in Directive 2019/2016 (European Commission, 2019) on the right-hand side. The graphic originates from `https://ec.europa.eu/commission/presscorner/detail/en/ip_21_818`

minimum efficiency requirements. Before the revision, participating households had to purchase A+++ models to be eligible to claim the subsidy. Since the introduction of the new scale, households can chose from any model ranked D or higher. While the choice set of participants only included one energy class before the revision, it increased to four classes after. If households are attentive to the label, differences in efficiency of models in the choice set must have become more salient to households.

5.7.1 Empirical strategy

The analysis uses the revision of the energy label as a treatment that increased the salience in efficiency differences between models. For this approach, two conditions need to be satisfied: first, households are attentive to energy classes on the EU Energy Label, and, second, the assistance program's minimum requirement for the level of energy efficiency of appliances did not get stricter after the revision. The first condition is necessary as the change from one to four compliant classes would only affect households' purchases if they pay attention to the label. As households need to comply with requirements to claim the subsidy, they are forced to pay attention to appliance labels to infer which models are available for purchase. Additionally, descriptive evidence from Section 5.4.1 suggests that households value higher energy classes. They purchase appliances from classes A to C, despite them being more expensive than D. Overproportional bunching of purchases at the lower end of energy classes demonstrates that households, while noting differences between energy classes, are not attentive to differences in efficiency within energy classes.[38] This finding implies that before the revision households tended to perceive all appliances in A+++ being of roughly the same efficiency level, whereas after, differences in efficiency levels between energy classes A to D (but not within) became salient.

The second condition of no increased stringency in the minimum efficiency requirements is needed to be able to attribute the entire effect of the label revision to salience. Increased salience of efficiency is expected to increase consideration of energy cost. If minimum efficiency requirements became stricter, i.e., the lower bound of D on the new scale was a stricter requirement than the lower bound of A+++ on the old scale, it would potentially affect purchase decisions in the same direction. Descriptive evidence from Section 5.4.2 suggests that minimum efficiency requirements did not become stricter after the revision. While all models that would have been compliant under the old requirements were also compliant under the new requirements, an additional range of models that would not have been compliant under the old requirements, qualifies for the program after the revision. Moreover, both the least efficient model still compliant as well as the median compliant model are less efficient in the choice set during the two years after the revision as compared to during the two years before the revision. As both conditions are satisfied, the revision effect on consideration of energy cost can be attributed to increased salience.

Identifying assumptions. To examine the impact of the revision on demand responses to energy cost and purchase prices, I use a comparison of means following the same empirical strategy as for estimation of the effect of the federal subsidy change. Correspondingly, identifying assumptions for the estimation are random assignment of households to the pre- and post-treatment period and comparability of investment decisions pre- and post-treatment. First, the change must be exogenous, randomly assigning households to the period before and after the change and with no possibility of manipulation by participating households whether they purchase under the old or new program requirements. Households would only have an incentive to manipulate timing if they knew in advance whether program requirements would change to their advantage or disadvantage. However, program officials decided only shortly before the new label's introduction for the lower-bound efficiency requirement to be class D, giving households little scope for manipulation. Moreover, households typically join a wait list before entering a program, with wait times varying by season and program location, making precise manipulation unlikely. Nonetheless, I empirically check for evidence of manipulation. The histogram of voucher issuances in the study period shows strong seasonal variation (see Figure A5.6 in the Appendix). Though the density of voucher

[38] Similar results have been documented for the general population (Andor et al., 2020; d'Adda et al., 2022).

issuances is not uniformly distributed, there is no discontinuity visible around the date of the label revision. Large dents are visible at other points of time coinciding with end-of-year breaks, the first Covid lockdown in March and April 2020 and the start of a new funding cycle in April 2022. To empirically show that there is no evidence for sorting into treatment, I test for discontinuities in household characteristics around the treatment date, allowing for different degrees of flexibility in RDD-style regressions (see Table A5.6 in the Appendix). I do not find evidence for discontinuities in household characteristics.

Second, investment decisions before and after the change must be comparable. Economic conditions within and outside the program underwent a few changes. Comparing household covariates shows that while the average number of persons per household does not change, total energy consumption, estimated savings from replacement and the age of the old appliance to be replaced slightly change in the pre- versus post-treatment period. As a robustness check, I reduce the sample bandwidth around the label revision in order to limit changes in covariates and compare temporally closest observations before and after.

Estimation. The aim of the estimation is to compare the trade-off ratio before versus after the label revision. The estimating equation is as follows:

$$\begin{aligned} Nsales_{mrq} =& \alpha_1 EC_{mrq} + \theta_1 P^{net}_{mrq} + \alpha_2 EC_{mrq} \times Revision_q + \theta_2 P^{net}_{mrq} \times Revision_q \\ & + \gamma_m + \delta_q + \zeta_r + \eta_{ys} + Brand_m \times \iota_y + Brand_m \times \kappa_s + Volume_m \times \kappa_s + \varepsilon_{mrq} \end{aligned} \tag{5.9}$$

It extends the basic specification in Equation 5.1 adding respective interactions between annual energy cost EC and the net purchase price P^{net} with a treatment indicator $Revision$, indicating the onset of the treatment period after the label revision. Fixed effects include model, region, quarter-year, year-state, manufacturing brand- and volume-year and brand-state. Standard errors are clustered by model. I estimate three specifications with different sampling restrictions. The first specification uses the full sample in a bandwidth of two years around the label revision in accordance with the comparison of means estimated for the subsidy change: from Q2 2019 to Q4 2020 for the pre-treatment period and from Q2 2021 to Q1 2023 for the post-treatment period. As the treatment date on 1 March 2021 is halfway into the first quarter of 2021, I drop these observations in the analysis. The second specification additionally drops observations from the state Saarland as a subsidy top-up program was introduced during the study period, and the change in the financial constraint could theoretically affect consideration of energy cost in parallel with the label change. The third specification uses only observations in a bandwidth of one year around the treatment date, from Q1 2020 to Q1 2022.

Trade-off ratio. The hypothesis to be tested is that the salience treatment significantly shifts m upward. The relevant statistical test is whether Δm is significantly positive. The shift in the ratio is calculated as follows:

$$\Delta m^{salience} = \frac{\alpha_1 + \alpha_2}{(\theta_1 + \theta_2)\rho(\frac{1-\rho^L}{1-\rho})} - \frac{\alpha_1}{\theta_1 \rho(\frac{1-\rho^L}{1-\rho})} \tag{5.10}$$

I calculate standard errors using the Delta method to test whether $\Delta m^{salience} > 0$ holds.

5.7.2 Results

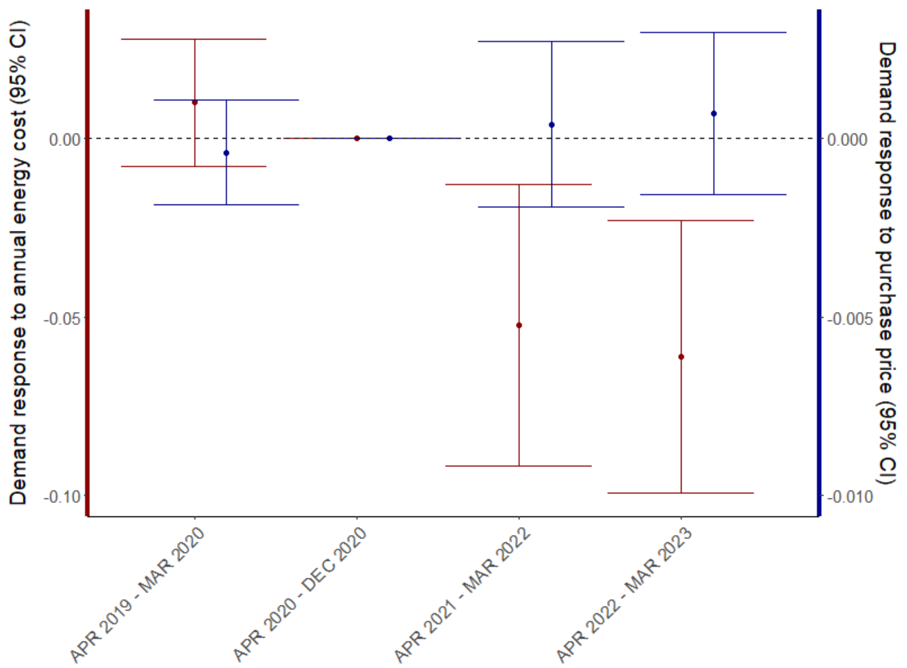


Figure 5.12: Change in demand responses: Revision of the EU Energy Label

Notes: This figure is created from a regression similar to Equation 5.9 that includes a categorical variable with four bins (for each year between Q2 2019 and Q1 2023) in place of the *Revision* indicator. The scale for demand responses to annual energy cost is provided on the left-hand x-axis and corresponding coefficients are drawn in red. The right-hand x-axis and blue coefficients show demand responses to the purchase price. The reference category (Q2 2020-Q4 2020) is the period before the label revision came into force.

As a preview to the regression results, Figure 5.12 illustrates how demand responses to annual energy cost (red coefficients, left-hand y-axis) and the purchase price (blue coefficients, right-hand y-axis) change due to the label revision. The pre-treatment coefficients in year t = -2 are not significant. With the implementation of the label revision, the coefficient on annual energy cost significantly drops, while the increase in the coefficient on the purchase price is small and insignificant. The direction of the change in the demand response to energy cost is as expected. Significant negative coefficients indicate an increase in the strength of the demand response, implying that households weigh energy cost relatively more after the label revision.

Table 5.8: Main results for the demand response estimation: Revision of the EU Energy Label

Dependent Variable:	N models sold		
Model:	(1)	(2)	(3)
Annual energy cost	0.0341^{**}	0.0343^{**}	0.0226
	(0.0140)	(0.0140)	(0.0148)
Net purchase price	-0.0015	-0.0015	-0.0019^{***}
	(0.0010)	(0.0010)	(0.0007)
Annual EC × Rescale	-0.0397^{**}	-0.0398^{**}	-0.0288^{*}
	(0.0163)	(0.0163)	(0.0161)
Net price × Rescale	3.37×10^{-5}	3.61×10^{-5}	0.0015^{*}
	(0.0013)	(0.0013)	(0.0009)
Estimator	GLM Poisson	GLM Poisson	GLM Poisson
Fixed effects	Yes	Yes	Yes
Observations	282,205	282,027	143,564
Change in trade-off ratio:	6.21^{**}	6.22^{**}	6.22
$\frac{\alpha_1+\alpha_2}{(\theta_1+\theta_2)\rho(\frac{1-\rho^L}{1-\rho})} - \frac{\alpha_1}{\theta_1\rho(\frac{1-\rho^L}{1-\rho})}$	(3.33)	(3.34)	(10.65)
p-value H0: $m = 0$	0.031	0.031	0.280

Notes: Estimation is done with the sample from Q2 2019 to Q1 2023 in column (1), leaving out observations from the state Saarland in column (2), and with the sample from Q1 2020 to Q1 2022 in column (3). Observations in Q1 2021 are excluded in all specifications. Fixed effects include model, region, quarter-year, state-year, volume-state, manufacturing brand-state and manufacturing brand-year fixed effects. Standard errors are clustered at the model level. The estimation drops 7,719 observations in column (1), 7,716 in column (2), and 23,369 in column (3) respectively due to fixed effects cells that contain only 0 outcomes. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 5.8 presents the results of the formal regression estimating Equation 5.9. As graphically illustrated, the shift in the demand response to annual energy cost is significant negative across all three specifications, while the shift in the response to the purchase price is insignificant in columns (1) and (2) and significant positive in (3). Calculating the change in m using the formula in Equation 5.10 yields a positive value of 6.2 for $\Delta m^{salience}$ that is significant at 5% significance for columns (1) and (2). Δm in column (3) is of the same magnitude, albeit insignificant, as the power of the estimation is lower. Comparing the ratio before versus after the label revision in column (1), m is at -5.32 and significantly different from full consideration (H0: $m = 1$, H1: $m < 1$, p-value = 0.061) before. After the revision, the ratio shifts to a positive value of 0.89 and is not significantly different from full consideration (H0: $m = 1$, H1: $m < 1$, p-value = 0.737).

5.8 Conclusion

This paper investigates whether households in the lower tail of the income distribution consider lifetime energy cost when purchasing household appliances. Analyzing data of 20,601 purchase decisions from Germany's largest energy efficiency assistance program, I find that on average households do not rationally consider lifetime energy cost at conservative assumptions regarding discount rate and appliance lifetime. My analysis is based on demand response estimation at the regional market level, comparing the relative magnitude of responses to purchase prices and energy cost. As the ratio between the demand responses to energy cost and the purchase price indicates less than full consideration of energy cost, I conclude that households do not put the same weight on energy cost they put on purchase prices.

An important factor to explain this result is inattention to energy efficiency. Using the revision of the EU Energy Label that increased the salience for differences in energy efficiency between energy classes to participating households, I find that increased salience significantly increases consideration of energy cost. The finding indicates that inattention to energy efficiency matters for low consideration of energy cost. To test the relevance of financial constraints, I use exogenous variation in program subsidies to analyze how an increased budget for the appliance purchase affects consideration of energy cost. Using various changes in the federal subsidy and top-up subsidy state programs, I estimate a comparison of means, a two-way fixed effects and Difference-in-Difference designs. I do not find evidence for an effect on the consideration of energy cost in any of these estimations and conclude that financial constraints are not binding in this setting.

My results contrast findings in the valuation literature that find (close to) full consideration of energy cost in the general population for purchase decisions regarding household appliances, cars and heating systems, including studies that investigate income heterogeneities. This points to a strong heterogeneity in decision making of lowest-income consumers where significant under-consideration of energy cost is present. However, the choice of the optimal level of energy efficiency is important, in particular for lowest-income households as each Euro in energy savings tends to have a higher marginal value. Moreover, as energy expenses represent a larger share of income at the bottom of the income distribution, the burden from higher energy bills due to lower efficiency of durables is especially grave.

The findings have implications for the choice of policy instruments to increase energy efficiency if policymakers want to implement efficient but equitable policies. Carbon taxes, making up a portion of energy cost, are only effective if households consider energy cost in their purchase decisions. If consideration is low in the lower tail of the income distribution, the general efficiency of carbon taxes may not be inhibited strongly. However, the distributional consequences for low-income households need to be carefully considered. Households that do not consider energy cost when purchasing durables carry a double burden from higher energy and carbon cost. If these households do not choose the optimal level of energy efficiency at the time of purchase, they will have to carry the additional burden over the entire lifetime of the durable.

The analysis of factors that drive the result of under-consideration of energy cost has implications for policy instruments to support households in their choices. If inattention to energy efficiency is a main issue, policymakers need to focus on providing assistance to correct bias in decision making. The analysis shows that the design of simple information provision instruments, such as labels, can help to correct households' decisions and mitigate inattention. Though financial constraints do not seem binding for the consideration of energy cost in this setting, they may matter in purchases of larger energy-using durables with a higher investment cost.

Implications for future research arise from questions of external validity of the findings: to what extent can the results be transferred to purchases of larger durables? While the inefficiency premium in energy cost for household appliances can be significant, choosing a suboptimal level of energy efficiency in larger durables has more severe economic consequences, such as the extra gasoline cost for a less efficient car or the premium on heating bills of poorly insulated dwellings. Inattention to energy efficiency may be a less relevant factor for larger investments, as the cost of gasoline or heating bills are more salient than the contribution of individual household appliances to the electricity bill. However, the relevance of financial constraints likely increases for those investments that carry a much larger investment cost. Further research needs to assess how these factors impact consideration of energy cost for different classes of durables.

Chapter 6

Impacts of ownership changes on emissions and industrial production: Evidence from Europe

Coauthor(s): Robert Germeshausen

Abstract: Firm ownership is a major determinant for the economic performance of firms, and emissions of pollutants are often by-products of industrial production. We investigate the impact of ownership on pollutant emissions of firms and their industrial facilities in Europe jointly with their output, productivity, and other key economic outcomes. To disentangle the influence of ownership from other firm characteristics, we analyse the effects of ownership changes in an event-study approach. We find that facilities and firms do not change their emissions and emissions intensity if they remain in operation after a change in ownership. Firms that shut down after acquisition strongly reduce their emissions via reductions in output. The reductions cannot be attributed to the ownership change as they already start before acquisition. There is no evidence for transfers in pollution abatement technologies between target and acquiring parent company. Overall, we do not find environmental benefits from ownership changes.

Publications: This chapter has been accepted for publication at the *Journal of the Association of Environmental and Resource Economists*. A previous version of this chapter is available as ZEW Discussion Paper No. 23-016.

Acknowledgments: We thank Kathrine von Graevenitz and two anonymous referees at JAERE for valuable comments. This paper also benefited from feedback by participants in a seminar at ZEW Mannheim, in a session at the VfS Annual Conference 2022 as well as in INTEGRATE project meetings. Funding by the German Ministry of Education and Research (BMBF) through INTEGRATE (Project number 01LP1928B) is gratefully acknowledged. We thank Lea Helms and Vincent Kühn for outstanding research assistance.

6.1 Introduction

Corporate ownership affects the economic performance of firms, such as productivity (e.g., Commander and Svejnar, 2011; Li, 2013) and innovation activity (e.g., Aghion et al., 2013; Clo et al., 2020). Ownership can affect knowledge transfer and management practices within firms (e.g., Burstein and Monge-Naranjo, 2009; Alcacer and Zhao, 2012; Bloom et al., 2013) as well as internal goal setting (e.g., Shleifer, 1998). Ownership changes, e.g., through mergers and acquisitions (M&A), often influence production and investment decisions. Furthermore, these changes reallocate funds across firms, thereby impacting even aggregate economic outcomes (David, 2021). However, against the background of climate change and pollution as two major societal challenges, it is unclear how these changes in firms' economic performance and overall economic outcomes translate to environmental impacts of firms, such as the emission of pollutants.

In this paper, we analyze the impact of ownership changes on emissions of industrial facilities and firms in Europe jointly with their output, productivity and other key economic outcomes.[1] In 2019, the industrial sectors had a share of about 16 percent of Europe's total employment (Eurostat, 2022a) and about 18 percent in its gross domestic product (GDP) (Eurostat, 2022b), but were also responsible for a substantial share in Europe's pollution: about 48 percent of total greenhouse gas emissions (EEA, 2022), 28 percent of nitrogen oxide emissions, and 81 percent of sulphur oxide emissions (EEA, 2021). These numbers underline the importance of the industrial sectors for economic but also environmental outcomes in Europe. Also in 2019, around 17 500 M&A deals (Thomson Reuters, 2019a) with a volume of €991 billion were made in the European economy (Thomson Reuters, 2019b).

To shed light on the impact of ownership changes on emissions and economic performance, we use ZEW's ME-FINE dataset, which combines emission information of industrial facilities from the European Pollutant Release and Transfer Register (E-PRTR) and financial indicators of firms from Bureau van Dijk's Orbis database. Our sample includes about 6,000 industrial facilities[2] associated to 4,600 firms[3] in the EU15[4] plus Hungary and Norway from 2007 to 2016.[5]

[1]Industrial refers to facilities and firms active in the manufacturing and energy supply sectors.

[2]Facility is the reporting unit in EPER/E-PRTR and describes "one or more installations on the same site that are operated by the same natural or legal person" (Regulation (EC) No 166/2006).

[3]A firm is the observational unit in Orbis defined by the Bureau van Dijk identifier. In our sample, the mean and median number of facilities per firm are 1.4 and 1, respectively.

[4]Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, and the United Kingdom.

[5]2007 is the first year of the emissions reporting in E-PRTR. ZEW's ME-FINE dataset also includes emissions data for 2001 and 2004 from the E-PRTR's predecessor, the European Pollutant Emissions Registry (EPER). We restrict the sample to the time period from 2007 to 2016 for a more

Since ownership changes and firm decisions, such as input and output choices, are likely endogenous, we use an event study design, exploiting variation in the timing of ownership changes among all units that experience a change in ownership during our observation period. In our sample, 47 percent of facilities and 43 percent of firms experience at least one ownership change between 2007 and 2016. Since ownership changes occur at different years across units, we address treatment effect heterogeneity by applying the estimator proposed by Sun and Abraham (2021) in addition to conventional two-way fixed effects models. We use only the within-variation in facilities' and firms' emissions and ownership status by including individual and a variety of year fixed effects.

In the context of large polluting industrial facilities in the European Union (EU), we investigate the effect of ownership changes on firms' and their industrial facilities' total emissions. On average, emissions decrease both in the lead up to and following an ownership change. The decrease after acquisition is at about 46 percent at the facility and at about 55 percent at the firm level. We differentiate between firms and facilities that remain in operation and those that close down in the years after an ownership change. Firms and facilities that remain in operation have insignificant pre-trends and neither change their emissions nor their emissions intensity of output after changing owners. Firms and facilities that close down after acquisition strongly reduce their emissions via output reductions. However, the falling trend in emissions and output starts already before the ownership change, so that we cannot causally attribute the emissions reductions to the acquisition.

Aggregate emissions in the acquiring parent company increase after the target joins but emissions intensity remains constant. The acquisition of the new facilities does not affect either emissions or emissions intensity of other facilities in the acquiring firm. This indicates that no transfer in pollution abatement technologies takes place between the target and the acquiring parent company. However, the acquisition appears to provide positive spillovers in terms of increases in productivity, operating profits and intangible fixed assets to other facilities in the acquiring parent company. In sum, acquisitions seem to be a zero-sum game that neither harms nor benefits the environment.

Our paper contributes to the large literature on the importance of corporate structure and ownership for firm performance and to the smaller literature on the effects of ownership on environmental performance. In a study closely related to ours, Jacqz (2021) finds that newly acquired facilities in the United States reduce their (toxic) emissions to the air, mainly driven by operational changes. Similar to our study, she uses an event study design. Two further US studies provide evidence that the ownership structure

comprehensive coverage and consistent definitions of pollutant emissions in those countries over time.

of facilities seems to matter for their emissions level: Grant and Jones (2003) compare emissions by subsidiaries and non-subsidiaries in the US and find that the former facilities pollute significantly more. Akey and Appel (2021) study how the degree of parent company liability affects pollution by subsidiaries in the US; they find that stronger liability protection for parents leads to increases in toxic pollutants emitted by subsidiaries. Several studies look at outcomes other than facility-level emissions related to pollution: Aden et al. (1999) study pollution abatement expenditures of foreign- and domestically-owned manufacturing plants in Korea. They find that domestically-owned plants spend more on abatement equipment than plants with some level of foreign ownership. Conversely, Albornoz et al. (2009) find that foreign direct investment (FDI) has a positive effect on the implementation of environmental management systems by Argentinean manufacturing firms. Ning and Wang (2018) find that FDI reduces local pollution intensity via spillovers at the prefectural city level in China.

The effect of mergers and acquisitions, specifically FDI, on other outcomes of firm performance has been studied more extensively. Most studies find a positive effect of foreign ownership on firm productivity (Javorcik, 2004; Haskel et al., 2007; Arnold and Javorcik, 2009; Newman et al., 2015). However, Aitken and Harrison (1999) find negative productivity spillovers on domestically-owned firms so that the net productivity increase from FDI is small. Harris and Robinson (2002) find that foreign-owned companies purchase the most productive facilities but productivity declines after the acquisition. Wang and Wang (2015) find no additional gains from FDI; both foreign and domestic acquisitions increase productivity of the target facilities equally.

The effect of acquisitions on output and employment depends on the context. Siegel and Simons (2010) find that Swedish firms in the manufacturing sector reduce output and employment after acquisition, while Wang and Wang (2015) find that foreign, but not domestic, acquisitions increase the output and employment of Chinese target firms. Also Arnold and Javorcik (2009) find a positive effect of foreign acquisitions on employment in Indonesian manufacturing firms. Conversely, Li (2013) finds that employment drops in US facilities after acquisition. Chen (2011) compares the effect of foreign and domestic acquisitions on target firms' profits and finds FDI to increase profits more compared to domestic acquisitions.

We contribute to the literature being the first to provide evidence on the role of ownership changes for emissions of firms in Europe, jointly with the impact of ownership changes on a wide range of firms' economic performance indicators. We use a novel data set combining information from Orbis and the E-PRTR. This enables us, as compared to single-country studies, to extend our analysis to a major economic region with a wide range of countries, allowing us to draw broader conclusions less dependent on country-specific peculiarities. Our findings differ from Jacqz (2021) who finds evidence for operational changes in newly acquired facilities in the US that reduce toxic emissions. Both studies cover a similar time period and the context of large facilities reporting to the Toxic Release Inventory is similar to our setting where large facilities report to the E-PRTR. The institutional context provides the most apparent difference between the settings studied. However, more evidence needs to be generated to provide a clear picture of the differences in the EU and US context and which factors contribute to the different result. Contrary to findings by Aden et al. (1999), Albornoz et al. (2009), and Ning and Wang (2018) that foreign ownership (in the form of FDI) impacts pollution abatement expenditures in Korea, the implementation of environmental management systems in Argentina and local pollution intensity in China respectively, we do not find that foreign acquisitions differ from domestic acquisitions in their impact on pollutant emissions and emissions intensity. However, in this comparison, countries, time periods and institutional settings vary widely so that it is unclear which factors drive the difference in results. Furthermore, observing outcomes at different aggregation levels, such as at the facility, firm and parent company level, we can distinguish between those three levels of aggregation and shed light on potential reallocation effects emissions and production indicators across facilities and firms within the parent company, and assess the impact on productivity and profits. Our paper is the first to provide evidence on reallocation effects of emissions after acquisitions which provides evidence on environmental technology transfers also from target to acquiring parent company, in addition to transfers from acquiring firm to target. In that, we go beyond the analysis by Jacqz (2021) who limits her analysis to the facility and firm level and does not consider the impact of acquisition on the parent companies.

The remainder of this paper is structured as follows: The data is described in Section 6.2. Our empirical strategy for the analysis of ownership changes is outlined in Section 6.3. Section 6.4 presents our results and Section 6.5 concludes.

6.2 Data

Our main data source is ZEW's ME-FINE data set which combines emissions data from the EPER/E-PRTR and financial information from Orbis (Germeshausen et al., 2022).[6] ME-FINE includes firms in the manufacturing and energy supply sectors (NACE Rev. 2: 10 - 35) in the EU-15 plus Norway and Hungary and covers about 70 percent of observations reported in EPER/E-PRTR in those sectors and countries. We use observations from the period 2007-2016[7], covering 6,097 facilities and 4,669 firms. For

[6] The documentation of the dataset also includes an index decomposition at the sector level for the period 2007-2016 that separates scale, composition and technique effects on the evolution of total emissions.

[7] From 2007 on, the E-PRTR reports information on pollutant emissions annually. The EPER is the predecessor which reports pollutant emissions for the years 2001 and 2004.

this period, ME-FINE covers about 87 percent of total E-PRTR observations in these sectors and countries. Furthermore, we add ownership links between firms and their parent company as reported in Orbis.

We divide the data set into three levels: facilities, firms and parent companies. At the facility level, facility-year observations contain information on reported emissions, on the associated firm and the parent company as well as the sector code (NACE Rev. 2). At the firm level, firm-year observations contain information on reported emissions (aggregated over all their E-PRTR facilities), financial indicators, the parent company, and the sector code. At the parent company level, parent company year observations contain information on reported emissions aggregated over all their E-PRTR facilities and financial indicators aggregated over all their firms with E-PRTR facilities. Reported emissions during our observation period stem from E-PRTR. Reporting emissions is mandatory for facilities in specific economic sectors that exceed capacity and pollutant-specific thresholds. These thresholds are set such that about 90 percent of the emissions of each of the 91 pollutants in E-PRTR is covered.This means that our aggregation at the firm and parent company level also only contains facility-level observations that release pollutant amounts beyond the threshold. Further information on the reporting procedures and data quality is provided in Appendix A.

Table 6.1 provides summary statistics on the outcome variables in our estimation sample at the facility, firm and parent company level. This sample includes only facilities and firms with one ownership change from 2007 to 2016. Total emissions is an aggregated measure which sums physical emission quantities over all pollutants reported to E-PRTR, whereby the quantity of each pollutant is divided by its pollutant-specific reporting threshold. Emissions intensity at the firm level scales total emissions by operating revenues in thousand euro (EUR). Operating revenues are deflated by two-digit sectoral (NACE Rev. 2) producer price indices from Eurostat. To obtain firm- and time-specific values for total factor productivity, we estimate a value added production function using firm investment as a proxy variable following Wooldridge (2009).[8]

[8]Total factor productivity estimates are highly correlated to estimates obtained by applying the methods by Olley and Pakes (1996) and Ackerberg et al. (2015). However, in the case of Ackerberg et al. (2015) the coefficient of capital input is negative. Therefore, we use the estimates obtained from following Wooldridge (2009) as reference.

Table 6.1: Summary statistics for final sample, 2007 - 2016

	N	Mean	St. Dev.	Min	P25	P75	Max
Variables at the facility level							
Total emissions	11,819	118.6	1,247.8	0.0	1.3	19.3	39,926.4
Variables at the firm level							
Total emissions	6,979	83.5	1,137.9	0.00	1.0	16.7	64,302.0
Operating revenues ('000 EUR)	6,210	838,806	9,204,588	0	20,411	227230	261,279,167
Emissions intensity	5,783	0.069	1.101	0.000	0.000	0.016	70.162
Total factor productivity	3,286	9.3	0.7	5.4	9.0	9.7	14.9
Number of employees	5,939	1,139	8,456	0	49	508	195,826
Tangible fixed assets ('000 EUR)	6,167	299,240	3,931,596	0	4,261	73,112	163,911,425
Labor expenditures ('000 EUR)	5,671	71,417	596,300	0	2,306	28,962	14,189,731
R&D expenditures ('000 EUR)	305	51,623	147,706	0	0	10,625	978,666
Intangible fixed assets ('000 EUR)	5,520	72,931	701,664	0	6	2,316	15,685,382
Variables at the parent company level							
Total emissions	2,612	202.7	1,408.2	0.0	2.9	59.7	35,201.3
Operating revenues ('000 EUR)	2,584	965,397	3,289,382	0	23,731	628,181	54,484,828
Emissions intensity	2,136	0.080	0.767	0.0	0.001	0.026	21.390
Total factor productivity	1,514	7.3	0.7	2.0	7.0	7.6	10.2
Number of employees	2,584	2,175	9,671	0	49	1,132	198,980
Tangible fixed assets ('000 EUR)	2,584	38,6025	1,561,469	0	5,225	220,251	25,848,393
Labor expenditures ('000 EUR)	2,583	125,575	658,641	0	1,704	56,949	12,188,843
R&D expenditures ('000 EUR)	2,584	13,818	170,347	0	0	0	453,9012
Intangible fixed assets ('000 EUR)	2,584	117,280	1,138,166	0	0	6,606	33,422,925

Notes: Total emissions is the sum over the quantities of all pollutants each divided by its reporting threshold. Emissions intensity refers to total emissions divided by deflated operating revenues at the firm level, and to total emissions divided by deflated operating revenues multiplied by 100,000 at the parent company level.

We define a change in ownership for both firms and facilities as a change in their parent company from one year to the next.[9] In total, we observe 2,621 changes of firm ownership in the sample. This corresponds to 1.3% of all M&A events recorded in the Zephyr Database for the EU15 plus Hungary and Norway for the period 2008-2016 (Zephyr Database, 2023). 978 firms experience one change, while in total 655 firms experience multiple ownership changes. 2,697 facilities experience at least one change in ownership, of which 1,525 change only once.[10] In our analysis, we only consider firms and facilities with one ownership change event.

There is considerable heterogeneity in the distribution of ownership change events over sectors and countries. In absolute terms, we observe most ownership changes in

[9]That means that we do not observe ownership changes according to our definition that happened in 2007, the first year in our sample, since we do not observe ownership in 2006.

[10]For facilities, we count changes in the global ultimate owner as indicated by Orbis. If the global ultimate owner is unknown, we assign the associated firm as the global ultimate owner.

German, French and Spanish firms. Scaling the number of observed changes by the absolute number of observations for each country in our sample, heterogeneity is much less pronounced and, in relative terms, we observe most changes in Greek, Portuguese, German and Luxembourgian firms (see Figure A6.1 in the Appendix). The absolute number of ownership changes is highest in NACE sectors 20 (Manufacture of chemicals and chemical products), 23 (Manufacture of other non-metallic mineral products) and 24 (Manufacture of basic metals). In relative terms, the share of ownership changes is highest in sectors 27 (Manufacture of electrical equipment) and 33 (Repair and installation of machinery and equipment; see Figure A6.2 in the Appendix). The distribution of ownership change events over the years is more uniform. Both the absolute number and the percentage of changes is higher in 2008, but remains almost stable thereafter (see Figure A6.3 in the Appendix).

Total emission reports are unbalanced in our sample. We define facilities as active in years in which they report a positive amount of emissions. In years for which facilities do not report any emissions, facilities could either have closed down or they could have emitted pollutants below the reporting thresholds.[11] We consider facilities with missing emission reports for a facility-year observation as active as long as the facilities report again in a later year in the sample. If facilities do not report again until the last year in our sample we assume they have closed down. To proxy their exit in the data, inactive facilities remain as zero-values in the sample for up to four years (at the latest until 2016) after their last reporting year, similar to the approach used by Jacqz (2021).[12] At the firm level, we apply the same procedure. Since firms' emissions are aggregated over all their facilities, we consider a firm to have exited only if none of its facilities reports again in a later year during the sample period. The largest share of facilities reports from 2007 on, only a small share of facilities enters the sample in later years. Later entries are relatively evenly distributed across years. The largest share of facilities reports until 2016, and similarly earlier exits are rather uniformly distributed.[13] At the firm level, most entries are recorded in the first two years and least entries in the later years. The majority of firms in our sample survive until the end of our observation period. The number of firm exits varies over time.[14]

Our sample consists of the overlap of E-PRTR and Orbis. Given the emissions reporting threshold in E-PRTR, we observe emission reports from rather large firms. With

[11]Since reporting positive emission amounts is censored below the threshold (there are no reported emissions below the pollutant-specific threshold), we investigate the impact of this censoring by considering two different imputation strategies, i.e., either imputing missing values by zero or by the threshold value, as robustness checks.

[12]Results are qualitatively similar if we replace zero-values with the pollutant-specific threshold at the facility level since facilities could still emit up to this amount without reporting obligation. Using both approaches provides us with an upper and lower bound of emissions.

[13]See Figure A6.4 in the Appendix.

[14]See Figure A6.5 in the Appendix.

respect to Orbis, its coverage differs across the globe due to different national reporting requirements and firm structures. Bajgar et al. (2020) find that firms in Orbis are rather large, old and productive. While these characteristics of E-PRTR and Orbis facilitate the assignment of E-PRTR facilities to Orbis firms in the ME-FINE data set, it has to be considered in the interpretation of our results. Our final estimation sample is not necessarily representative of the overall economy but focuses on rather large industrial facilities and firms.

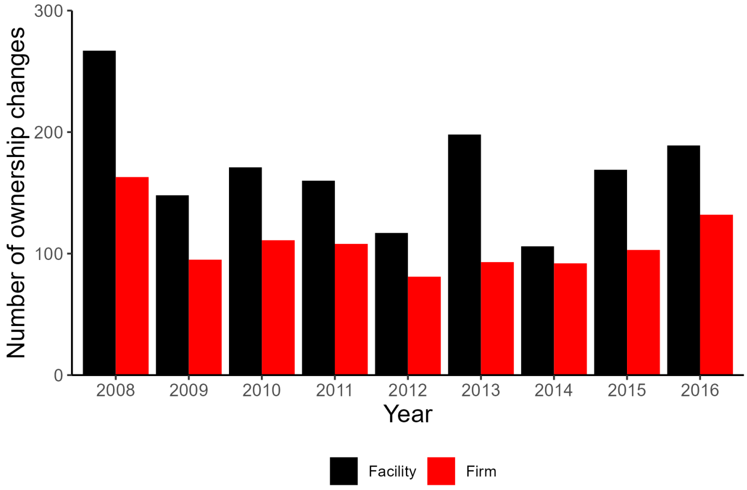
6.3 Empirical strategy

We aim to identify the effect of a change in ownership (parent company change) on pollutant emissions and economic outcomes of firms and their facilities. In our sample, we observe 978 firms and 1,525 facilities whose parent company changes once during the period 2007-2016. Our empirical strategy relies on fixed unit characteristics at the facility and firm level which allows us to use only within-unit variation to identify the effect of ownership changes. The events are distributed over 9 years so that treatment adoption – change in ownership in our case – is "staggered". Our method is closely related to Jacqz (2021) who investigates a similar question in the US context.[15]

For our event study of ownership changes, we use the Sun and Abraham interaction-weighted estimator that is robust to treatment effects heterogeneity (Sun and Abraham, 2021). The estimator interacts treatment group and relative time dummies which are then aggregated to obtain the average treatment effect for the treated for each period. In our setting, we have nine treated groups of units (firms and facilities) whose parent company changed in the respective year 2008 to 2016. Figure 6.1 shows how the ownership change events are distributed over the sample period for facilities and firms. Each treatment group has observations in up to 10 periods relative to the treatment period.

[15]Jacqz (2021) uses plant-level data from the EPA's Toxic Release Inventory for the period 2001-2019 to investigate the effect of corporate acquisition on facility-level air pollution and its firm level distribution.

Figure 6.1: Distribution of ownership changes over time at facility and firm level



Notes: This figure shows number of ownership changes at the facility and firm level in each year for facilities and firms with only one ownership change during our observation period.

Based on Sun and Abraham (2021) the regression for our event study is:

$$Y_{it} = \alpha_i + \lambda_{ct} + \mu_{st} + \sum_{e \notin C} \sum_{l \neq -1} \delta_{el}(1\{E_i = e\} * D_{it}^l) + \varepsilon_{it},$$

where the outcome Y_{it} is aggregated emissions or economic outcomes of unit i in year t. D_{it}^l indicates the relative period of the observation, unit i being l periods away from year of treatment E in year t, and $1\{E_i = e\}$ indicates the treatment group that unit i belongs to. The specification interacts these indicators, but omits interactions with the last group of units with ownership change in 2016 because these units do not have a not-yet-treated control group, and with the reference period $l = -1$ to avoid issues of multicollinearity. δ_{el} represents the group-specific average treatment effect on the treated. α_i, λ_{ct}, and μ_{st} capture unit-specific, country-year and sector-year fixed effects, respectively.

To form the interaction-weighted estimator, δ_{el} is weighted with sample shares of each group in each period $Pr\{E_i = e|E_i \in [-l, T-l]$. The resulting weighted average estimate normalized for the number of periods after treatment g is then:

$$\hat{v}_g = \frac{1}{|g|} \sum_{l \in g} \sum_{e} \delta_{el} \hat{Pr}\{E_i = e|E_i \in [-l, T-l].$$

We employ the Sun and Abraham estimator to identify the effect of an ownership change event on total emissions, emissions intensity and economic outcomes of firms and on total emissions of facilities. We apply the inverse hyperbole sine transformation to the outcome so that we can interpret the effects in percentage changes.[16] Our preferred specification estimates the ownership change effect using 4 leads and lags around the treatment year.[17] We include only treated firms or facilities so that the later treated units act as controls for the earlier treated units. Firms or facilities with more than one ownership change event during the period 2007-2016 are excluded. We cluster the standard errors at the respective unit level.

The main identifying assumptions for the event study estimation to produce an unbiased effect of ownership changes on facility and firm indicators are, first, for the control group to have parallel trends in the outcomes of interest in the absence of treatment, and, second, that treatment timing is random, i.e., it is not associated with firm characteristics that also affect outcomes of interest. The first assumption of parallel trends connects to the empirical challenge of finding a valid counterfactual for facilities and firms that are acquired. Firms and facilities with an ownership change event may systematically differ from firms and facilities that keep their parent company over the entire period. Moreover, firms and facilities with more than one event may also be systematically different. We check empirically whether the groups of firms differ systematically in observable characteristics. We find small differences in capital, long-term debt, total emissions, employment and intangible fixed assets (see Figure A6.6 in the Appendix). We deal with this issue by omitting firms and facilities with no ownership change and those with multiple ownership changes over the sample period. We inspect the pre-treatment coefficients in the event study to check if pre-trends are parallel.

The second assumption of treatment timing being unrelated to facility and firm characteristics cannot be tested empirically. We argue that the assumption is reasonable in our context:[18] M&A processes usually take a significant amount of time and it is ex-ante not predictable whether ownership will change within the same year or with considerable delay in the negotiations. Moreover, in our sample a significant share of the acquisitions happens in bundles where several facilities or firms change from one parent company to another jointly in the same year. Acquisition decisions taken at an aggregate level tend to be more independent of the performance of individual firms and even more so of facilities.[19] In addition, we provide suggestive evidence that the

[16] We use the hyperbole sine transformation instead of the natural logarithm to deal with zero values when facilities do not report emissions or economic indicator values are equal to zero.

[17] We bin the first and the last lag following Schmidheiny and Siegloch (2020). Hence, we assume that effects remain constant before and after these years, respectively.

[18] Other studies that investigate the effect of M&A on firm-level outcomes and use variation in timing of ownership changes in event study settings are Jacqz (2021) and Blonigen and Pierce (2016).

[19] Of all firms that change their owner once during the study period, at least 26% are acquired in bundles of two or more firms. We can only provide this lower bound share from our data as we do not observe firms not included in the E-PRTR which are potentially also part of bundle deals but whose pollutant emissions are below the E-PRTR reporting thresholds.

assumption appears to hold in our context by testing whether observed firm characteristics provide any predictive power for the timing of ownership changes. Except for intangible fixed assets, we do not find any observed firm characteristics to significantly predict treatment timing (see Table A6.1 in the Appendix).

Firms and facilities with an earlier change in ownership could also differ from firms with later changes if the reasons for ownership changes differ over time, e.g., via the financial crisis which had its strongest impact at the beginning of the sample period. Similarly, merger waves could be sector-specific and their timing could differ across industries. Environmental policy regulation that came into force during the study period could additionally introduce a trend in emissions and emissions intensity over time.[20] We address these issues by including sector-year and country-year fixed effects.

Anticipation effects are another threat to identification if the prospect of a change in ownership affects reported emissions and economic outcomes of firms or facilities before an acquisition. If the effect of a change in ownership manifests through a change in management practice or a technology transfer, the effect is implausible to affect emissions before an acquisition. It could however be advantageous for firms in a merger process to play down their emissions in the negotiations and report lower emissions. On the other hand, firms could ramp up production and increase output to appear more profitable for potential investors. Such anticipation effects would be visible in the pre-treatment coefficients close to treatment. We do not find evidence for an anticipation effect. We find however significant pre-treatment coefficients for some of the outcomes at the firm level three to two years before an ownership change event. In these cases, we must be cautious to interpret the coefficient as an isolated effect of the ownership change since the coefficient may reflect also other differential trends.

Shocks that affect both emissions and the propensity for an ownership event of firms and facilities can also bias the estimate. If a positive demand shock leads parent companies to buy up promising firms that will expand in the coming years, the estimate of emission reductions will be biased downward. If a negative demand shock leads parent companies to sell low-performing firms which would otherwise have closed down, the change in ownership delays the closure so that the estimate will be biased upward. Arnold (2019) and Jacqz (2021) counter this source of bias by focusing on ownership changes of larger firms which are less affected by local demand shocks. Our sample mainly consists of

[20] Relevant environmental regulation that affects pollutants included in the E-PRTR is, first, the Large Combustion Plant EU Directive (European Parliament and European Council, 2001) that specified emission limits for SO_2, NO_x and dust from 2008 on. The regulation was binding mainly for plants in Southern and Eastern European economies that still operated on older technologies (European Environmental Agency, 2019). Second, the EU ETS (European Parliament and European Council, 2003) regulated CO2 emissions via a cap-and-trade system. However, the price per ton of CO2 was very low over the entire study period and the timing and strictness of the regulation was uniform across countries in the sample.

large firms as well.

6.4 Results

We first present results on the effect of ownership changes on our main outcomes of interest, emissions and emissions intensity (see Section 6.4.1). Second, we show how ownership changes affect output and production inputs (see Section 6.4.2). Third, we investigate spillovers of the ownership changes on total factor productivity, operating profits and intangible fixed assets (see Section 6.4.3). For each set of outcomes, we move from the more granular levels of observation, the facility and firm, to the aggregated level, the parent company. In section 6.4.4, we discuss implications of our results.

6.4.1 Emissions and emissions intensity

Total emissions of facilities decrease steadily after an ownership change. In the third year after a change, their emissions decrease by about 50 percent (see Figure 6.2). The point estimate of the average effect of an ownership change on total emissions at the facility level is at negative 37 percent (see Table 6.2).[21] The estimates using the Sun and Abraham (2021) approach that we report here are larger in absolute terms as compared to the two-way fixed effects estimates.[22] Results are similar when we impute missing emission values at the facility level with either zero or the pollutant-specific thresholds (see Table A6.2 in the Appendix).

A falling trend in emissions is visible already in the pre-treatment period before the ownership change, but the slope is less steep than in post treatment and none of the coefficients is significant. We still conduct a sensitivity test using the Rambachan and Roth (2023) approach. The test shows that the reduction in emissions after changing ownership adjusted for the pattern in pre-trends is significant as long as the deviation from parallel trends in the post-treatment period is of similar size or smaller than the maximum violation observed in the pre-treatment period.[23] Since the E-PRTR data does not provide information on output and other industrial indicators of industrial facilities, we cannot investigate the impact of ownership changes on production at the

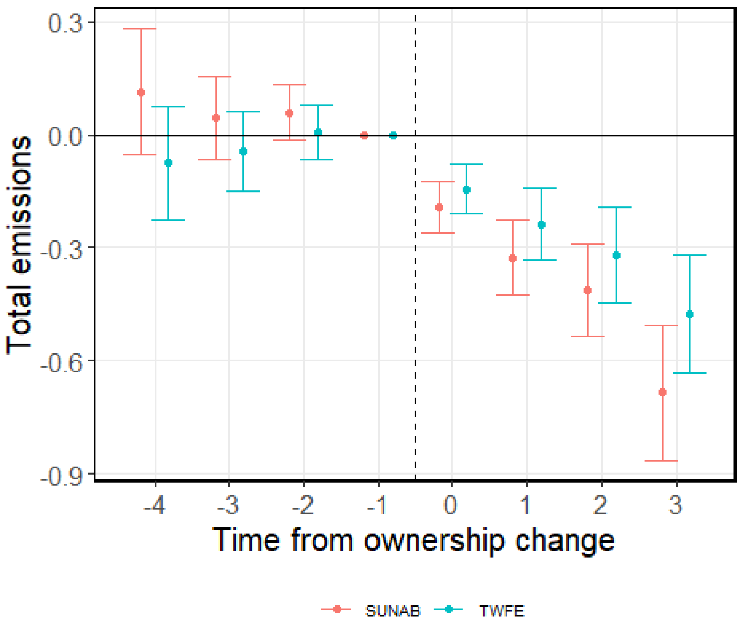
[21] For interpretation of the coefficients x from the log/ihs-linear specifications, we use the formula $(e^x - 1) * 100$ to retrieve the percentage change estimate.

[22] The Sun Abraham estimator is larger in magnitude for total emissions. For some other outcomes it is the other way around. A deviation of the results of the Sun Abraham estimator from those of the naïve two-way fixed effects estimator indicates that treatment effects vary across units and cohort effects over the years are not constant. Consequently, the results for the two-way fixed effects estimator will be biased. In this section, we focus on reporting and discussing results from the Sun Abraham estimator but show results for the plain two-way fixed effects estimates in the event study plots for comparison.

[23] See Appendix C for a short introduction to the method and Figure A6.63 for the results of the sensitivity check.

facility level.

Figure 6.2: Effect on total emissions at the facility level

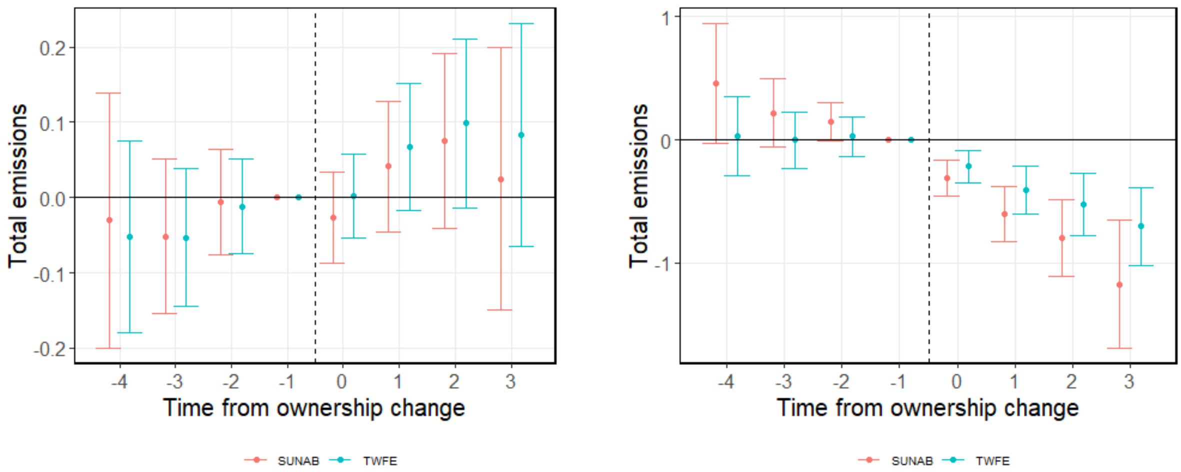


Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions. The inverse hyperbolic sine transformation is applied to the independent variable.

We separate the sample into facilities that remain in operation and facilities that are shut down after changing ownership to see how much of the reduction in emissions is due to shutdowns.[24] Figure 6.3 shows that emissions in facilities that remain in operation after changing owners do not change significantly while facilities that are shut down reduce their emissions strongly. The average reduction in exiting facilities is significant negative at 54 percent (see Table 6.2). A part of the facilities is only closed down after more than three years so that we do not see a reduction in the range of 100 percent. The reduction in emissions appears to be at least partially driven by a falling trend that already starts in the pre-treatment period so that the drop in emissions in exiting facilities may not be attributable to the ownership change. In a sensitivity check, we find that the effect is only significant if the deviation from parallel trends post treatment is half or less of the maximum violation observed in the pre-treatment period.[25]

[24] As explained in Section 6.2, we identify shutdowns of facilities and firms as those that stop reporting emissions in the observation period.
[25] See Figure A6.64 in the Appendix for the results of the sensitivity check.

Figure 6.3: Effect on total emissions at the facility level by subsample



(a) Remaining facilities

(b) Exiting facilities

Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions. The inverse hyperbolic sine transformation is applied to the independent variable.

Table 6.2: Aggregate effects on facilities

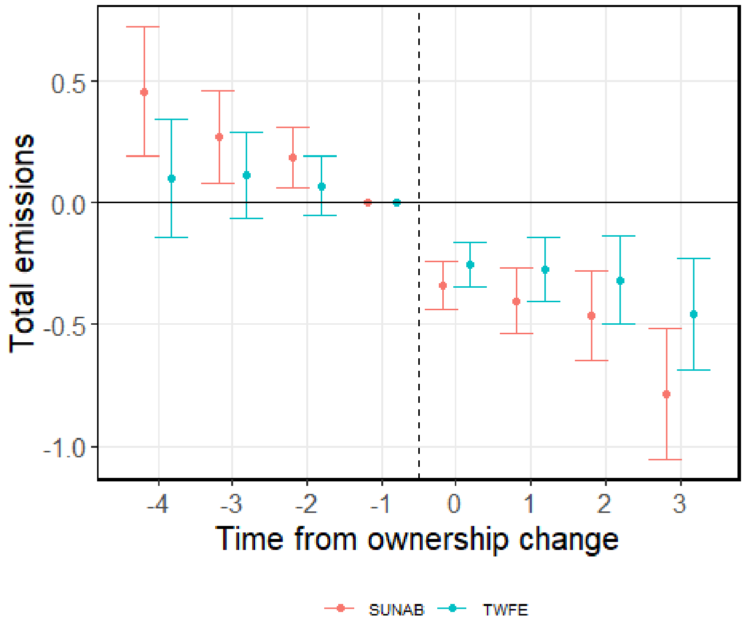
	Full sample			**Remaining facilities**			**Exiting facilities**		
Dependent variables	**ATT**	**SE**	**N**	**ATT**	**SE**	**N**	**ATT**	**SE**	**N**
Total emissions	-0.455***	(0.058)	10,624	0.025	(0.055)	6,416	-0.786***	(0.157)	4,208

Notes: The first column denotes the dependent variable with an inverse hyperbolic sine transformation. The table shows three separate event study regressions on the full sample of firms, on remaining firms as well as on exiting firms. For each regression, we report the point estimate of the aggregated effect of the event study following Sun and Abraham (2021) (ATT), the standard error (SE) and the number of observations (N). Standard errors clustered at the facility level are in parenthesis. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Up to this point, we assume that every ownership change has a similar impact, neglecting differences across types of ownership changes and across new owner characteristics. There is a large literature highlighting the role of foreign direct investments for firm performance. With respect to domestic and foreign ownership changes, we find similar effects to the average effects in the full sample of facilities presented above and confidence intervals overlap for all ownership type groups. These results suggest no large differences across different owner types (see Figure A6.7 and the more detailed explanation in the Appendix). Furthermore, we only find limited effect heterogeneity across sectors, with the exception of the sector manufacture of motor vehicles which experiences a larger reduction in total emissions compared to other sectors (see Figure A6.8 in the Appendix).

Facilities that reduce their emissions after a change in ownership are often part of a larger firm which potentially owns many industrial facilities. Investigating these firms for whom we have data on output and an abundance of financial and economic performance indicators allows us to also capture the effect of ownership changes on emissions intensity, production and spillovers, while this data is not available at the facility level.

Figure 6.4: Effect on total emissions at the firm level



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions. The inverse hyperbolic sine transformation is applied to the independent variable.

Firms reduce their total emissions on average by 42 percent which is in the range of the effect at the facility level (see Figure 6.4 and Table 6.3).[26] After the third year, emissions decrease by about 54 percent.[27] The decline in emissions that is observed after the ownership change, starts already in the years before the acquisition event. This may point to additional underlying trends not driven by the ownership change. We run a sensitivity check and find that the effect is only significant if the deviation from parallel trends post treatment is half or less of the maximum violation observed in the pre-treatment period.[28]

When investigating the samples of firms remaining in operation versus firms that are closed down, the results look similar to the facility results (see Figure 6.5 and Table 6.3). Firms that remain in operation do not change their emissions. The falling trend observed both before and after changing ownership appears to be solely driven by

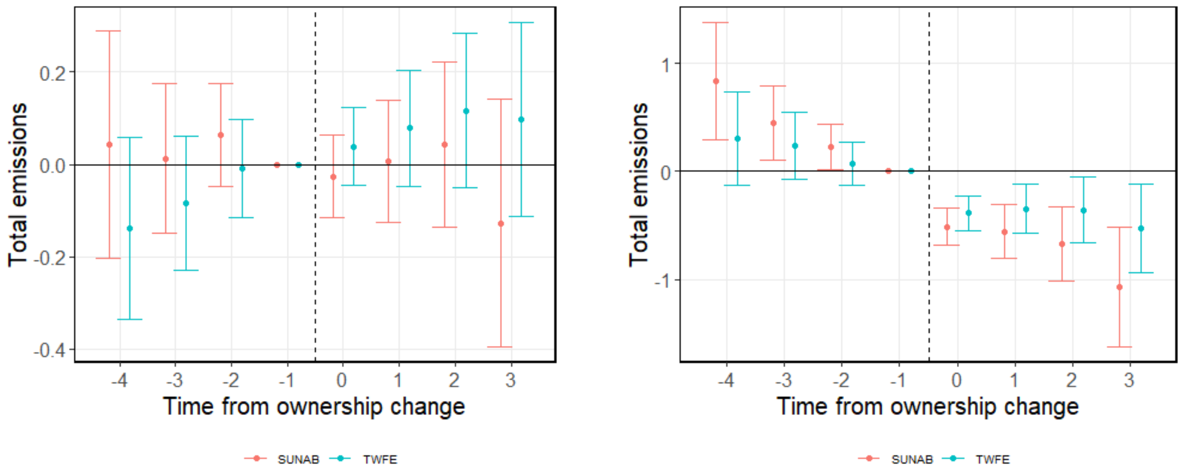
[26] As the firm level emissions are an aggregation of facility level emissions, firm results are a reweighing of facility results.

[27] The effect is robust to imputation of missing values with both zero-values and threshold-values (see Table A6.3).

[28] See Figure A6.65 in the Appendix for the results of the sensitivity check.

the firms that shut down subsequently. The sensitivity check shows that the effect for exiting firms is only significant if the coefficient is adjusted with a deviation from parallel trends that is less than half of the maximum violation observed in the pre-treatment period.[29]

Figure 6.5: Effect on total emissions at the firm level by subsample



(a) Remaining firms (b) Exiting firms

Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions. The inverse hyperbolic sine transformation is applied to the independent variable.

As in the case of facilities, we do not find differential effects based on the type of ownership change, i.e., from domestic to domestic, domestic to foreign, foreign to domestic or foreign to foreign owner, or across sectors (see Figure A6.9 and Figure A6.10 in the Appendix).[30]

After a change in ownership, facilities and firms appear to reduce their emissions through shutdowns of firms and their facilities that were already reducing emissions before acquisition, while facilities and firms that remain in operation do not change emissions. But how does their integration in the new parent company affect aggregate environmental performance at the parent company level and other industrial facilities within the new parent company? To shed light on this question, we aggregate all industrial facilities reporting to the E-PRTR at the parent company level for each year. Then, we re-run the event studies for the subset of parent companies that have already owned industrial plants before they acquired a new one and acquired new industrial facilities only once.[31] On average, total emissions of parent companies increase significantly and

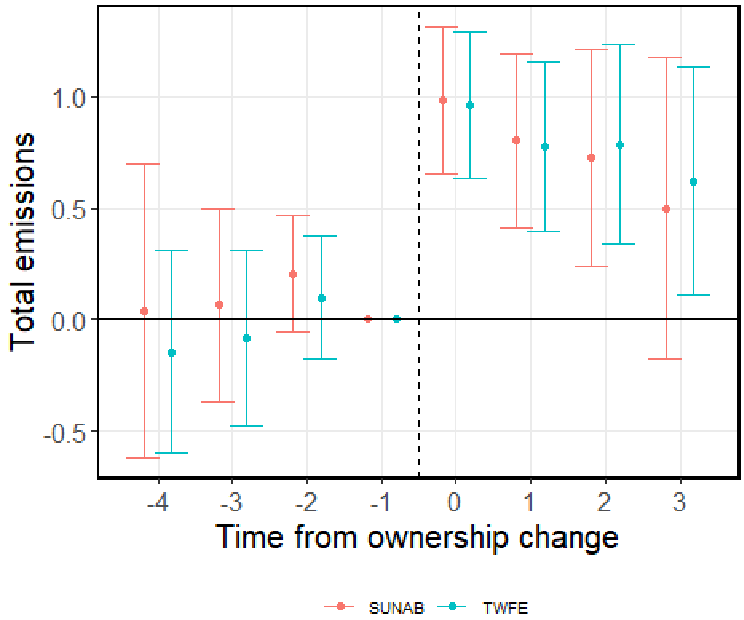
[29]See Figure A6.66 in the Appendix for the results of the sensitivity check.

[30]Furthermore, we explore whether effects differ for firms with one vs. many facilities and with above vs. below median parent company emissions. We do not observe any differential effects for these groups. Results are available upon request from the authors.

[31]The ownership change event starts in the year in which the parent company acquires a new

strongly after acquiring a new industrial facility (see Table 6.4). Furthermore, there is a downward trend over time after the change in ownership (see Figure 6.6), mirroring the results at the facility and the firm level. For parent companies that do not shut down their target after acquisition, the results are very similar (see Figure A6.11 in the Appendix).[32]

Figure 6.6: Effect on total emissions at the parent company level



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

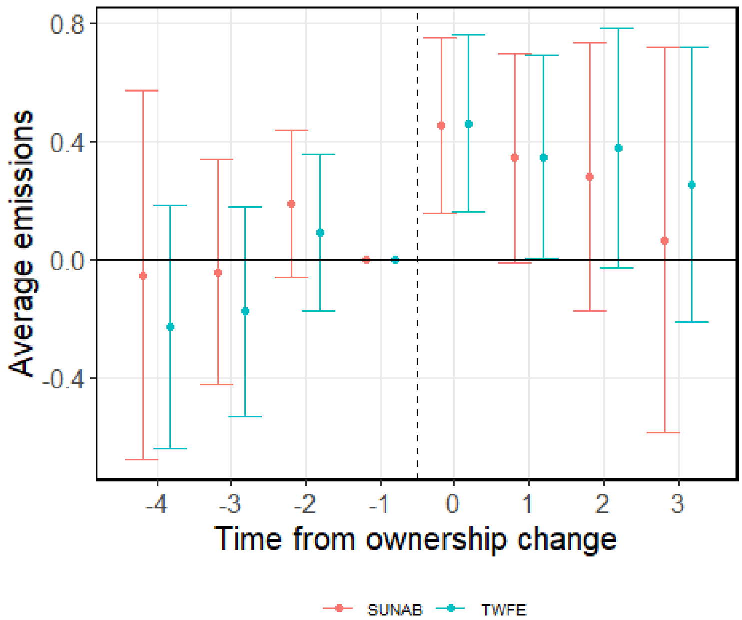
We also estimate the effect on average emissions at the facility level over all facilities in the parent company to see whether the newly acquired facility differs from the other facilities. The effect on average emissions per facility is positive, albeit insignificant after the first year, with a falling trend (see Figure 6.7). This indicates that emissions of the newly acquired facilities are not significantly above average. For the subsample of parent companies that do not shut down their target after acquisition, the increase after the ownership change is smaller and also insignificant except for the year in which the ownership change happens (see Figure A6.12 in the Appendix). To trace out whether any reallocation in emissions happens across new and old facilities within a parent company after acquisition, we also look at the effect of ownership changes on the other

facility. In the event studies on the parent company level, we only include parent company and year fixed effects since we cannot unambiguously assign countries and sectors to the parent company.

[32]For parent companies that acquire more than one facility in the same year, we consider them in this sample if they do not shut down any of their targets. We do not run a separate regression for parent companies that close down at least one of their targets as the sample gets too small and estimation too imprecise.

facilities that were part of the acquiring parent company before the acquisition. For these facilities, emissions do not change after acquisition (see Figure 6.8 and Table 6.5).[33]

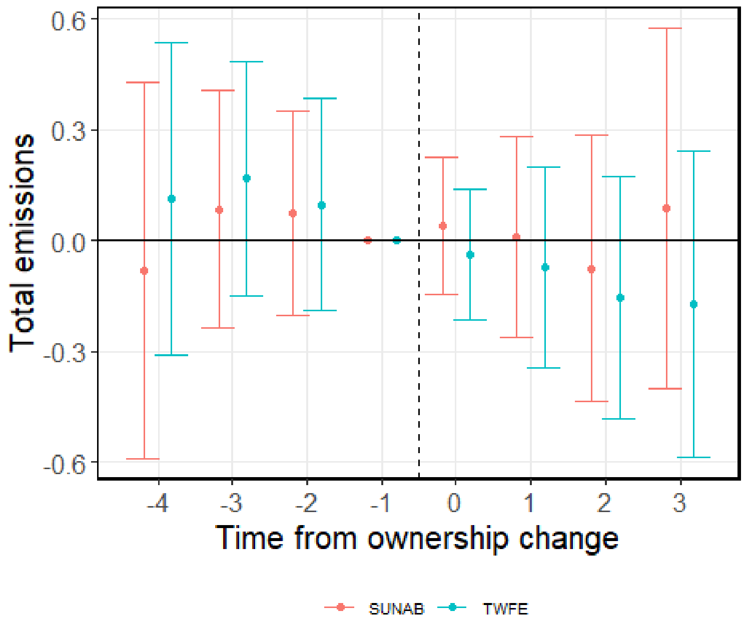
Figure 6.7: Effect on total emissions per industrial facility at the parent company level



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions per industrial facility at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

[33]We do not run separate regressions for the sample of other facilities in parent companies that do not close down their acquisitions versus parent companies that shut down acquisitions as the sample size gets too small.

Figure 6.8: Effect on total emissions in other industrial facilities of the acquiring parent company



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions per industrial facility at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

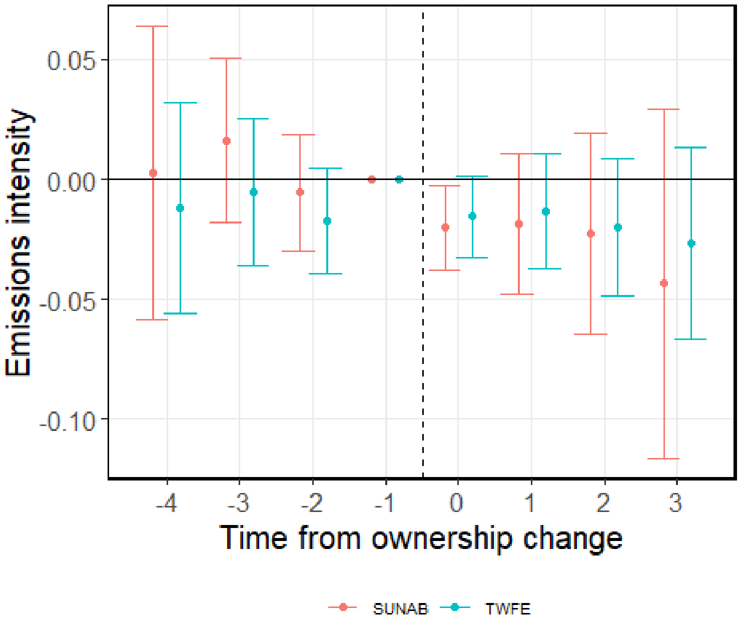
At the firm level, we have data available to estimate the effect of ownership changes on emissions intensity (emissions over operating revenues). Firms' emissions intensity seems relatively unaffected. A falling trend is observed in the post-treatment period but the magnitude of reduction is negligible. Only the effect in the year of ownership change is significant while all other years and the aggregate effect are insignificant. The small decrease appears to be solely driven by exiting firms. For firms that remain in operation, emissions intensity remains virtually constant (see Figure A6.13 in the Appendix).[34]

Also for the acquiring parent company both at the aggregated level and per facility, emissions intensity does not change as the new facility joins the parent company and the results do not differ considerably for the sample of only parent companies that do not shut down their targets after acquisition (see Figures A6.14 to A6.17 and Table 6.4). That means that the newly acquired facilities' emissions intensity is not significantly above average. Finally, for other facilities in the acquiring parent company emissions

[34] Additionally, we calculate how emissions in firms with an ownership change had evolved if emissions intensity had remained constant and only output had changed (scale effect) and how emissions would have evolved if output had remained constant and only emissions intensity had changed (technique effect) comparing output and emissions two years before an ownership change versus two years after. Total emissions increased by 13 percent. Via the isolated effect of output expansion (scale), emissions would have increased by 65 percent and, via a reduction in the emissions intensity, emissions would have decreased by 32 percent.

intensity remains unchanged as well after the new facility joins the group (see Figure A6.18 in the Appendix). Consequently, we do not find evidence for a technology transfer between the acquired facility and other facilities in the parent company that would affect environmental performance.

Figure 6.9: Effect on emissions intensity at the firm level



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on the emissions intensity, i.e., total emissions scaled by deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

In summary, we find that emission reductions observed for the full sample are driven by exiting facilities and firms that decrease emissions already before ownership changes. For facilities and firms that do not shut down, emissions remain constant before and after acquisition. Emissions intensity does not change significantly, neither for exiting firms nor for those remaining in operation. At the parent company level, aggregate emissions increase significantly. However, average emissions per facility do not change. Emissions intensity remains constant, both at the aggregate level and per facility. These findings indicate that newly acquired facilities are of similar size and they produce at similar emissions intensity as facilities already in possession of the parent company. Moreover, there is no evidence for spillovers of the acquisition on the latter facilities as their emissions and emissions intensity remains constant as well.

Table 6.3: Aggregate effects on firms

	Full sample			Remaining firms			Exiting firms		
Dependent variables	**ATT**	**SE**	**N**	**ATT**	**SE**	**N**	**ATT**	**SE**	**N**
Total emissions	-0.549***	(0.085)	6,272	-0.053	(0.085)	3,097	-0.748***	(0.166)	3,175
Output	-1.484***	(0.300)	5,547	-0.952**	(0.365)	2,791	-1.494***	(0.523)	2,756
Emissions intensity	-0.030	(0.022)	5,175	-0.005	(0.017)	2,457	-0.058	(0.052)	2,718
Total factor productivity	0.004	(0.005)	2,945	0.001	(0.006)	1,598	0.005	(0.009)	1,347
Operating profits	-0.751*	(0.386)	4,191	-0.533*	(0.341)	2,168	-1.172	(0.732)	2,023
Labor input	-0.474***	(0.099)	5,264	-0.297**	(0.111)	2,609	-0.516**	(0.191)	2,655
Capital input	-1.217***	(0.292)	5,500	-0.701*	(0.352)	2,761	-1.446**	(0.515)	2,739
Labor expenditures	-1.152***	(0.301)	5,057	-0.663*	(0.288)	2,480	-1.176*	(0.529)	2,577
Intangible fixed assets	-0.148	(0.392)	5,004	-0.029	(0.482)	2,671	0.040	(0.709)	2,333

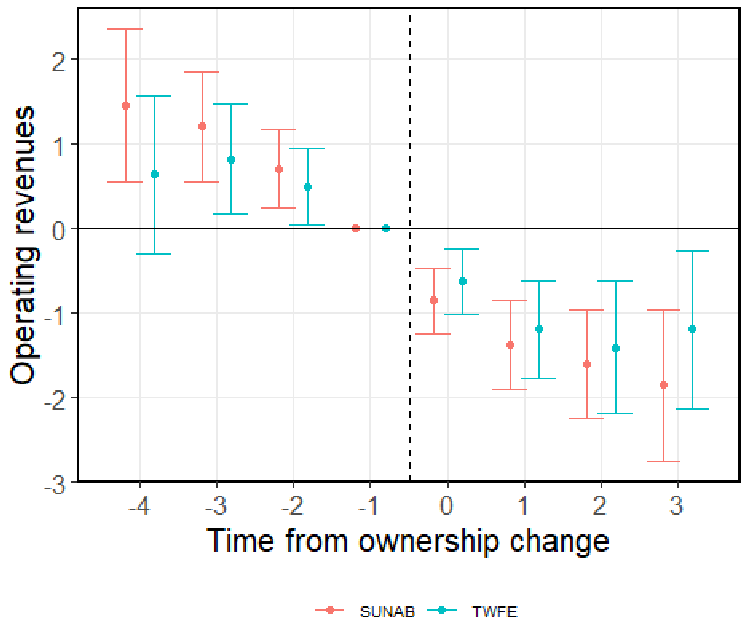
Notes: The first column denotes the respective dependent variables each with an inverse hyperbolic sine transformation. Each line represents three separate event study regressions on the full sample of firms, on remaining firms as well as on exiting firms. For each regression, we report the point estimate of the aggregated effect of the event study following Sun and Abraham (2021) (ATT), the standard error (SE) and the number of observations (N). Standard errors clustered at the firm level are in parenthesis. Output refers to deflated operating revenues, emissions intensity to total emissions divided by output, labor input to number of employees, capital input to deflated tangible fixed assets, respectively. As a robustness check, Table A6.5 in the Appendix reports the same estimations on the 50 percent largest firms. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

6.4.2 Output and production inputs

We also take a look at output and production inputs after ownership changes which are potential drivers of the emission reductions that we observe. At the firm level, output as proxied by operating revenues shows a relatively steep falling pre-trend before the change in ownership that continues after (see Figure 6.10). The average decrease after treatment is at 77 percent. Similar to the results for emissions, the falling pre-trend appears to be predominantly driven by the firms that are closed down in the years after the ownership change (see Figure A6.20 in the Appendix). For those firms, pre-trends fall steeply and the estimated reduction after the ownership change is at 78 percent. For firms that remain in operation, the coefficient in the year before treatment is marginally significant, but no clear pre-trend is visible. The reduction in output after the change in ownership is smaller, but still sizable at 61 percent. We run sensitivity checks for all three samples that adjust the treatment effects for violation in parallel trends before treatment. We find for each of the effects in all firms, exiting firms and remaining firms, respectively, that coefficients are only significant when they are adjusted by less than half of the maximum violation of parallel trends in the pre-treatment period.[35] This indicates the observed reductions in output were rather not causally driven by the ownership changes but would – at least partially – have happened without acquisition as well.

[35]See Figures A6.67 to A6.69 in the Appendix.

Figure 6.10: Effect on output at the firm level



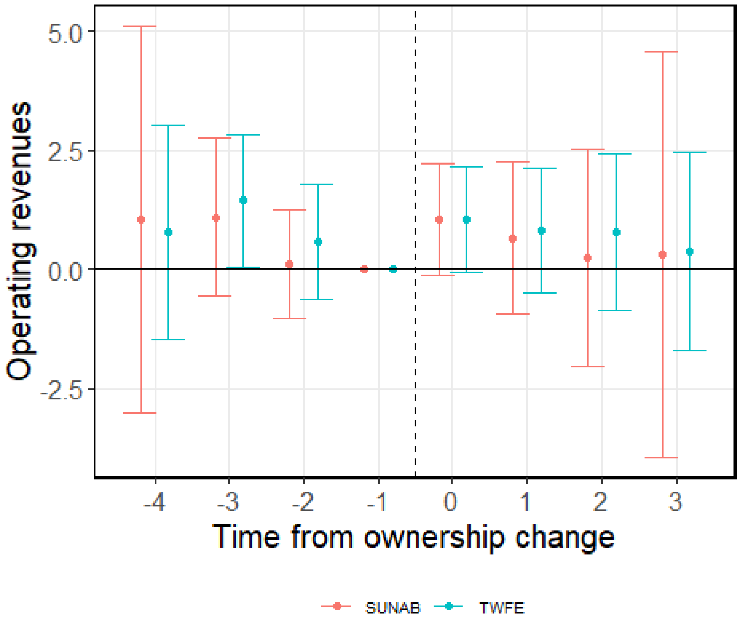
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

In the Orbis data, we also observe labour and capital production inputs as well as labour expenses. The results for these outcomes mirror the ones for the output: a steep falling trend before the change in ownership that continues after (see Figures A6.21 to A6.23 in the Appendix). A sensitivity test that adjusts post-treatment coefficients for violations in parallel trends before changing ownership finds that effects turn insignificant at adjustments of less than half of the maximum violation observed before treatment for either of the three outcomes.[36] When separating the sample into firms that remain in operation versus those that close down, also for these outcomes the reductions are stronger for exiting firms. However, reductions are still sizable for firms remaining in operation (see Table 6.3).

At the level of acquiring parent companies, the effect of an ownership change on output is not significant. However, the estimation is rather imprecise as standard errors are very large and the same goes for the estimation on the subsample of parent companies that do not shut down their acquisition (see Figure 6.11 and Table 6.4). Labour and capital input as well as labour expenses increase significantly after acquisition, but the increases are smaller in the sample focusing on acquisitions that are not closed down - and the increase in labor expenses is not significant for the latter (see Figures A6.24 to A6.26 and Table 6.4).

[36]See Figures A6.70 to A6.72 in the Appendix.

Figure 6.11: Effect on output at the parent company level



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Per industrial facility of the acquiring parent company, output does not change significantly, but again standard errors are quite large for this estimation. Capital input and labor expenditures significantly increase and the increases are smaller in the sample focusing on acquisitions that are not closed down. This indicates that newly acquired facilities employ above average amounts of capital input and labor expenditures, and this effect is stronger for facilities that are closed down subsequent to the ownership change. Labor input relative to the number of facilities does not significantly change after the acquisition in neither of the samples (see Figures A6.28 to A6.30 in the Appendix and Table 6.4). For other facilities in the parent company, the acquisition of the new facility neither affects output nor production inputs (see Figures A6.31 to A6.34 in the Appendix and Table 6.5).

Table 6.4: Aggregate effects on acquiring parent companies

	Full sample			Remaining firms		
Dependent variables	**ATT**	**SE**	**N**	**ATT**	**SE**	**N**
Total emissions	0.742***	(0.217)	2,274	0.673*	(0.261)	1,444
Output	0.592	(1.186)	2,248	0.048	(1.440)	1,433
Emissions intensity	0.020	(0.028)	1,858	0.037	(0.033)	1,163
Total factor productivity	-0.007	(0.017)	1,305	-0.002	(0.020)	816
Operating profits	0.888	(2.124)	2,248	1.693	(1.785)	1,433
Labor input	0.834***	(0.286)	2,248	0.567	(0.346)	1,433
Capital input	2.701***	(0.774)	2,248	2.094*	(0.972)	1,433
Labor expenditures	2.270***	(0.670)	2,248	1.761*	(0.847)	1,433
Intangible fixed assets	1.477**	(0.601)	2,248	0.880	(0.709)	1,433
Total emissions per industrial facility	0.273	(0.204)	2,274	0.302	(0.240)	1,444
Output per industrial facility	0.256	(1.119)	2,248	-0.200	(1.363)	1,433
Emissions intensity per industrial facility	0.016	(0.021)	1,858	0.027	(0.026)	1,163
Total factor productivity per industrial facility	-0.435***	(0.076)	1,305	-0.277***	(0.067)	816
Operating profits per industrial facility	0.547	(2.036)	2,248	1.384	(1.713)	1,433
Labor input per industrial facility	0.426	(0.260)	2,248	0.253	(0.313)	1,433
Capital input per industrial facility	2.252**	(0.750)	2,248	1.741*	(0.940)	1,433
Labor expenditures per industrial facility	1.864**	(0.643)	2,248	1.430*	(0.814)	1,433
Intangible fixed assets per industrial facility	1.112*	(0.578)	2,248	0.615	(0.686)	1,433

Notes: The first column denotes the respective dependent variables each with an inverse hyperbolic sine transformation. Each line represents two separate event study regressions on the full sample of acquiring parent companies as well as on acquiring parent companies that do not close down their target. For each regression, we report the point estimate of the aggregated effect of the event study following Sun and Abraham (2021) (ATT), the standard error (SE) and the number of observations (N). Standard errors clustered at the firm level are in parenthesis. Output refers to deflated operating revenues, emissions intensity to total emissions divided by output, labor input to number of employees, capital input to deflated tangible fixed assets, respectively. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

In summary, the trends in output and production inputs mirror emissions at the firm level: these outcomes already decrease before acquisition and continue falling afterwards. These trends are stronger for firms that are shut down after they change owners, confirming that emissions reductions are indeed driven by output reductions. For firms that remain in operation, the falling trends before and after acquisition are less pronounced. At the level of the acquiring parent company, production inputs increase. The increase is smaller for firms purchased by parent companies that keep their targets in operation, again mirroring the pattern of emissions. On average at the facility level, capital input and labor expenses increase, but the increase is less if acquired facilities remain in operation. Other facilities' inputs and output within the same parent company is not affected which suggests that production is not reallocated across facilities within a parent company after acquisition.

6.4.3 Total factor productivity, operating profits and intangible fixed assets

The finding that emissions of facilities and firms that remain in operation after acquisition do not change suggests that no technology is transferred to the acquisition that would affect its environmental performance. We also do not find emissions of other facilities in the acquiring company to change which indicates that no environmental technology is transferred to them from the acquisition either. In this section, we check whether the ownership change induces any other types of spillovers. To this end, we look at total factor productivity, operating profits and intangible fixed assets.

At the firm level, productivity remains virtually constant after ownership changes and the result remains the same when separating firms that remain in operation from firms that exit (see Figures A6.35 and A6.36 in the Appendix and Table 6.3).[37] Operating profits drop significantly by 52 percent - however, the falling trend starts already in the pre-treatment period so that the reduction can be at most partially attributed to the ownership change (see Figure A6.37 in the Appendix). A sensitivity check shows that the effect in the third year after ownership change turns even less significant when adjusting for violations in parallel trends before treatment.[38] The decrease appears to be driven by firms that remain in operation: for those, the reduction is significant and pre-trends are insignificant. For firms that shut down, the falling trend before treatment is even more pronounced and seems to continue after the change in ownership, but the average effect is not significant (see Figure A6.38 in the Appendix and Table 6.3). As for total factor productivity, the results on intangible fixed assets including, e.g., patents, copyrights, trademarks and goodwill, are all insignificant: both for the full sample and the split samples, the effect is close to zero (see Figures A6.39 and A6.40 in the Appendix and Table 6.3).

At the level of acquiring parent companies, the effect of an ownership change aggregated over all facilities on total factor productivity and operating profits is insignificant, and the same goes for the subsample of only parent companies that do not shut down their acquisition (see Figures A6.41 and A6.42 and Table 6.4). However, intangible fixed assets increase strongly and permanently after the acquisition (see Figure A6.43 in the Appendix). This results seems to be driven by parent companies that shut down their targets after acquisition as the effect is not significant for the sample where targets remain in operation (see Table 6.4).

Per industrial facility in the acquiring parent company, total factor productivity drops

[37]We use our estimated TFP values as an outcome variable in the event study regressions. We run the regression with the sample of firms for which capital input is available.

[38]See Figure A6.73 in the Appendix.

by 35 percent which means that the newly acquired facilities are on average less productive than the rest of facilities. The drop in the sample of parent companies that do not shut down their acquisitions is at 24 percent a bit smaller – acquired firms that are shut down tend to be less productive than the ones that remain in operation (see Figure A6.44 and Table 6.4). Operating profits per facility are not affected by the acquisition of the parent company, not either for the sample of parent companies that let their acquisitions continue operation (see Figure A6.45 and Table 6.4). Intangible fixed assets significantly and strongly increase per facility, meaning that acquired facilities hold above average intangibles, but this is only the case for the full sample. In the sample of only parent companies that do not shut down their acquisition, the effect is insignificant (see Figure A6.46 and Table 6.4). That indicates that parent companies tend to shut down facilities that hold above average intangible fixed assets.

For other facilities within the acquiring parent company, there appear to materialize spillovers from the acquisition: both operating profits and intangibles increase strongly and significantly (see Figures A6.48 and A6.49 in the Appendix and Table 6.5). The effect on total factor productivity shows a minor increase but the years before treatment indicate an increasing trend as well so that the increase is unlikely driven by the ownership change (see Figure A6.47 in the Appendix and Table 6.5).

Table 6.5: Aggregate effects on other industrial firms of acquiring parent companies

Dependent variables	**ATT estimate**	**Std. Error**	**N**
Total emissions	0.032	(0.153)	1,864
Output	-0.088	(0.715)	1,691
Emissions intensity	0.006	(0.019)	1,474
Total factor productivity	0.014*	(0.007)	972
Operating profits	3.625**	(1.479)	1,665
Labor input	-0.275	(0.184)	1,578
Capital input	-0.109	(0.639)	1,658
Labor expenditures	-0.101	(0.575)	1,534
Intangible fixed assets	0.995*	(0.522)	1,551

Notes: The first column denotes the respective dependent variables each with an inverse hyperbolic sine transformation. Each line represents a separate event study regression. Output refers to deflated operating revenues, emissions intensity to total emissions divided by output multiplied to 100,000, labor input to number of employees, capital input to tangible fixed assets, respectively. The second and third columns show the point estimates and standard errors of the aggregated effect of the event study following Sun and Abraham (2021). The fourth column contains the number of observations. Standard errors clustered at the firm level are in parenthesis. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

In summary, the results suggest that parent companies close down targets with below-average productivity and above-average intangible fixed assets. These intangibles subsequently appear to be transferred to other facilities owned by the acquiring parent company, potentially driving the increases in productivity and operating profits observed in these facilities.

6.4.4 Discussion

We observe two different patterns in facilities and firms after an ownership change. On the one hand, a significant share of them reduces their emissions strongly alongside a strong decrease in production and production inputs via shutdowns. Emissions intensity, total factor productivity, operating profits and intangible fixed assets in these firms do not change significantly before the shutdown. These facilities and firms appear to have been on a downward trajectory already before acquisition so that we cannot attribute the reduction in emissions to the change in ownership. On the other hand, emissions and emissions intensity of facilities and firms that remain in operation after acquisition do not change significantly and reductions in output and inputs are smaller. Productivity and intangibles are not affected either, but profits decrease significantly.

At the parent company level, the dichotomy between parents that close down at least one of their targets versus parents that continue operation in all of their tragets after acquisition is visible as well. In the former group without shutdowns, the increase in production after acquisition both in the aggregate and per facility is smaller then for the full sample. Moreover, the drop in production is smaller per industrial facility and intangibles do not change significantly while they increase in the full sample. For other facilities in the acquiring parent company we are not able to investigate the dichotomy due to sample restrictions. Here, we observe that the acquisition significantly increases productivity, profits and intangibles in other facilities in the parent company. Emissions intensity is not affected at any level in any of the samples.

We conclude from these results that neither total emissions nor emissions intensity are affected by ownership changes. Even emissions reductions via shutdowns of facilities and firms after acquisition do not seem to be caused by ownership changes as emissions already start falling several years before an acquisition. The transfer of technologies between the acquiring and the acquired firm is often discussed as rationale for mergers and acquisitions in the literature. Our results do not provide evidence for a transfer of environmental technologies between acquiring and acquired firm in either direction as emissions intensity remains constant at all entity levels. Likewise, we do not find strong evidence for non-environmental technology transfers. Productivity in the other facilities of acquiring parent companies increases, but average productivity of targets is lower than in these other facilities. Our results for European industrial firms and facilities differ from the findings by Jacqz (2021) that US-American facilities that continue operation after an M&A event reduce emissions of toxic chemicals, hinting at operational changes and technology transfers as reason for the observed reductions.

The significant share of close-downs in the years after acquisition seems to follow a rationale other than technology transfers. One explanation in the literature is that acquiring mother firms want to reduce output in sectors where there are oligopoly rents to harvest. Previous empirical work in various settings finds a tendency of acquiring firms to shut down a significant proportion of their targets after acquisition. Several studies report that the probability of shutdown after acquisition is higher if target and acquiring firm are not active in the same sector. Blonigen and Pierce (2016) find that the exit probability of acquired US plants is higher than for plants in the control group and that the shutdowns are predominantly in constellation where the acquiring firm does not operate in the same sector as the target. Kaplan and Weisbach (1992) study shutdowns of targets over a longer time horizon of up to 18 years after acquisition and find that 44 percent of acquired firms close down over this period. Diversifying acquisitions that are active in another industry than the acquirer are close to four times more likely to shut down. Maksimovic et al. (2011) find that in acquisitions of US targets, 19 percent are closed down after the third year. The likelihood of shutdown is lower for targets in the same industry as the acquirer, larger targets and larger acquirers. These findings point to objectives other than market power prevalent in these settings studied. Cunningham et al. (2021) document objectives of market power to matter in shutdown decisions finding that 5 to 7 percent of acquisitions in the pharmaceutical sector are killer acquisitions that are supposed to discontinue the targets' innovations and kill future competition. Davis et al. (2014) find productivity gains to be the rationale of shutdowns after private-equity buyouts, where less productive targets are closed down after acquisition.

To explore whether market power is a rationale for shutdowns in our sample, we conduct two empirical tests. First, we check whether we observe exits after acquisition predominantly in specific sectors. In Section 6.4.2, we find that the output reduction is relatively uniform over sectors (see Figure A6.19 in the Appendix). Moreover, we compare whether the share of firms that change ownership in each NACE2 sector is in the same range as the share of firms that shuts down after acquisition in the same sector. Doing this, we hope to uncover whether there are specific sectors that most of the shutdowns observed in our sample can be attributed to. We do not see strong differences in shares of firms and shutdowns for any sector, but nevertheless we conduct a two-sided proportion test to test for statistically significant differences in the shares. We find a few sectors in which the share out of all exiting firms is significantly higher than the share out of all firms with ownership change. But neither of the differences in sector shares is economically significant.[39]

As a second empirical check, we investigate whether a higher share of acquisitions that happen within the same sector is shut down subsequently as compared to acquisitions where the target and the acquiring company are predominantly active in different sectors. If the rationale for shutdowns is market power, we would mainly observe them by parent companies that are active in the same sector. We define an acquisition within the same sector as the target being predominantly active in the same sector as the majority of firms owned by the acquiring parent company according to their NACE2 classification. We find that the largest share, 91.4 percent of acquisitions, are within-sector. We conduct a t-test (p value = 0.5744) and do not find a significant difference in the share of within-sector acquisitions between the firms being shut down and the firms remaining in operation. The shares of within-sector acquisitions are 90.4 percent for exiting firms and 91.6 percent for firms remaining in operation. Our empirical checks do not provide evidence for that a significant share of firms in our sample is shut down strategically after acquisition to gain market power.

An alternative explanation for the rationale of closing down targets after acquisition is the transfer of intangible fixed assets from the acquired to the acquiring firm. Intangible fixed assets are assets of non-physical nature, such as intellectual property, licenses, trademarks or patents. While the stock of intangibles does not significantly change in acquired firms that shut down subsequently, the increase in the aggregate stock of intangibles in the parent company seems to be driven by this subset of firms. Meanwhile, intangibles do not change in parent companies that do not shut down any of their targets. The stock of intangibles increases in other facilities of acquiring parent companies, probably via transfers of intangibles from the acquired firm that is about to shut down to these other facilities. Literature suggests that the transfer of intangibles is often an objective of mergers and acquisitions as it is expected to create value for the acquiring company, either directly or via exploiting differences in tax rates which would not necessarily affect production (Juranek et al., 2018; Mamun et al., 2021; Filipovic and Wagner, 2023; Juranek et al., 2023). We look at profits before taxes so that a potential effect on taxes paid is not observable. We see positive spillovers on operating profits and productivity which may be driven by a transfer of intangibles. Evidence on the effect of acquisitions on intangibles in other settings is not conclusive.[40]

[39]The sectors for which the difference in shares is statistically significant are manufacture of electrical equipment, manufacture of printed goods, beverage production, food and animal feed, and collection, treatment and disposal of waste, recycling. Detailed results on the sector shares and the proportion tests are available upon request from the authors.

[40]Lerner et al. (2011) study the effect of leveraged buyouts on innovation activities measured by patenting activity and find a positive impact. Amess et al. (2016) look at the effect of private equity-backed leveraged buyouts on the patent stock and find it to increase as a result of the acquisition. Conversely, Cumming et al. (2020) find a negative effect of public-to-private buyouts on patents and patent citations. Haucap et al. (2019) find a negative effect of horizontal mergers in the pharmaceutical

The acquired facilities and firms do not all exit in the year of their acquisition. As can be deduced for the step-wise reduction in output and emissions over the four years since the ownership change, a substantial share of acquired firms and their facilities only exit after two or three years. Potential reasons for the grace period that the parent companies grant their newly acquired firms could be rigid labour markets in the form of strong labour protection laws in some of the sample countries that do not allow for a quicker shutdown of large entities as they are present in our sample. Previous literature finds that the likelihood of shutdown after acquisition is smaller for larger entities.[41] Along the same lines, in our sample the propensity to be shut down after acquisition is smaller for the largest facilities and firms. However, reductions in total emissions, output and production inputs as well as profits already start in the years before the ownership change. The mechanisms driving these patterns are unclear.

6.5 Conclusion

We estimate the impact of ownership changes on pollutant emissions and economic performance indicators of industrial firms in Europe. We find a robust decrease in total emissions of newly acquired facilities and firms, which is exclusively driven by facilities and firms closing down in the years after acquisition rather than changes in abatement technology. Acquired firms that remain in operation do not change their total emissions and emissions intensity and neither do other firms in the acquiring parent company. From an environmental perspective, these acquisitions are a zero-sum game that neither harms nor benefits the environment. The shutdowns would benefit the environment if the firms had continued operation in the absence of the ownership change. However, the observed reductions in emissions and output that start prior to the change in ownership indicate that the acquisition may not have been the cause for the shutdowns.

Even though we use a comprehensive data set and cover a major industrial continent with different countries, more research is needed to investigate these effects in other settings. Similarly, future research should try to disentangle even more deeply the mechanisms of the effects on emissions and economic performance indicators and assess their consequences on the firm distribution.

Finally, our research highlights that – absent comprehensive pollution regulation – environmental components could deserve more attention when discussing the costs

market in Europe on patenting of the merged entity.

[41] McGuckin and Nguyen (1995) observe for acquisitions in the US food manufacturing sector that larger facilities are more likely to be purchased than closed when they are performing poorly. Maksimovic et al. (2011) find that the likelihood of shutdown in acquisitions of US targets is lower for larger targets and larger acquirers.

and benefits of ownership changes as well as potentially play a more prominent role in M&A regulation.

References

Abdulai, A. (2023). Information acquisition and the adoption of improved crop varieties. *American Journal of Agricultural Economics 105*(4), 1049–1062.

Achtziger, A., P. M. Gollwitzer, and P. Sheeran (2008). Implementation intentions and shielding goal striving from unwanted thoughts and feelings. *Personality and Social Psychology Bulletin 34(3)*, 381–393.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6).

Aden, J., A. Kyu-Hong, and M. T. Rock (1999). What is driving the pollution abatement expenditure behavior of manufacturing plants in Korea? *World Development 27*(7), 1203–1214.

Aghion, P., J. Van Reenen, and L. Zingales (2013). Innovation and institutional ownership. *American Economic Review 103*(1), 277–304.

Aitken, B. J. and A. E. Harrison (1999). Do domestic firms benefit from direct foreign investment? Evidence from Venezuela. *American Economic Review 89*(3), 605–618.

Akey, P. and I. Appel (2021). The limits of limited liability: Evidence from industrial pollution. *Journal of Finance 76*(1), 5–55.

Alberini, A., W. Gans, and D. Velez-Lopez (2011). Residential consumption of gas and electricity in the US: The role of prices and income. *Energy Economics 33*(5), 870–881.

Albornoz, F., M. A. Cole, R. J. Elliott, and M. G. Ercolani (2009). In search of environmental spillovers. *World Economy 32*(1), 136–163.

Alcacer, J. and M. Zhao (2012). Local R&D strategies and multilocation firms: The role of internal linkages. *Management Sciences 58*(4).

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

Allcott, H. (2013). The welfare effects of misperceived product costs: Data and calibrations from the automobile market. *American Economic Journal: Economic Policy 5*(3), 30–66.

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

Allcott, H. and M. Greenstone (2012). Is there an energy efficiency gap? *Journal of Economic Perspectives 26*(1), 3–28.

Allcott, H. and M. Greenstone (2017). Measuring the welfare effects of residential energy efficiency programs. *NBER Working Paper 23386*.

Allcott, H. and J. B. Kessler (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics 11*(1), 236–276.

Allcott, H., S. Mullainathan, and D. Taubinsky (2014). Energy policy with externalities and internalities. *Journal of Public Economics 112*, 72–88.

Allcott, H. and D. Taubinsky (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review 105*(8), 2501–2538.

Allcott, H. and N. Wozny (2014). Gasoline prices, fuel economy, and the energy paradox. *Review of Economics and Statistics 96*(5), 779–795.

Altmann, S., C. Traxler, and P. Weinschenk (2021). Deadlines and memory limitations. *Management Science*.

Ambrose, A., W. Baker, E. Batty, and A. Hawkins (2019). Reaching the 'hardest to reach' with energy advice.

Ameli, N. and N. Brandt (2015). 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)*.

Amess, K., J. Stiebale, and M. Wright (2016). The impact of private equity on firms' patenting activity. *European Economic Review 86*, 147–160.

Anderson, S. T., R. Kellogg, J. M. Sallee, and R. T. Curtin (2011). Forecasting gasoline prices using consumer surveys. *American Economic Review 101*(3), 110–114.

Andersson, J. J. (2019). Carbon taxes and CO_2 emissions: Sweden as a case study. *American Economic Journal: Economic Policy 11*(4), 1–30.

Andor, M. A., D. H. Bernstein, and S. Sommer (2021). Determining the efficiency of residential electricity consumption. *Empirical Economics 60*, 2897–2923.

Andor, M. A., A. Gerster, J. Peters, and C. M. Schmidt (2020). Social norms and energy conservation beyond the US. *Journal of Environmental Economics and Management 103*, 102351.

Andor, M. A., A. Gerster, and S. Sommer (2020). Consumer inattention, heuristic thinking and the role of energy labels. *The Energy Journal 41*(1), 83–112.

Ankney, K. et al. (2021). Do credit constraints explain the energy efficiency gap? Evidence from the US new vehicle market.

Arnold, D. (2019). Mergers and acquisitions, local labor market concentration, and worker outcomes.

Arnold, J. M. and B. S. Javorcik (2009). Gifted kids or pushy parents? Foreign direct investment and plant productivity in Indonesia. *Journal of International Economics 79*(1), 42–53.

Bajgar, M., G. Berlingieri, S. Calligaris, C. Criscuolo, and J. Timmis (2020). Coverage and representativeness of Orbis data. OECD Science, Technology and Industry Working Papers 2020/06.

Banerjee, A., R. Hanna, B. A. Olken, E. Satriawan, and S. Sumarto (2022). Electronic food vouchers: Evidence from an at-scale experiment in Indonesia. *American Economic Review 113(2)*, 514–47.

Banerjee, A. V., S. Cole, E. Duflo, and L. Linden (2007). Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics 122*(3), 1235–1264.

Barreca, A. I., M. Guldi, J. M. Lindo, and G. R. Waddell (2011). Saving babies? Revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics 126*(4), 2117–2123.

BDEW (2024). German Association of Water and Energy Industries: Strompreisanalyse Juli 2024. Technical report. Accessed 6th August, 2024.

Berkouwer, S. B. and J. T. Dean (2022). Credit, attention, and externalities in the adoption of energy efficient technologies by low-income households. *American Economic Review 112*(10), 3291–3330.

Bertrand, M., D. Karlan, S. Mullainathan, E. Shafir, and J. Zinman (2010). What's advertising content worth? Evidence from a consumer credit marketing field experiment. *The Quarterly Journal of Economics 125*(1), 263–306.

Bertrand, M., E. F. Luttmer, and S. Mullainathan (2000). Network effects and welfare cultures. *The Quarterly Journal of Economics 115*(3), 1019–1055.

Bertrand, M., S. Mullainathan, and E. Shafir (2004). A behavioral-economics view of poverty. *American Economic Review 94*(2), 419–423.

Bertrand, M., S. Mullainathan, and E. Shafir (2006). Behavioral economics and marketing in aid of decision making among the poor. *Journal of Public Policy & Marketing 25*(1), 8–23.

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

Blasch, J., N. Boogen, C. Daminato, M. Filippini, et al. (2021). Empower the consumer! Energy-related financial literacy and its implications for economic decision making. *Economics of Energy & Environmental Policy 10*(2), 149–181.

Blonigen, B. A. and J. R. Pierce (2016). Evidence for the effects of mergers on market power and efficiency. *National Bureau of Economic Research*.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013). Does management matter? Evidence from India. *Quarterly Journal of Economics 128*(1).

Bollinger, B., J. Burkhardt, and K. Gillingham (2018). Peer effects in water conservation: Evidence from consumer migration. *National Bureau of Economic Research*.

Bollinger, B. and K. Gillingham (2012). Peer effects in the diffusion of solar photovoltaic panels. *Marketing Science 31*(6), 900–912.

Bonan, J., C. Cattaneo, G. d'Adda, and M. Tavoni (2021). Can social information programs be more effective? The role of environmental identity for energy conservation. *Journal of Environmental Economics and Management 108*, 102467.

Boomhower, J. and L. W. Davis (2014). A credible approach for measuring inframarginal participation in energy efficiency programs. *Journal of Public Economics 113*, 67–79.

Bradford, D., C. Courtemanche, G. Heutel, P. McAlvanah, and C. Ruhm (2017). Time preferences and consumer behavior. *Journal of Risk and Uncertainty 55*, 119–145.

Brandon, A., P. J. Ferraro, J. A. List, R. D. Metcalfe, M. K. Price, and F. Rundhammer (2017). Do the effects of nudges persist? Theory and evidence from 38 natural field experiments. Technical report, National Bureau of Economic Research.

Brent, D. A. and M. B. Ward (2018). Energy efficiency and financial literacy. *Journal of Environmental Economics and Management 90*, 181–216.

Bronchetti, E. T., T. S. Dee, D. B. Huffman, and E. Magenheim (2013). When a nudge isn't enough: Defaults and saving among low-income tax filers. *National Tax Journal 66*(3), 609–634.

Brot-Goldberg, Z., T. Layton, B. Vabson, and A. Y. Wang (2023). The behavioral foundations of default effects: Theory and evidence from Medicare Part D. *American Economic Review 113*(10), 2718–2758.

Bruckmeier, K., R. T. Riphahn, and J. Wiemers (2021). Misreporting of program take-up in survey data and its consequences for measuring non-take-up: New evidence from linked administrative and survey data. *Empirical Economics 61*, 1567–1616.

Bucher-Koenen, T., A. Hackethal, J. Koenen, and C. Laudenbach (2021). Gender differences in financial advice.

Bundesagentur für Arbeit. Bedarfsgemeinschaften und deren Mitglieder 2020. `https://statistik.arbeitsagentur.de/SiteGlobals/Forms/Suche/Einzelheftsuche_Formular.html?nn=1460284&topic_f=gs-asu-sgbii-rev`. Accessed: 2022-05-19.

Bundesverband der Energie- und Wasserwirtschaft e.V. (BDEW). Energiemarkt Deutschland 2019. `https://www.bdew.de/service/publikationen/bdew-energiemarkt-deutschland-2019/`. Accessed: 2022-05-19.

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

Burdina, M., R. S. Hiller, and N. E. Metz (2017). Goal attainability and performance: Evidence from Boston marathon qualifying standards. *Journal of Economic Psychology 58*, 77–88.

Burstein, A. T. and A. Monge-Naranjo (2009). Foreign know-how, firm control, and the income of developing countries. *Quarterly Journal of Economics 124*(1).

Busse, M. R., C. R. Knittel, and F. Zettelmeyer (2013). Are consumers myopic? Evidence from new and used car purchases. *American Economic Review 103*(1), 220–256.

Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal 17*(2), 372–404.

Calvet, L. E., J. Y. Campbell, and P. Sodini (2009). Measuring the financial sophistication of households. *American Economic Review 99*(2), 393–98.

Chapman, J., E. Snowberg, S. Wang, and C. Camerer (2024). Looming large or seeming small? Attitudes towards losses in a representative sample. *Review of Economic Studies*, rdae093.

Chen, C. (2023). Framing energy-efficiency programs: A survey experiment. *Energy Policy 183*, 113776.

Chen, W. (2011). The effect of investor origin on firm performance: Domestic and foreign direct investment in the United States. *Journal of International Economics 83*(2), 219–228.

Chester, L. (2016). Studying low-income households: Challenges and issues. *Handbook of research methods and applications in heterodox economics*, 341–363.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. *American Economic Review 99*(4), 1145–1177.

Chlond, B., T. Goeschl, and M. Kesternich (2025). Money versus procedures—Evidence from an energy efficiency assistance program. *Journal of Environmental Economics and Management 130*, 103080.

Chlond, B., T. Goeschl, M. Kesternich, and M. Werthschulte (2024). Transporting behavioral insights to low-income households: A field experiment on energy efficiency investments.

Christensen, P., P. Francisco, E. Myers, and M. Souza (2023). Decomposing the wedge between projected and realized returns in energy efficiency programs. *Review of Economics and Statistics 105*(4), 798–817.

Clo, S., M. Florio, and F. Rentocchini (2020). Firm ownership, quality of government and innovation: Evidence from patenting in the telecommunication industry. *Research Policy 49*(5).

Cohodes, S. R. and K. S. Parham (2021). Charter schools' effectiveness, mechanisms, and competitive influence. *NBER Working Paper No. 28477*.

Commander, S. and J. Svejnar (2011). Business environment, exports, ownership, and firm performance. *Review of Economics and Statistics 93*(1), 309–337.

Cumming, D., R. Peter, and M. Tarsalewska (2020). Public-to-private buyouts and innovation. *British Journal of Management 31*(4), 811–829.

Cunningham, C., F. Ederer, and S. Ma (2021). Killer acquisitions. *Journal of Political Economy 129*(3), 649–702.

Dahabreh, I. J. and M. A. Hernán (2019). Extending inferences from a randomized trial to a target population. *European Journal of Epidemiology 34*, 719–722.

David, J. M. (2021). The aggregate implications of mergers and acquisitions. *Review of Economic Studies 88*, 1796–1830.

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

Davis, S. J., J. Haltiwanger, K. Handley, R. Jarmin, J. Lerner, and J. Miranda (2014). Private equity, jobs, and productivity. *American Economic Review 104*(12), 3956–3990.

DeGolia, A. H., E. H. Hiroyasu, and S. E. Anderson (2019). Economic losses or environmental gains? Framing effects on public support for environmental management. *PLoS One 14*(7), e0220320.

Degtiar, I. and S. Rose (2023). A review of generalizability and transportability. *Annual Review of Statistics and Its Application 10*, 501–524.

Della Valle, N. and P. Bertoldi (2021). Mobilizing citizens to invest in energy efficiency. https://op.europa.eu/en/publication-detail/-/publication/b27bcfaa-5320-11ec-91ac-01aa75ed71a1/language-en.

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

Döbbeling-Hildebrandt, N., K. Miersch, T. M. Khanna, M. Bachelet, S. B. Bruns, M. Callaghan, O. Edenhofer, C. Flachsland, P. M. Forster, M. Kalkuhl, et al. (2024). Systematic review and meta-analysis of ex-post evaluations on the effectiveness of carbon pricing. *Nature communications 15*(1), 4147.

Domurat, R., I. Menashe, and W. Yin (2021). The role of behavioral frictions in health insurance marketplace enrollment and risk: Evidence from a field experiment. *American Economic Review 111*(5), 1549–1574.

Dörre, K. (2015). Tests for the underclass: The social effects of activating labour market policy in Germany. In *The New Social Division: Making and Unmaking Precariousness*, pp. 83–100. Springer.

Dubois, U. and A. Sinea (2023). Methodological challenges in energy poverty research. *International Journal of Market Research 65*(2-3), 340–358.

Duflo, E., R. Glennerster, and M. Kremer (2007). Using randomization in development economics research: A toolkit. *Handbook of Development Economics 4*, 3895–3962.

Duma, N., J. Vera Aguilera, J. Paludo, C. L. Haddox, M. Gonzalez Velez, Y. Wang, K. Leventakos, J. M. Hubbard, A. S. Mansfield, R. S. Go, et al. (2018). Representation of minorities and women in oncology clinical trials: Review of the past 14 years. *Journal of Oncology Practice 14*(1), e1–e10.

d'Adda, G., Y. Gao, and M. Tavoni (2022). A randomized trial of energy cost information provision alongside energy-efficiency classes for refrigerator purchases. *Nature Energy 7*(4), 360–368.

EEA (2021). Air quality in Europe 2021. `https://www.eea.europa.eu/publications/air-quality-in-europe-2021`. Accessed 25th January, 2022.

EEA (2022). EEA greenhouse gases - data viewer. `https://www.eea.europa.eu/data-and-maps/data/data-viewers/greenhouse-gases-viewer`. Accessed 25th January, 2022.

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

Engström, P., K. Nordblom, H. Ohlsson, and A. Persson (2015). Tax compliance and loss aversion. *American Economic Journal: Economic Policy 7*(4), 132–164.

Ericson, K. M. (2017). On the interaction of memory and procrastination: Implications for reminders, deadlines, and empirical estimation. *Journal of the European Economic Association 15*(3), 692–719.

Eryzhenskiy, I., L.-G. Giraudet, and M. Segu (2023). Success and failure of a zero-interest green loan program: Evidence from france. *Available at SSRN 4534399*.

Escobar, L. and A. Pedraza (2023). Active trading and (poor) performance: The social transmission channel. *Journal of Financial Economics 150*(1), 139–165.

Essien, U. R., S. B. Dusetzina, and W. F. Gellad (2021). A policy prescription for reducing health disparities—achieving pharmacoequity. *Jama 326*(18), 1793–1794.

European Commission (2003). Commission Directive 2003/66/EC of 3 July 2003 amending Directive 94/2/EC implementing Council Directive 92/75/EC with regard to energy labelling of household electric refrigerators, freezers and their combinations. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32003L0066. Accessed 7th March, 2024.

European Commission (2010). Commission Delegated Regulation (EU) No 1060/2010 of 28 September 2010 supplementing Directive 2010/30/EU of the European Parliament and of the Council with regard to energy labelling of household refrigerating appliances. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32010R1060. Accessed 7th March, 2024.

European Commission (2019). Commission Delegated Regulation (EU) 2019/2016 of 11 March 2019 supplementing Regulation (EU) 2017/1369 of the European Parliament and of the Council with regard to energy labelling of refrigerating appliances and repealing Commission Delegated Regulation (EU) No 1060/2010. https://eur-lex.europa.eu/eli/reg_del/2019/2016. Accessed 6th March, 2024.

European Commission (2023). 2030 climate targets. https://climate.ec.europa.eu/eu-action/climate-strategies-targets/2030-climate-targets_en. Accessed: 2025-03-12.

European Commission (2024). Energy poverty. https://energy.ec.europa.eu/topics/markets-and-consumers/energy-consumers-and-prosumers/energy-poverty_en. Accessed: 2025-03-17.

European Environmental Agency (2019). Assessing the effectiveness of EU policy on large combustion plants in reducing air pollutant emissions. https://www.eea.europa.eu/publications/effectiveness-of-eu-policy-on. Accessed 20th February, 2023.

European Parliament and European Council (2001). Directive 2001/80/EC of the European Parliament and of the Council of 23 October 2001 on the limitation of emissions of certain pollutants into the air from large combustion plants. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32001L0080. Accessed 20th February, 2023.

European Parliament and European Council (2003). Directive 2003/87/EC of the European Parliament and of the Council of 13 October 2003 establishing a scheme for greenhouse gas emission allowance trading within the Community and amending Council Directive 96/61/EC. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32003L0087. Accessed 20th February, 2023.

European Parliament and European Council (2006). Regulation (EC) No 166/2006 of the European Parliament and of the European Council. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32006R0166. Accessed 20th February, 2023.

European Union (2023). Directive (EU) 2023/1791 of the European Parliament and of the Council of 13 September 2023 on energy efficiency and amending Regulation (EU)

2023/955 (recast). https://eur-lex.europa.eu/eli/dir/2023/1791/oj/eng. Accessed: 2025-03-17.

Eurostat (2022a). Employment by sex, age and economic activity (LFSA_EGAN2_EEA). https://ec.europa.eu/eurostat/databrowser/view/LFSA_EGAN2__custom_1965790/default/table?lang=en. Accessed 25th January, 2022.

Eurostat (2022b). Gross value added and income by A*10 industry breakdowns (NAMA_10_A10). https://ec.europa.eu/eurostat/databrowser/view/NAMA_10_A10__custom_2987749/default/table?lang=en. Accessed: 2022-06-28.

Eurostat (2023). Electricity price statisics. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Electricity_price_statistics#Electricity_prices_for_household_consumers. Accessed: 2024-07-29.

Eurostat (2024). Quarterly greenhouse gas emissions in the EU. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Quarterly_greenhouse_gas_emissions_in_the_EU. Accessed: 2025-03-12.

Faberi, S., M. Presutto, R. Stamminger, R. Scialdoni, W. Mebane, and R. Esposito (2007). Preparatory studies for eco-design requirements of EUPS: Lot 13: Domestic refrigerators & freezers (Preparatory study). https://ekosuunnittelu.info/wp-content/uploads/2015/08/Lot_13_Final_Report_Taks_3-5.pdf. Accessed 21st May, 2024.

Fabra, N. and M. Reguant (2024). The energy transition: A balancing act. *Resource and Energy Economics 76*, 101408.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018). Global Evidence on Economic Preferences*. *The Quarterly Journal of Economics 133*(4), 1645–1692.

Falk, A., A. Becker, T. Dohmen, D. Huffman, and U. Sunde (2022). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *Management Science*.

Fang, X., L. Goette, B. Rockenbach, M. Sutter, V. Tiefenbeck, S. Schoeb, and T. Staake (2023). Complementarities in behavioral interventions: Evidence from a field experiment on resource conservation. *Journal of Public Economics 228*, 105028.

Fehr, D., G. Fink, and B. K. Jack (2022). Poor and rational: Decision-making under scarcity. *Journal of Political Economy 130*(11), 2862–2897.

Filipovic, Z. and A. F. Wagner (2023). The intangibles song in takeover announcements: Good tempo, hollow tune. *Swiss Finance Institute Research Paper* (19-03).

Finkelstein, A. and M. J. Notowidigdo (2019). Take-up and targeting: Experimental evidence from SNAP. *The Quarterly Journal of Economics 134*(3), 1505–1556.

Fischbacher, U., S. Schudy, and S. Teyssier (2021). Heterogeneous preferences and investments in energy saving measures. *Resource and Energy Economics 63*, 101202.

Fowlie, M., M. Greenstone, and C. Wolfram (2015). 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), 201–204.

Fowlie, M., M. Greenstone, and C. Wolfram (2018). Do energy efficiency investments deliver? Evidence from the Weatherization Assistance Program. *The Quarterly Journal of Economics 133(3)*, 1597–1644.

Francesconi, M. and J. James (2021). None for the road? Stricter drink driving laws and road accidents. *Journal of Health Economics 79*, 102487.

Frondel, M., G. Kussel, and S. Sommer (2019). Heterogeneity in the price response of residential electricity demand: A dynamic approach for Germany. *Resource and Energy Economics 57*, 119–134.

Gal, D. and D. D. Rucker (2018). The loss of loss aversion: Will it loom larger than its gain? *Journal of Consumer Psychology 28*(3), 497–516.

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

Gerarden, T., R. Newell, and R. Stavins (2017). Assessing the energy-efficiency gap. *Journal of Economic Literature 55(4)*, 1486–1525.

Germeshausen, R., B. Chlond, K. Tchorzewska, and K. von Graevenitz (2022). MEFINE: Mannheim European panel on Financial Indicators and Emissions. Mimeo.

Giandomenico, L., M. Papineau, and N. Rivers (2022). A systematic review of energy efficiency home retrofit evaluation studies. *Annual Review of Resource Economics 14*(1), 689–708.

Gillingham, K. and P. Huang (2021). Air pollution, health, and racial disparities: Evidence from ports.

Gillingham, K. and K. Palmer (2014). Bridging the energy efficiency gap: Policy insights from economic theory and empirical evidence. *Review of Environmental Economics and Policy*.

Gillingham, K. T., S. Houde, and A. A. Van Benthem (2021). Consumer myopia in vehicle purchases: Evidence from a natural experiment. *American Economic Journal: Economic Policy 13*(3), 207–238.

Giraudet, L.-G., S. Houde, and J. Maher (2018). Moral hazard and the energy efficiency gap: Theory and evidence. *Journal of the Association of Environmental and Resource Economists 5*(4), 755–790.

Goeschl, T. (2019). Cold case: The forensic economics of energy efficiency labels for domestic refrigeration appliances. *Energy Economics 84*, 104468.

Goldstein, B. A., M. Phelan, N. J. Pagidipati, R. R. Holman, M. J. Pencina, and E. A. Stuart (2019). An outcome model approach to transporting a randomized controlled trial results to a target population. *Journal of the American Medical Informatics Association 26*(5), 429–437.

Gonzales, M. H., E. Aronson, and M. A. Costanzo (1988). Using social cognition and persuasion to promote energy conservation: A quasi-experiment. *Journal of Applied Social Psychology 18*(12), 1049–1066.

Gopalan, M. and M. A. Pirog (2017). Applying behavioral insights in policy analysis: Recent trends in the United States. *Policy Studies Journal 45*(S1), S82–S114.

Grant, D. and A. W. Jones (2003). Are subsidiaries more prone to pollute? New evidence from the EPA's Toxic Release Inventory. *Social Science Quarterly 84*(1), 162–173.

Gravert, C. A. (2021). Reminders as a tool for behavior change. *Published in Behavioral Science in the Wild, Toronto, Canada: University of Toronto Press*.

Graziano, M. and K. Gillingham (2015). Spatial patterns of solar photovoltaic system adoption: the influence of neighbors and the built environment. *Journal of Economic Geography 15*(4), 815–839.

Guyton, J., D. S. Manoli, B. Schafer, and M. Sebastiani (2016). Reminders & recidivism: Evidence from tax filing & EITC participation among low-income nonfilers.

Hahn, R. W. and R. D. Metcalfe (2021). Efficiency and equity impacts of energy subsidies. *American Economic Review 111*(5), 1658–88.

Hallsworth, M. (2023). A manifesto for applying behavioural science. *Nature Human Behaviour 7*(3), 310–322.

Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics 148*, 14–31.

Halpern, D. and M. Sanders (2016). Nudging by government: Progress, impact & lessons learned. *Behavioral Science & Policy 2*(2), 53–65.

Hancevic, P. I. and H. H. Sandoval (2022). Low-income energy efficiency programs and energy consumption. *Journal of Environmental Economics and Management 113:102656.*

Harding, M. and A. Hsiaw (2014). Goal setting and energy conservation. *Journal of Economic Behavior & Organization 107*, 209–227.

Harris, R. and C. Robinson (2002). The effect of foreign acquisitions on total factor productivity: Plant-level evidence from UK manufacturing, 1987–1992. *Review of Economics and Statistics 84*(3), 562–568.

Harrison, G. W., M. I. Lau, and M. B. Williams (2002). Estimating individual discount rates in Denmark: A field experiment. *American Economic Review 92*(5), 1606–1617.

Harrison, G. W. and J. A. List (2004). Field experiments. *Journal of Economic Literature 42*(4), 1009—1055.

Harrison, G. W., K. Morsink, and M. Schneider (2020). Do no harm? The welfare consequences of behavioural interventions. *CEAR Working Paper 2020*.

Haskel, J. E., S. C. Pereira, and M. J. Slaughter (2007). Does inward foreign direct investment boost the productivity of domestic firms? *The Review of Economics and Statistics 89*(3), 482–496.

Hassett, K. A. and G. E. Metcalf (1995, jun). Energy tax credits and residential conservation investment: Evidence from panel data. *Journal of Public Economics 57*(2), 201–217.

Hastings, J. S. and J. M. Weinstein (2008). Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly Journal of Economics 123*(4), 1373–1414.

Haucap, J., A. Rasch, and J. Stiebale (2019). How mergers affect innovation: Theory and evidence. *International Journal of Industrial Organization 63*, 283–325.

Haushofer, J. and E. Fehr (2014). On the psychology of poverty. *Science 344*(6186), 862–867.

Hausman, C. and D. S. Rapson (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics 10*, 533–552.

Hausman, J. A. (1979). Individual discount rates and the purchase and utilization of energy-using durables. *The Bell Journal of Economics*, 33–54.

Heutel, G. (2019). Prospect theory and energy efficiency. *Journal of Environmental Economics and Management 96*, 236–254.

Holzmeister, F., J. Huber, M. Kirchler, and R. Schwaiger (2022). Nudging debtors to pay their debt: Two randomized controlled trials. *Journal of Economic Behavior & Organization 198*, 535–551.

Homar, A. R. and L. K. Cvelbar (2021). The effects of framing on environmental decisions: A systematic literature review. *Ecological Economics 183*, 106950.

Hossain, T. and J. Morgan (2006). ... plus shipping and handling: Revenue (non) equivalence in field experiments on ebay. *The BE Journal of Economic Analysis & Policy 6*(2), 0000102202153806371429.

Hotard, M., D. Lawrence, D. D. Laitin, and J. Hainmueller (2019). A low-cost information nudge increases citizenship application rates among low-income immigrants. *Nature Human Behaviour 3*(7), 678–683.

Hotz, V. J., G. W. Imbens, and J. H. Mortimer (2005). Predicting the efficacy of future training programs using past experiences at other locations. *Journal of Econometrics 125*(1-2), 241–270.

Houde, S. (2018a). Bunching with the stars: How firms respond to environmental certification. *CER-ETH–Center of Economic Research at ETH Zurich Working Paper 18*, 292.

Houde, S. (2018b). How consumers respond to product certification and the value of energy information. *The RAND Journal of Economics 49*(2), 453–477.

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

Houde, S. and E. Myers (2019). Heterogeneous (mis-) perceptions of energy costs: Implications for measurement and policy design. *National Bureau of Economic Research*.

Houde, S. and E. Myers (2021). Are consumers attentive to local energy costs? Evidence from the appliance market. *Journal of Public Economics 201*, 104480.

Hurley, M. (2007). Who's on whose margins? In *Researching the margins: Strategies for ethical and rigorous research with marginalised communities*, pp. 160–189. Springer.

Ifcher, J. and H. Zarghamee (2011). Happiness and time preference: The effect of positive affect in a random-assignment experiment. *American Economic Review 101*(7), 3109–3129.

Jacqz, I. (2021). Pollution and acquisition: Emissions and distributional effects of mergers.

Jaffe, A. and R. Stavins (1994). The energy efficiency gap: What does it mean? *Energy policy 22(10)*, 804–810.

Javorcik, B. (2004). Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. *American Economic Review 94*(3), 605–627.

Juranek, S., D. Schindler, and G. Schjelderup (2018). Transfer pricing regulation and taxation of royalty payments. *Journal of Public Economic Theory 20*(1), 67–84.

Juranek, S., D. Schindler, and A. Schneider (2023). Royalty taxation under tax competition and profit shifting. *Canadian Journal of Economics/Revue canadienne d'économique 56*(4), 1377–1412.

Kaplan, S. N. and M. S. Weisbach (1992). The success of acquisitions: Evidence from divestitures. *The Journal of Finance 47*(1), 107–138.

Karlan, D., M. McConnell, S. Mullainathan, and J. Zinman (2016). Getting to the top of mind: How reminders increase saving. *Management Science 62*(12), 3393–3411.

Kesselring, A. (2023). Energy labels in the European Union: Consumer inattention and producer responses. *Available at SSRN 4395430*.

Khanna, T. M., G. Baiocchi, M. Callaghan, F. Creutzig, H. Guias, N. R. Haddaway, L. Hirth, A. Javaid, N. Koch, S. Laukemper, et al. (2021). A multi-country meta-analysis on the role of behavioural change in reducing energy consumption and CO2 emissions in residential buildings. *Nature Energy 6*(9), 925–932.

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

Köster, T. (2023). *Arbeitslosigkeit unter inklusionstheoretischer Betrachtung nach Phelps: Eine Fallstudie*, Volume 19. Walter de Gruyter GmbH & Co KG.

Kőszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics 121*(4), 1133–1165.

Kraft-Todd, G. T., B. Bollinger, K. Gillingham, S. Lamp, and D. G. Rand (2018). Credibility-enhancing displays promote the provision of non-normative public goods. *Nature 563*(7730), 245–248.

Kühberger, A. (1998). The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes 75*(1), 23–55.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics 112*(2), 443–478.

Laibson, D. and J. A. List (2015). Principles of (behavioral) economics. *American Economic Review 105*(5), 385–390.

Leard, B. (2018). Consumer inattention and the demand for vehicle fuel cost savings. *Journal of Choice Modelling 29*, 1–16.

Lee, H., K. Calvin, D. Dasgupta, G. Krinner, A. Mukherji, P. Thorne, C. Trisos, J. Romero, P. Aldunce, K. Barret, et al. (2023). IPCC, 2023: Climate Change 2023: Synthesis Report, Summary for Policymakers. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.

Lees, J. (2014). *Ethnographier la précarité énergétique: Au-delà de l'action publique, des mises à l'épreuve de l'habiter*. Ph. D. thesis, EHESS.

Lerner, J., M. Sorensen, and P. Strömberg (2011). Private equity and long-run investment: The case of innovation. *The Journal of Finance 66*(2), 445–477.

Levine, D. I., T. Beltramo, G. Blalock, C. Cotterman, and A. M. Simons (2018). What impedes efficient adoption of products? Evidence from randomized sales offers for fuel-efficient cookstoves in Uganda. *Journal of the European Economic Association 16*(6), 1850–1880.

Levinson, A. (2019). Energy efficiency standards are more regressive than energy taxes: Theory and evidence. *Journal of the Association of Environmental and Resource Economists 6*(1), 7–36.

Levinson, A. and L. Sager (2023). Who values future energy savings? Evidence from American drivers. *Journal of the Association of Environmental and Resource Economists 10*(3), 717–751.

Levinson, A. and E. Silva (2022). The electric Gini: Income redistribution through energy prices. *American Economic Journal: Economic Policy 14*(2), 341–365.

Li, P., Y. Lu, and J. Wang (2020). The effects of fuel standards on air pollution: Evidence from China. *Journal of Development Economics 146:102488.*

Li, X. (2013). Productivity, restructuring, and the gains from takeovers. *Journal of Financial Economics 109*(1), 250–271.

Linos, E., A. Prohofsky, A. Ramesh, J. Rothstein, and M. Unrath (2022). Can nudges increase take-up of the EITC? Evidence from multiple field experiments. *American Economic Journal: Economic Policy 14*(4), 432–452.

List, J. A. (2020). Non est disputandum de generalizability? A glimpse into the external validity trial. *NBER Working Paper w27535*.

List, J. A. and M. K. Price (2009). The role of social connections in charitable fundraising: Evidence from a natural field experiment. *Journal of Economic Behavior & Organization 69*(2), 160–169.

List, J. A., A. M. Shaikh, and Y. Xu (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics 22*(4), 773–793.

Locke, E. A. and G. P. Latham (1990). *A theory of goal setting & task performance.* Prentice-Hall, Inc.

Löschel, A., M. Rodemeier, and M. Werthschulte (2023). Can self-set goals encourage resource conservation? Field experimental evidence from a smartphone app. *European Economic Review 160*, 104612.

Lusardi, A. and O. S. Mitchell (2014). The economic importance of financial literacy: Theory and evidence. *American Economic Journal: Journal of Economic Literature 52*(1), 5–44.

Maksimovic, V., G. Phillips, and N. R. Prabhala (2011). Post-merger restructuring and the boundaries of the firm. *Journal of Financial Economics 102*(2), 317–343.

Mamun, A., D. Mishra, and L. Zhan (2021). The value of intangible capital transfer in mergers. *Journal of Economics and Business 117*, 106020.

Mani, A., S. Mullainathan, E. Shafir, and J. Zhao (2013). Poverty impedes cognitive function. *Science 341(6149)*, 976–980.

Marchionni, C. and S. Reijula (2019). What is mechanistic evidence, and why do we need it for evidence-based policy? *Studies in History and Philosophy of Science Part A 73*, 54–63.

Marshall, A. (1890). *Principles of economics*. Macmillan and Company.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142(2)*, 698–714.

McGuckin, R. H. and S. V. Nguyen (1995). On productivity and plant ownership change: New evidence from the Longitudinal Research Database. *The RAND Journal of Economics*, 257–276.

Mills, B. and J. Schleich (2012). Residential energy-efficient technology adoption, energy conservation, knowledge, and attitudes: An analysis of European countries. *Energy Policy 49*, 616–628.

Mugerman, Y., O. Sade, and M. Shayo (2014). Long term savings decisions: Financial reform, peer effects and ethnicity. *Journal of Economic Behavior & Organization 106*, 235–253.

Mullainathan, S. and E. Shafir (2013). *Scarcity: Why having too little means so much*. Macmillan.

Munshi, K. and J. Myaux (2006). Social norms and the fertility transition. *Journal of Development Economics 80*(1), 1–38.

Myers, E. (2019). Are home buyers inattentive? Evidence from capitalization of energy costs. *American Economic Journal: Economic Policy 11*(2), 165–188.

Newell, R. G. and J. Siikamäki (2015). Individual time preferences and energy efficiency. *American Economic Review 105*(5), 196–200.

Newman, C., J. Rand, T. Talbot, and F. Tarp (2015). Technology transfers, foreign investment and productivity spillovers. *European Economic Review 76*, 168–187.

Ning, L. and F. Wang (2018). Does FDI bring environmental knowledge spillovers to developing countries? The role of the local industrial structure. *Environmental and Resource Economics 71*(2), 381–405.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6).

Ong, Q., W. Theseira, and I. Y. Ng (2019). Reducing debt improves psychological functioning and changes decision-making in the poor. *Proceedings of the National Academy of Sciences 116*(15), 7244–7249.

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

Palmer, K. and M. Walls (2015). Limited attention and the residential energy efficiency gap. *American Economic Review 105*(5), 192–95.

Palmer, K. L., M. Walls, and L. O'Keeffe (2015). Putting information into action: what explains follow-up on home energy audits? *Resources for the Future Discussion Paper*, 15–34.

Park, J., W. Son, H. Moon, and J. Woo (2023). Nudging energy efficiency behavior: The effect of message framing on implicit discount rate. *Energy Economics 117*, 106485.

Pearl, J. and E. Bareinboim (2011). Transportability of causal and statistical relations: A formal approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Volume 25, pp. 247–254.

Pigou, A. (1920). *The economics of welfare.*

Qiu, Y., G. Colson, and C. Grebitus (2014). Risk preferences and purchase of energy-efficient technologies in the residential sector. *Ecological Economics 107*, 216–229.

Rambachan, A. and J. Roth (2023). A more credible approach to parallel trends. *Review of Economic Studies 90*(5), 2555–2591.

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

Rodemeier, M. and A. Löschel (2022). Information nudges, subsidies, and crowding out of attention: Field evidence from energy efficiency investments. *SSRN*.

Rubin, E. and M. Auffhammer (2024). Quantifying heterogeneity in the price elasticity of residential natural gas. *Journal of the Association of Environmental and Resource Economists 11*(2), 319–357.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates. *The Quarterly Journal of Economics 116*(2), 681–704.

Sahari, A. (2019). Electricity prices and consumers' long-term technology choices: Evidence from heating investments. *European Economic Review 114*, 19–53.

Sallee, J. M. (2014). Rational inattention and energy efficiency. *The Journal of Law and Economics 57*(3), 781–820.

Sallee, J. M., S. E. West, and W. Fan (2016). Do consumers recognize the value of fuel economy? Evidence from used car prices and gasoline price fluctuations. *Journal of Public Economics 135*, 61–73.

Schleich, J. (2019). Energy efficient technology adoption in low-income households in the European Union - What is the evidence? *Energy Policy 125*, 196–206.

Schleich, J., X. Gassmann, C. Faure, and T. Meissner (2016). Making the implicit explicit: A look inside the implicit discount rate. *Energy Policy 97*, 321–331.

Schleich, J., X. Gassmann, T. Meissner, and C. Faure (2019). A large-scale test of the effects of time discounting, risk aversion, loss aversion, and present bias on household adoption of energy-efficient technologies. *Energy Economics 80*, 377–393.

Schmidheiny, K. and S. Siegloch (2020). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. CEPR Discussion Paper No. 13477.

Schulte, I. and P. Heindl (2017). Price and income elasticities of residential energy demand in Germany. *Energy Policy 102*, 512–528.

Shaghaghi, A., R. S. Bhopal, and A. Sheikh (2011). Approaches to recruiting 'hard-to-reach' populations into research: A review of the literature. *Health Promotion Perspectives 1*(2), 86.

Shah, A. K., S. Mullainathan, and E. Shafir (2012). Some consequences of having too little. *Science 338*(6107), 682–685.

Shleifer, A. (1998). State versus private ownership. *Journal of Economic Perspectives 12*(4).

Shu, S. B. and A. Gneezy (2010). Procrastination of enjoyable experiences. *Journal of Marketing Research 47*(5), 933–944.

Siegel, D. S. and K. L. Simons (2010). Assessing the effects of mergers and acquisitions on firm performance, plant productivity, and workers: New evidence from matched employer-employee data. *Strategic Management Journal 31*(8), 903–916.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 99–118.

Sirin, S. M., E. Camadan, I. E. Erten, and A. H. Zhang (2023). Market failure or politics? Understanding the motives behind regulatory actions to address surging electricity prices. *Energy Policy 180*, 113647.

Sorensen, A. T. (2006). Social learning and health plan choice. *The Rand Journal of Economics 37*(4), 929–945.

Statistisches Bundesamt (Destatis). Fachserie. 15, Wirtschaftsrechnungen. 1, Laufende Wirtschaftsrechnungen, Einkommen, Einnahmen und Ausgaben privater Haushalte. https://www.statistischebibliothek.de/mir/receive/DESerie_mods_00000152. Accessed: 2023-01-10.

Statistisches Bundesamt (Destatis). Verbraucherpreisindex für Deutschland – Lange Reihen ab 1948. https://www.destatis.de/DE/Themen/Wirtschaft/Preise/Verbraucherpreisindex/Publikationen/Downloads-Verbraucherpreise/

verbraucherpreisindex-lange-reihen-pdf-5611103.html. Accessed: 2022-05-19.

Statistisches Bundesamt (Destatis) (2018). Wohnen in Deutschland – Zusatzprogramm des Mikrozensus 2018. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Wohnen/Publikationen/Downloads-Wohnen/wohnen-in-deutschland-5122125189005.html. Accessed: 2022-05-19.

Statistisches Bundesamt (Destatis) (2020). Mikrozensus: Haushalte und Familien 2020. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Haushalte-Familien/Publikationen/_publikationen-innen-haushalte.html. Accessed: 2022-05-19.

Statistisches Bundesamt (Destatis) (2023a). Haushalte nach Haushaltsgröße und Haushaltsmitgliedern. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Haushalte-Familien/Tabellen/1-2-privathaushalte-bundeslaender.html. Accessed: 2023-11-21.

Statistisches Bundesamt (Destatis) (2023b). Tabelle 61243-0001: Strompreise für Haushalte: Deutschland, Halbjahre, Jahresverbrauchsklassen, Preisarten.

Statistisches Bundesamt (Destatis) (2023c). Umweltökonmische Gesamtrechnungen. Stromverbrauch der privaten Haushalte nach Haushaltsgrößenklassen. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Umwelt/UGR/private-haushalte/Tabellen/stromverbrauch-haushalte.html. Accessed: 2023-30-10.

Statistisches Bundesamt (Destatis) (2023d). Wohnfläche. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Wohnen/Tabellen/tabelle-wo4-wohnflaeche.html. Accessed: 2023-11-21.

Stojanovski, O., G. W. Leslie, F. A. Wolak, J. E. H. Wong, and M. C. Thurber (2020). Increasing the energy cognizance of electricity consumers in Mexico: Results from a field experiment. *Journal of Environmental Economics and Management 102*, 102323.

Stolper, O. and A. Walter (2019). Birds of a feather: The impact of homophily on the propensity to follow financial advice. *The Review of Financial Studies 32*(2), 524–563.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199.

Sussman, R., M. Chikumbo, and R. Gifford (2018). Message framing for home energy efficiency upgrades. *Energy and Buildings 174*, 428–438.

Thaler, R. H. and C. R. Sunstein (2021). *Nudge*. Yale University Press.

Thomson Reuters (2019a). Number of M&A deals in Europe. Extracted from https://de.statista.com/statistik/daten/studie/972513/umfrage/anzahl-der-munda-deals-in-europa/. Accessed: 2022-06-28.

Thomson Reuters (2019b). Volume of M&A deals in Europe. Extracted from https://de.statista.com/statistik/daten/studie/972512/umfrage/volumen-der-munda-deals-in-europa/. Accessed: 2022-06-28.

Tonn, B., D. Carroll, S. Pigg, M. Blasnik, G. Dalhoff, J. Berger, E. Rose, B. Hawkins, J. Eisenberg, F. Ucar, et al. (2014). Weatherization works – Summary of findings from the retrospective evaluation of the US Department of Energy's Weatherization Assistance Program.

Train, K. (1985). Discount rates in consumers' energy-related decisions: A review of the literature. *Energy 10*(12), 1243–1253.

Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science 185*(4157), 1124–1131.

United Nations (2025). Climate Change – United Nations. https://www.un.org/en/global-issues/climate-change. Accessed: 2025-03-12.

US Department of Housing and Urban Development (2024). The Green and Resilient Retrofit Program (GRRP). https://www.hud.gov/GRRP. Accessed: 2024-01-08.

U.S. Energy Information Administration (2023). December 2023. *Electric Power Monthly*, 177.

U.S. Energy Information Administration (EIA) (2015). Air conditioning accounts for about 12% of U.S. home energy expenditures, 2018. https://www.eia.gov/todayinenergy/detail.php?id=36692. Accessed: 2022-05-19.

Verband der Elektro- und Digitalindustrie (ZVEI). Absatz von Kuehlgeraeten in Deutschland in den Jahren 2014 bis 2021. https://de.statista.com/statistik/daten/studie/557159/umfrage/absatz-von-kuehlgeraeten-in-deutschland/. Accessed: 2023-01-10.

Wang, J. and S. Matsumoto (2021). An economic model of home appliance replacement: Application to refrigerator replacement among Japanese households. *Environmental Economics and Policy Studies 24*, 29–48.

Wang, J. and X. Wang (2015). Benefits of foreign ownership: Evidence from foreign direct investment in China. *Journal of International Economics 97*(2), 325–338.

Warner, J. T. and S. Pleeter (2001). The personal discount rate: Evidence from military downsizing programs. *American Economic Review 91*(1), 33–53.

Westreich, D., J. K. Edwards, C. R. Lesko, E. Stuart, and S. R. Cole (2017). Transportability of trial results using inverse odds of sampling weights. *American Journal of Epidemiology 186*(8), 1010–1014.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters 104*(3).

Zephyr Database (2023). Zephyr M&A Deals and Rumors. Extracted from `https://www.bvdinfo.com/en-us/our-products/data/greenfield-investment-and-ma/zephyr?gclid=EAIaIQobChMIvsK6-OCt_QIVRoxoCR098AGQEAAYASAAEgJnPfD_BwE`. Accessed: 2022-06-23.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and statistics 85*(1), 9–23.

Chapter A2

Appendix for Chapter 2

A2.1 Tables

Table A2.1: McCrary Test results: procedural variation cutoff

Procedural variation	-1,392.44	-379.58	-1,003.81	-285.94	-153.20	-399.03	-319.24
(EE-RIG = 1)	(1,818.82)	(597.64)	(771.83)	(171.30)	(198.25)	(266.33)	(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 on January 1, 2018 when the subsidy level changes.

Table A2.2: McCrary Test results: subsidy variation cutoff

Subsidy variation	901.097	527.033	236.981	463.678	183.923	60.653	136.028
(€ 150 = 1)	(1,369.933)	(341.001)	(353.834)	(54.090)	(94.996)	(107.208)	(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 on February 1, 2019 when the enrolment procedure and voucher terms change. EE-RIG denotes the elective enrolment mode with rigid voucher terms.

Table A2.3: Mean comparison of eligible households before vs. after subsidy and procedural changes

	Procedural variation		
	Mean before 01 Jan 2009 - 31 Dec 2017	Mean after 01 Jan 2018 - 31 Jan 2019	Difference
Household variables			
No. inhabitants	2.97	2.68	0.30
	(0.02)	(0.01)	(0.02)
Living space in m^2	70.62	68.64	1.98
	(0.21)	(0.11)	(0.10)
Electricity price per kWh	0.276	0.274	0.001
	(0.0002)	(0.0001)	(0.0002)
Annual electricity consumption in kWh	3,069.5	3,000.1	69.4
	(15.9)	(8.7)	(18.2)
Old refrigerator variables			
Age in years	18.2	16.9	1.3
	(0.04)	(0.02)	(0.04)
Volume in liters	246.0	232.7	13.3
	(0.6)	(0.3)	(0.7)
Estimated savings from replacement	341.2	336.6	4.7
	(1.4)	(0.8)	(1.6)
	Subsidy variation		
	Mean before 01 Jan 2018 - 31 Jan 2019	Mean after 01 Feb 2019 - July 2019	Difference
Household variables			
No. inhabitants	2.97	2.96	0.009
	(0.02)	(0.01)	(0.02)
Living space in m^2	70.62	70.74	-0.12
	(0.21)	(0.20)	(0.29)
Electricity price per kWh	0.276	0.282	-0.006
	(0.0002)	(0.0001)	(0.0002)
Annual electricity consumption in kWh	3,069.5	3,136.1	-66.6
	(15.9)	(83.1)	(18.2)
Old refrigerator variables			
Age in years	18.2	17.8	0.4
	(0.04)	(0.04)	(0.06)
Volume in liters	246.0	251.5	-5.5
	(0.6)	(0.7)	(0.9)
Estimated savings from replacement	341.2	330.9	10.3
	(1.4)	(1.3)	(1.9)

Notes: * p<0.05, ** p<0.01, *** p<0.001.

Table A2.4: Mean comparison of households that replace refrigerators before/after procedural change

	Procedural variation - full sample		
	Mean before July 2017 - December 2017	Mean after January 2018 - June 2018	Difference
Household variables			
No. inhabitants	2.89	3.00	0.10
Living space in m^2	70.1	71.0	0.8
Electricity price per kWh	0.28	0.28	0.00
Annual electricity consumption in kWh	2,895	2,965	70
Old refrigerator variables			
Age in years	17.9	18.4	0.5
Volume in liters	240.9	249.3	8.5
Estimated savings from replacement	341.9	349.8	7.9
New refrigerator variables			
Price	462.0	475.5	13.5
Volume in liters	258.6	268.2	9.6
KWh consumption	138.8	141.7	3.0

	Procedural variation - non-NRW sample		
	Mean before July 2017 - December 2017	Mean after Jan 2018 - June 2018	Difference
Household variables			
No. inhabitants	2.45	2.63	0.19
Living space in m^2	65.4	66.8	1.4
Electricity price per kWh	0.28	0.28	0.00
Annual electricity consumption in kWh	2,467	2,509	42
Old refrigerator variables			
Age in years	18.1	18.6	0.5
Volume in liters	239.1	249.5	10.4
Estimated savings from replacement	331.5	334.2	2.6
New refrigerator variables			
Price	446.7	461.4	14.7
Volume in liters	241.2	257.5	16
KWh consumption	133.1	138.7	5.6

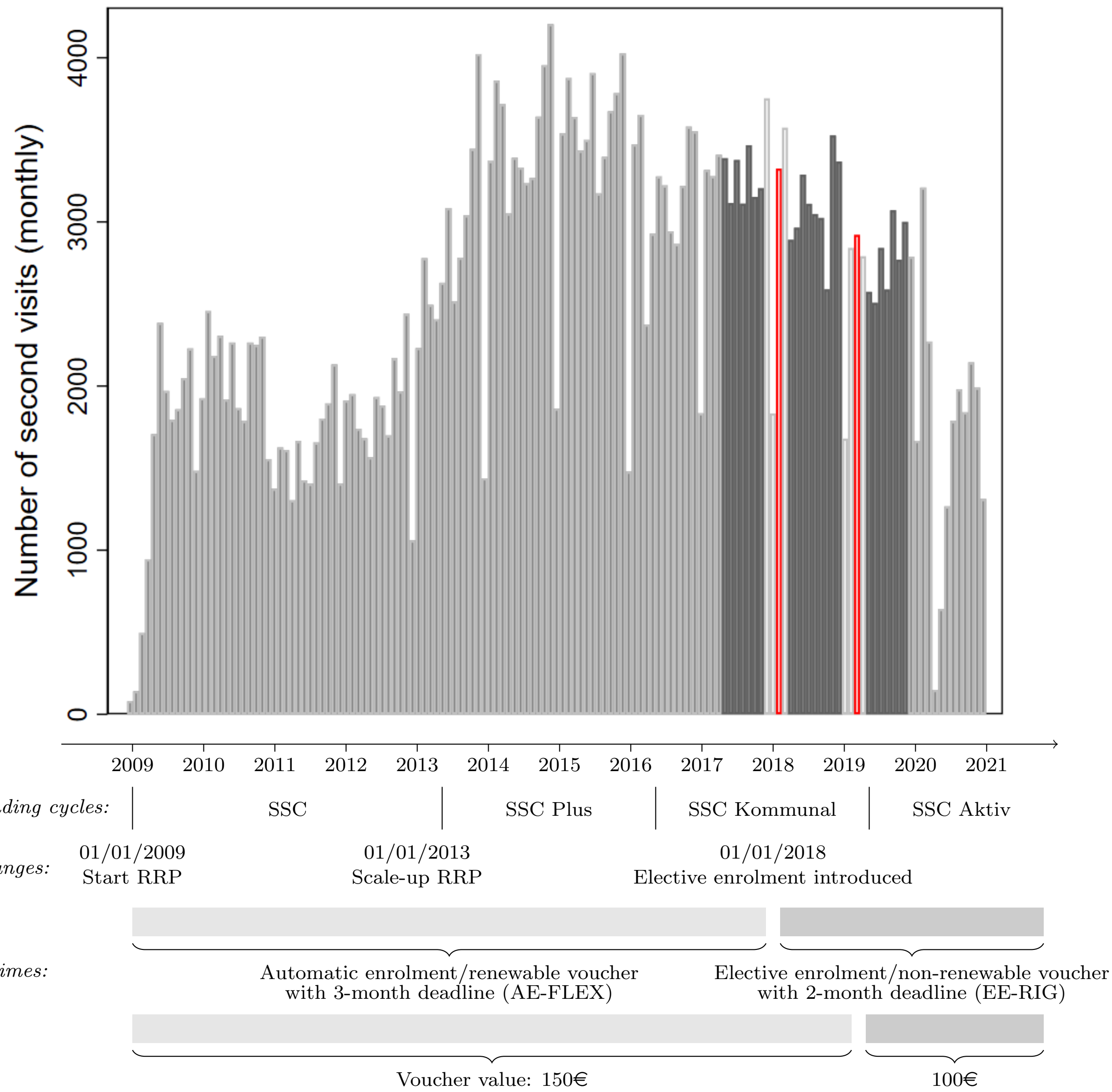
Table A2.5: Mean comparison of households that replace refrigerators before/after subsidy change

	Subsidy variation - full sample		
	Mean before August 2018 - January 2019	Mean after February 2019 - July 2019	Difference
Household variables			
No. inhabitants	3.10	2.92	-0.18
Living space in m^2	72.2	69.8	-2.4
Electricity price per kWh	0.28	0.28	0.00
Annual electricity consumption in kWh	2,964	2,829	-135
Old refrigerator variables			
Age in years	18.6	18.2	-0.4
Volume in liters	249.6	260.0	10.3
Estimated savings from replacement	357.8	342.8	-15.1
New refrigerator variables			
Price	490.1	483.2	6.8
Volume in liters	277.2	278.0	0.8
KWh consumption	145.4	145.8	0.4

	Subsidy variation - non-NRW sample		
	Mean before August 2018 - January 2019	Mean after February 2019 - July 2019	Difference
Household variables			
No. inhabitants	2.65	2.62	-0.04
Living space in m^2	67.5	66.5	-1.0
Electricity price per kWh	0.28	0.28	0.00
Annual electricity consumption in kWh	2,545	2,593	47
Old refrigerator variables			
Age in years	19.0	18.5	-0.4
Volume in liters	247.7	255.4	7.7
Estimated savings from replacement	346.7	341.7	-5.1
New refrigerator variables			
Price	465.7	466.5	0.8
Volume in liters	260.7	262.9	2.2
KWh consumption	140.5	140.6	0.1

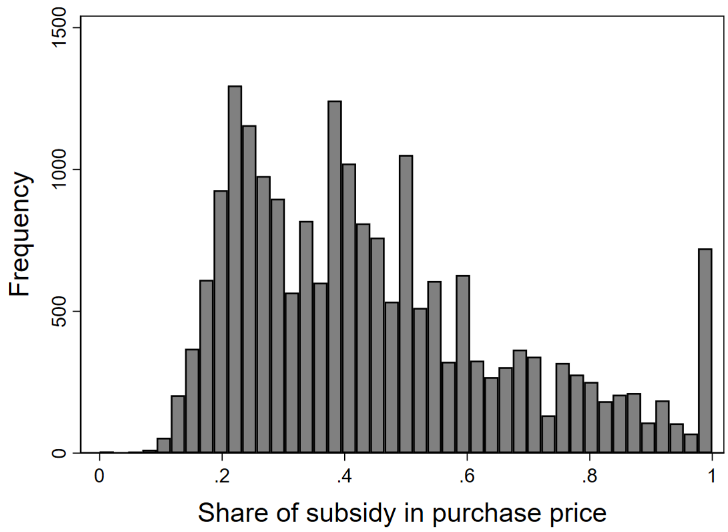
A2.2 Figures

Figure A2.1: Audit distribution over sample period and program timeline



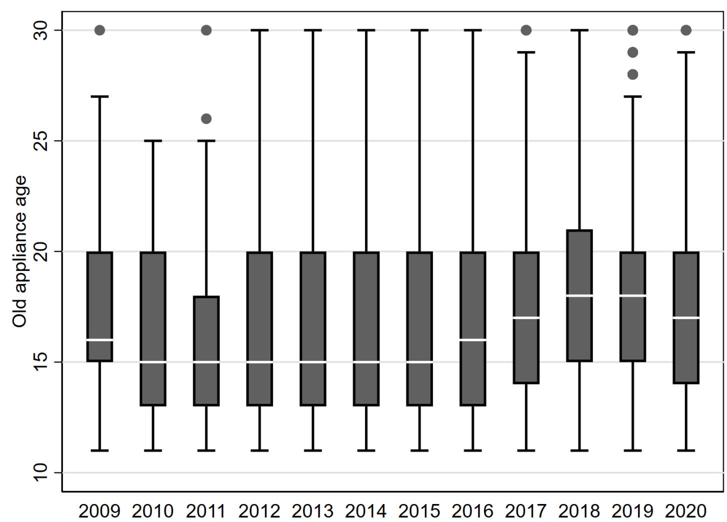
Notes: This figure shows the distribution of audits in the program over time, each bar mapping one month from January 2009 to December 2020. The number of monthly audits increases up to 2015 and remains on the high level until it slightly decreases from 2018 on. The dramatic dip during the second quarter 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. To estimate the effect of varying procedures and subsidy levels (see red-colored bars), in our main RDD specification we use data from April 2017 to October 2019 (see dark-colored bars) leaving out data in the interim periods directly before and after the program changes (see light gray-colored bars). 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.

Figure A2.2: Share of subsidy in purchase price of new refrigerator



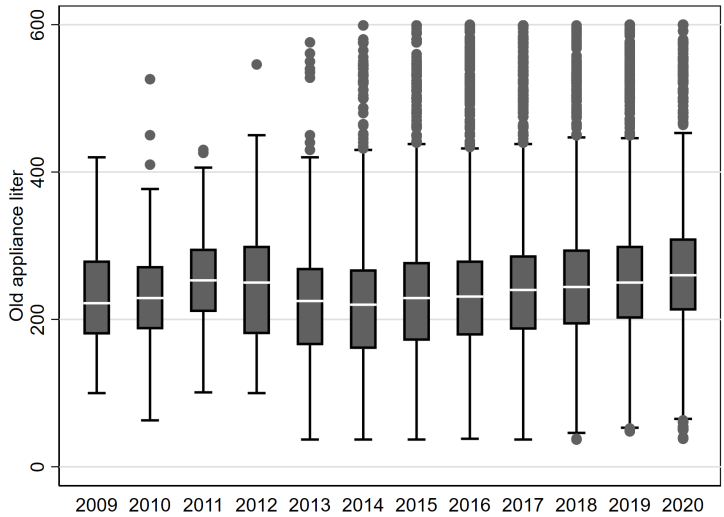
Notes: This figure shows the share that the replacement subsidy covers of the total purchase price of the new refrigerator. The subsidies considered here include the federal subsidy of €150 up to 2017 and €100 as of 2018 respectively as well as the complementary programs by four state governments as listed in Section 2.2.

Figure A2.3: Age in years of old refrigerators



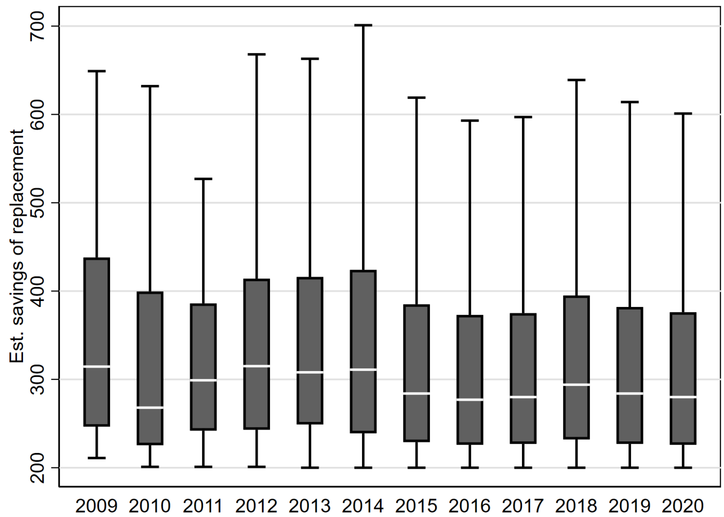
Notes: This figure shows the age distribution of old refrigerators for each year in the sample period separately. The figure was created with the sample of for replacement eligible households.

Figure A2.4: Volume in liter of old refrigerators



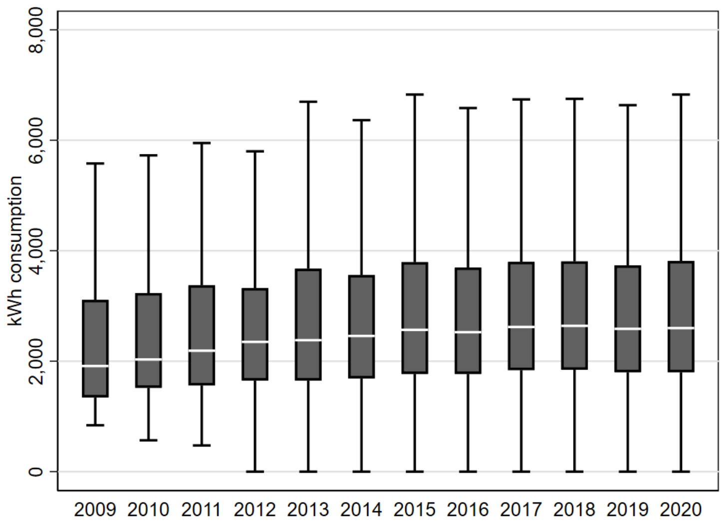
Notes: This figure shows the volume distribution of old refrigerators in liter for each year in the sample period separately. The figure was created with the sample of for replacement eligible households.

Figure A2.5: Estimated savings after replacement



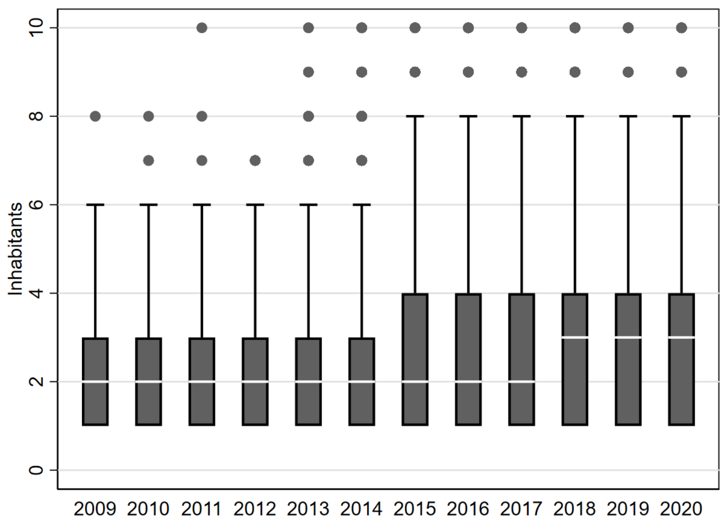
Notes: This figure shows the distribution of estimated savings after replacement for each year in the sample period separately. Note that outliers are not shown. The figure was created with the sample of for replacement eligible households.

Figure A2.6: Annual electricity consumption of households



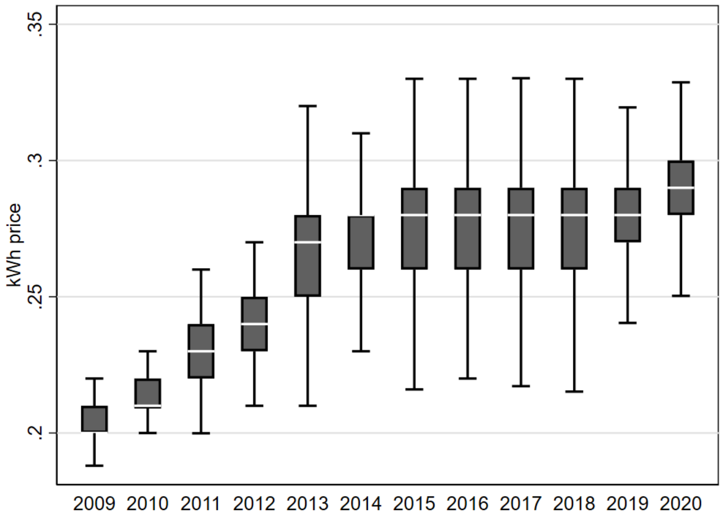
Notes: This figure shows the distribution of the annual electricity consumption of households for each year in the sample period separately. Note that outliers are not shown. The figure was created with the sample of for replacement eligible households.

Figure A2.7: Household structure



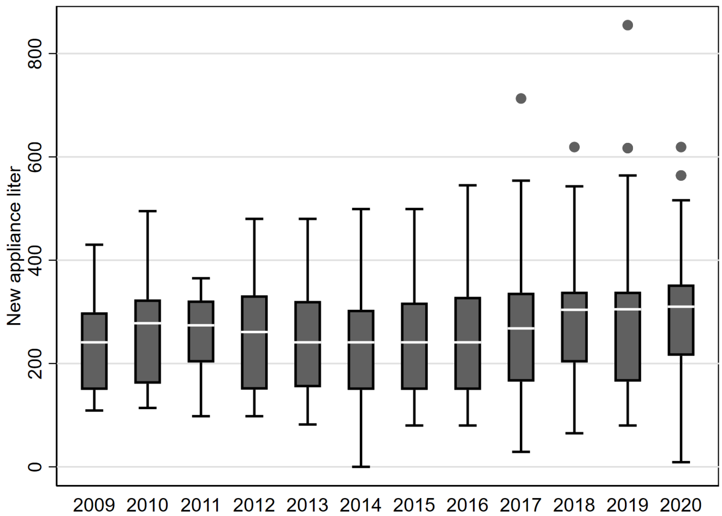
Notes: This figure shows the distribution of the number of inhabitants for each year in the sample period separately. The figure was created with the sample of for replacement eligible households.

Figure A2.8: Electricity price per kWh



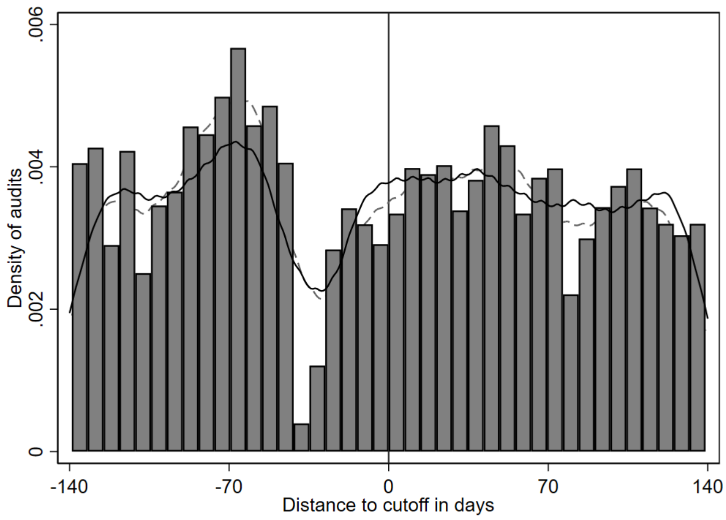
Notes: This figure shows the electricity price per kWh that households pay for each year in the sample period separately. The figure was created with the sample of for replacement eligible households.

Figure A2.9: Volume in liter of new refrigerator



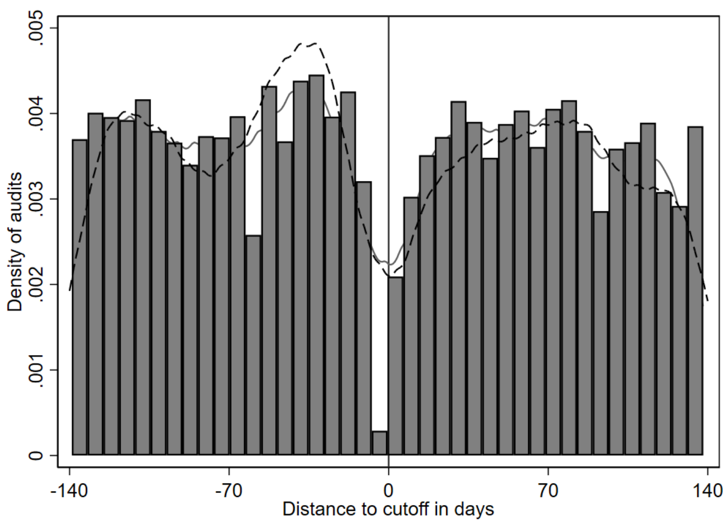
Notes: This figure shows the volume in liter of new appliances that households buy for each year in the sample period separately. The figure was created with the sample of households that replace their refrigerators.

Figure A2.10: Audit density around the change in the subsidy level



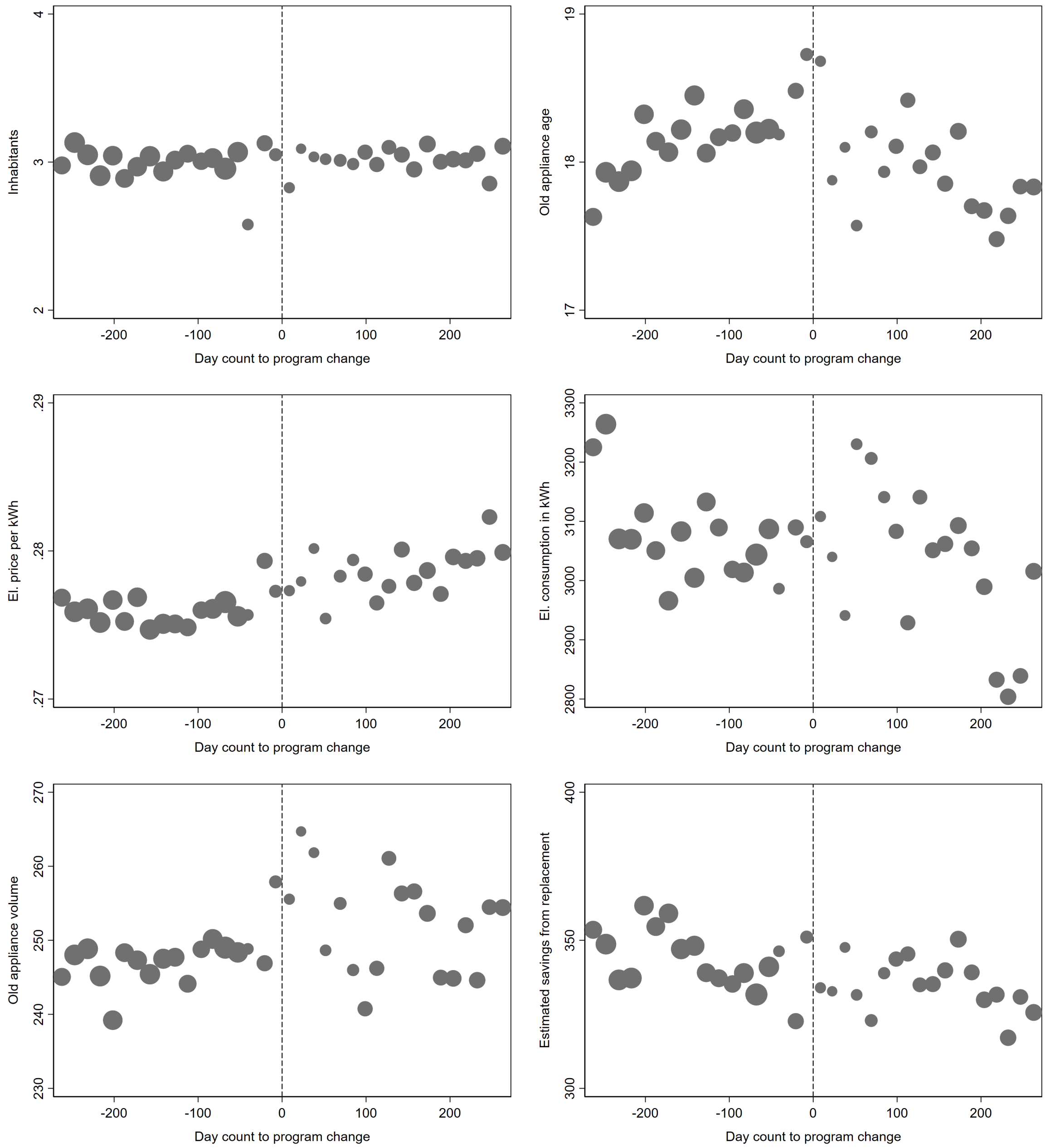
Notes: This figure shows the density (bars) and Kernel density (dashed line) of audits (second home visits) in a bandwidth of 20 weeks around the regime change. No bunching is apparent on either side of the cutoff (we would expect bunching to occur on the left side if households wanted to sort themselves into the regime with the higher subsidy level). A sharp drop in the audit density 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 to two weeks. To demonstrate that this pattern is usual we also provide the Kernel density of audits in the year before (2018) during the same season (solid line). Both Kernel densities are well aligned.

Figure A2.11: Audit density around the procedural change



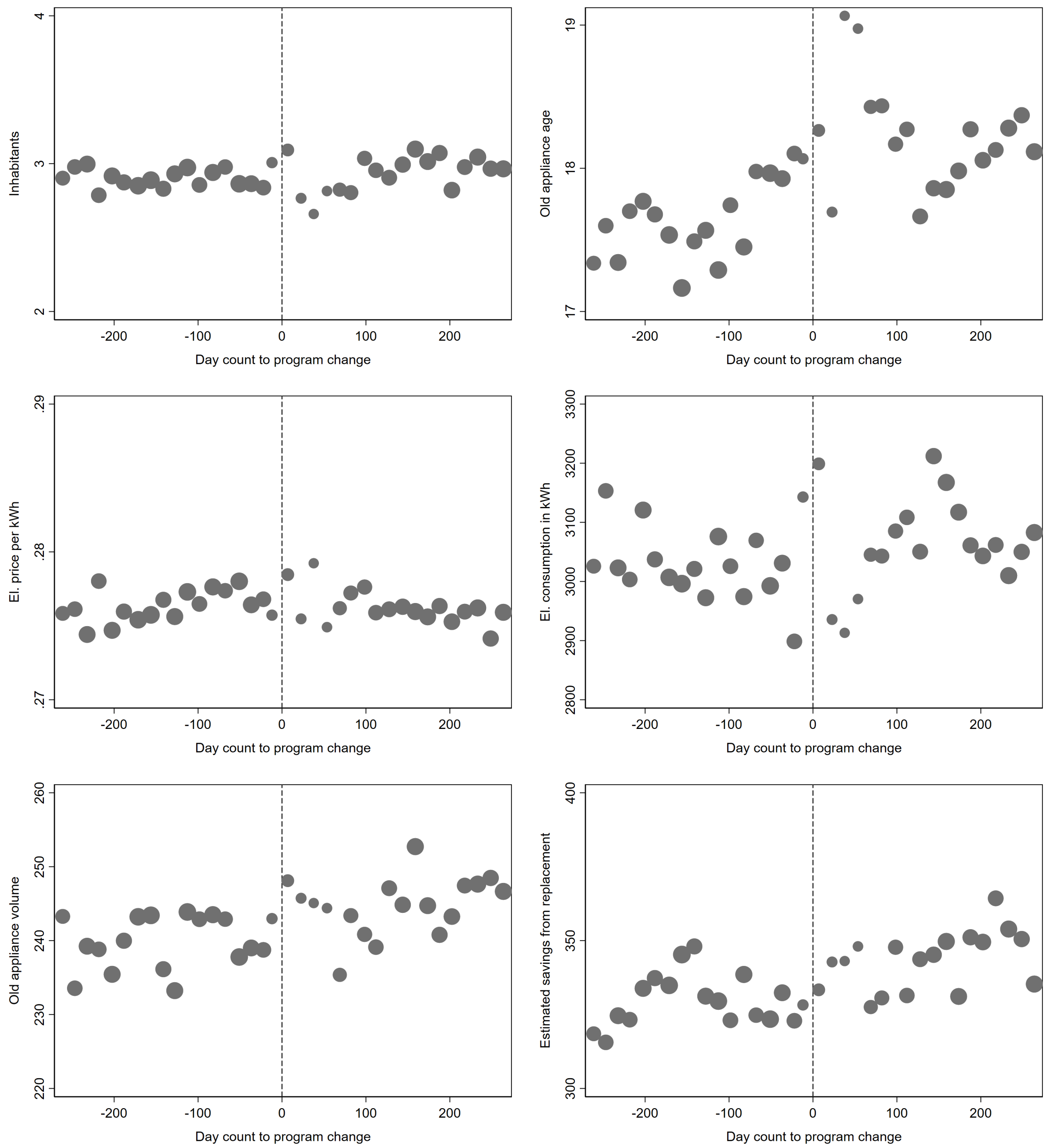
Notes: This figure shows the density (bars) and Kernel density (solid line) of audits (second home visits) in a bandwidth of 20 weeks around the regime change. A sharp drop in density appears directly before the regime change which coincides with the Christmas and end-of-year break when most local branches close for one to two weeks.We do not expect bunching to occur; the more attractive renewable 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 cutoff. To demonstrate that this pattern is usual we also provide the Kernel density of audits in the year after (2019) during the same season (dashed line). Both Kernel densities are well aligned.

Figure A2.12: Discontinuity check of covariates at change in subsidy level



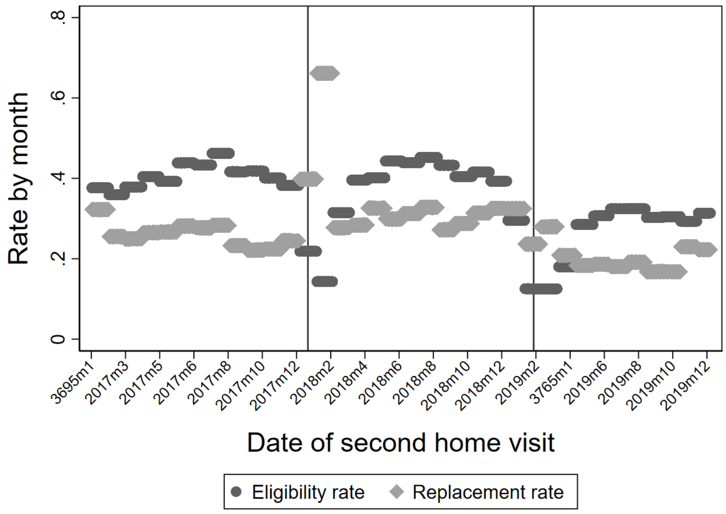
Notes: The figures show 15-day bin averages of household covariates in a bandwidth of 270 days around the change in the subsidy level.

Figure A2.13: Discontinuity check of covariates at change in program procedures



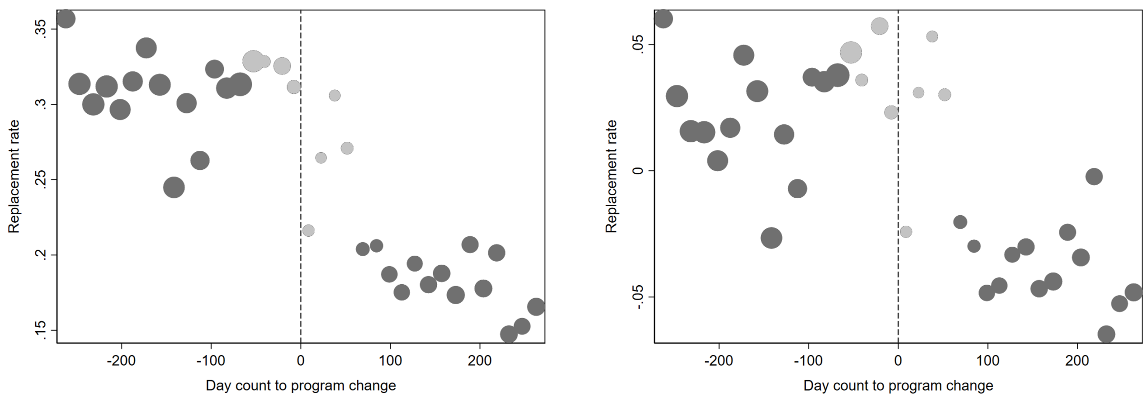
Notes: The figures show 15-day bin averages of household covariates in a bandwidth of 270 days around the change in program procedures.

Figure A2.14: Eligibility ratio around the program changes



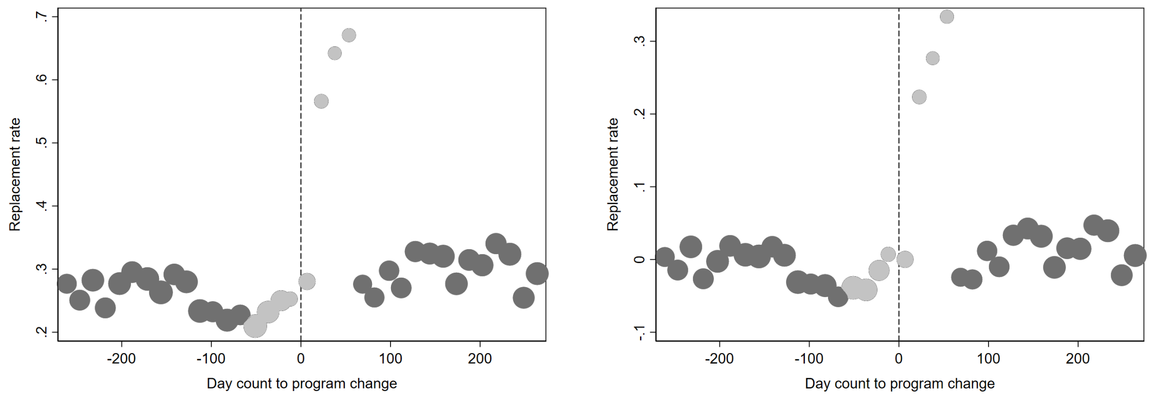
Notes: This figure shows the monthly ratio of households that are found eligible for replacement and receive an invitation letter out of all audited households. Around both regime changes (procedural change and change in the subsidy level), the eligibility ratio (in dark-gray) drops considerably. In the data, we observe that this is not driven by fewer households complying with the eligibility criteria for the replacement scheme (refrigerator older than 10 years, annual estimated savings of at least 200 kWh). Rather, not all households that fulfill the criteria receive an invitation 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 information letter at the first regime change and due to the end and start of a new funding phase at the second regime change.

Figure A2.15: Subsidy variation, replacement rate: Interim period included



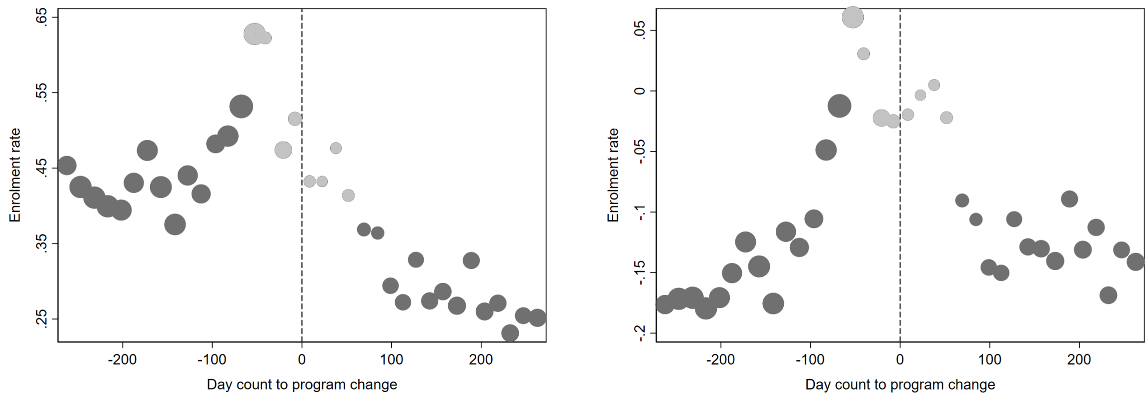
Notes: This figure shows the rate of households that successfully complete refrigerator replacement out of all households that are eligible for the RRP over time around the reduction of the voucher value by €50 on April 1, 2019. The left figure shows the raw data plots, the right figure shows the residuals as a result from the ALL first stage. Replacement rates are binned and averaged over 15 days in a bandwidth of 270 days around the program change. The light-grey dots mark the data that fall in the interim period of 2 months around the change. In our main specification, we exclude this data from the analysis in a RD Donut Design. For a detailed discussion of eligibility and replacement rates in the interim period, see Section 2.4.

Figure A2.16: Procedures variation, replacement rate: Interim period included



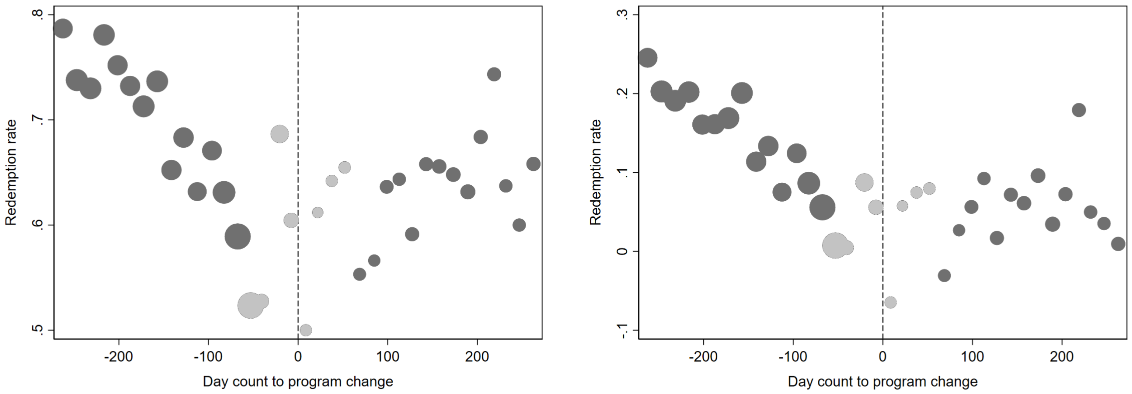
Notes: This figure shows the rate of households that successfully complete refrigerator replacement out of all households that are eligible for the RRP over time around the change in procedures on January 1, 2018. The left figure shows the raw data plots, the right figure shows the residuals as a result from the ALL first stage. Replacement rates are binned and averaged over 15 days in a bandwidth of 270 days around the program change. The light-grey dots mark the data that fall in the interim period of 2 months around the change. In our main specification, we exclude this data from the analysis in a RD Donut Design. For a detailed discussion of eligibility and replacement rates in the interim period, see Section 2.4.

Figure A2.17: Subsidy variation, enrolment rate: Interim period included



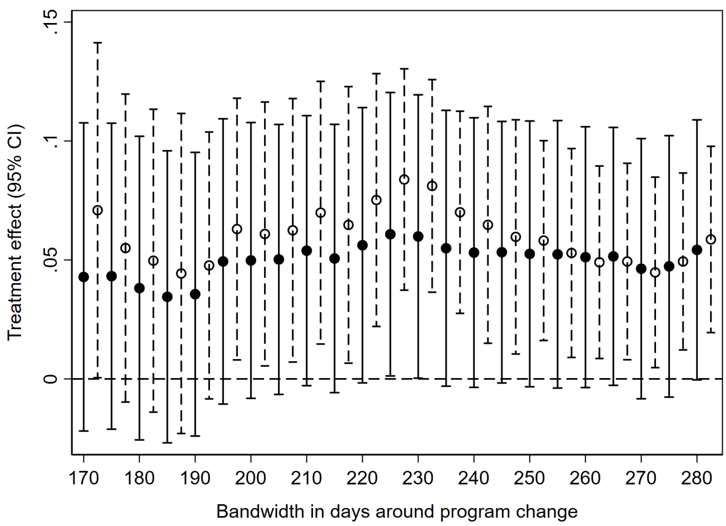
Notes: This figure shows the rate of households that request a voucher for refrigerator replacement and enrol in the program out of all households that are eligible for the RRP over time around the reduction of the voucher value by €50 on April 1, 2019. The left figure shows the raw data plots, the right figure shows the residuals as a result from the ALL first stage. Enrolment rates are binned and averaged over 15 days in a bandwidth of 270 days around the program change. The light-grey dots mark the data that fall in the interim period of 2 months around the change. In our main specification, we exclude this data from the analysis in a RD Donut Design. For a detailed discussion of eligibility and replacement rates in the interim period, see Section 2.4.

Figure A2.18: Subsidy variation, redemption rate: Interim period included



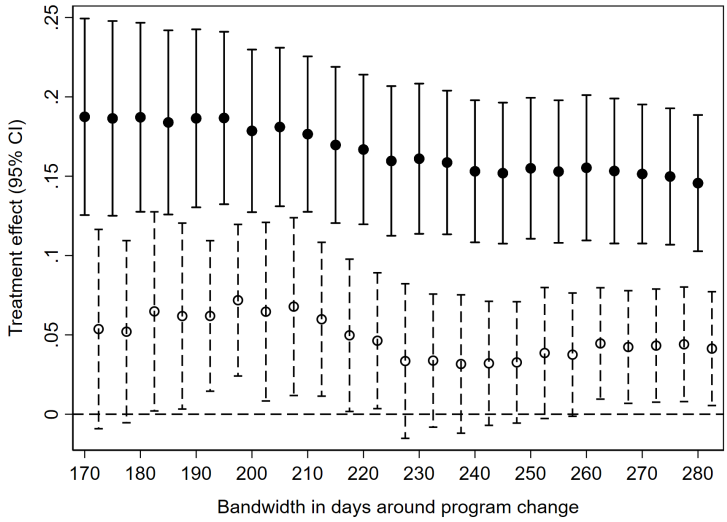
Notes: This figure shows the rate of households that successfully complete refrigerator replacement out of all households that have requested a voucher and enroled into the program over time around the reduction of the voucher value by €50 on April 1, 2019. The left figure shows the raw data plots, the right figure shows the residuals as a result from the ALL first stage. Redemption rates are binned and averaged over 15 days in a bandwidth of 270 days around the program change. The light-grey dots mark the data that fall in the interim period of 2 months around the change. In our main specification, we exclude this data from the analysis in a RD Donut Design. For a detailed discussion of eligibility and replacement rates in the interim period, see Section 2.4.

Figure A2.19: Effect of subsidy change on replacement as function of bandwidth



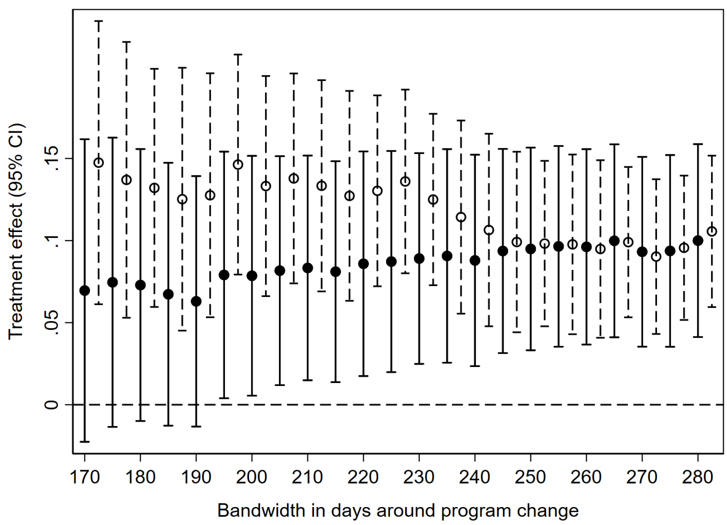
Notes: This figure shows the effect of the change in subsidy value (+€50) on the replacement rate for different bandwidths. The coefficients with filled dots and solid confidence interval markers show the effect for model 2 and the coefficients with hollow dots and dashed confidence interval markers show the effect for model 3 in Table 2.2.

Figure A2.20: Effect of procedural change on replacement as function of bandwidth



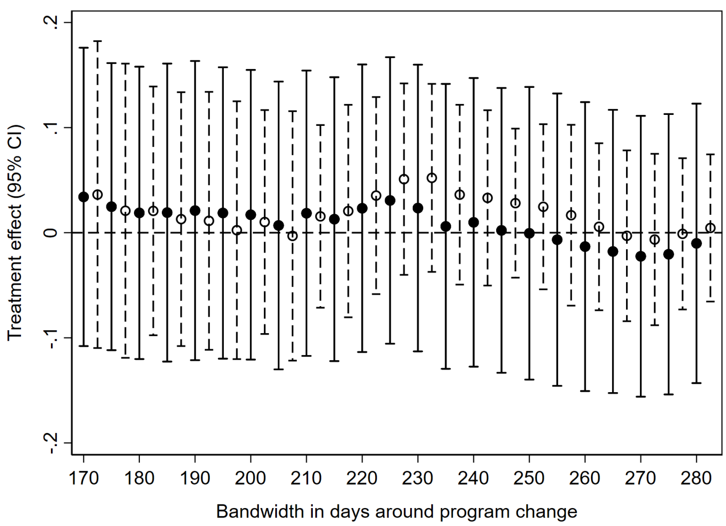
Notes: This figure shows the effect of the change in procedures from automatic enrolment and flexible voucher terms to elective enrolment and rigid voucher terms on the replacement rate for different bandwidths. The coefficients with filled dots and solid confidence interval markers show the effect for model 2 and the coefficients with hollow dots and dashed confidence interval markers show the effect for model 3 in Table 2.2.

Figure A2.21: Effect of subsidy change on enrolment as function of bandwidth



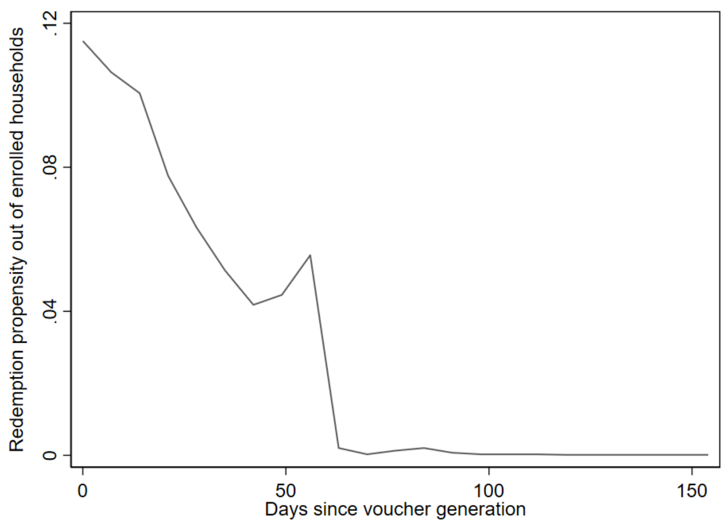
Notes: This figure shows the effect of the change in subsidy value (+€ 50) on the enrolment rate for different bandwidths. The coefficients with filled dots and solid confidence interval markers show the effect for model 2 and the coefficients with hollow dots and dashed confidence interval markers show the effect for model 3 in Table 2.4.

Figure A2.22: Effect of subsidy change on redemption as function of bandwidth



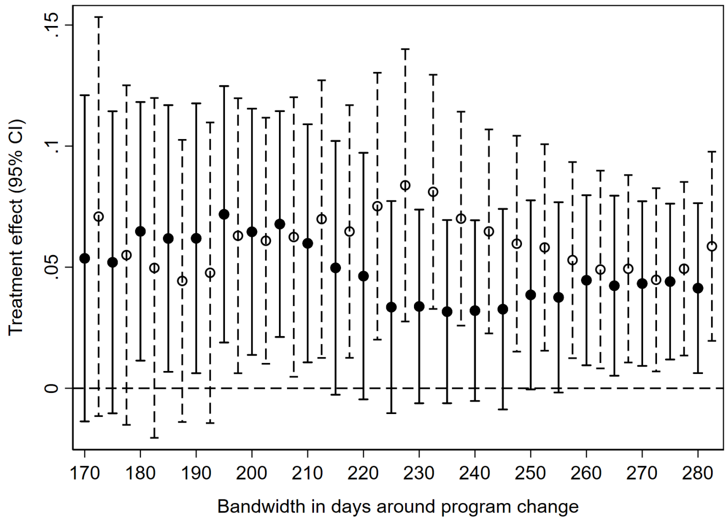
Notes: This figure shows the effect of the change in subsidy value (+€ 50) on the redemption rate for different bandwidths. The coefficients with filled dots and solid confidence interval markers show the effect for model 2 and the coefficients with hollow dots and dashed confidence interval markers show the effect for model 3 in Table 2.5.

Figure A2.23: Redemption propensity of enrolled households in EE-RIG



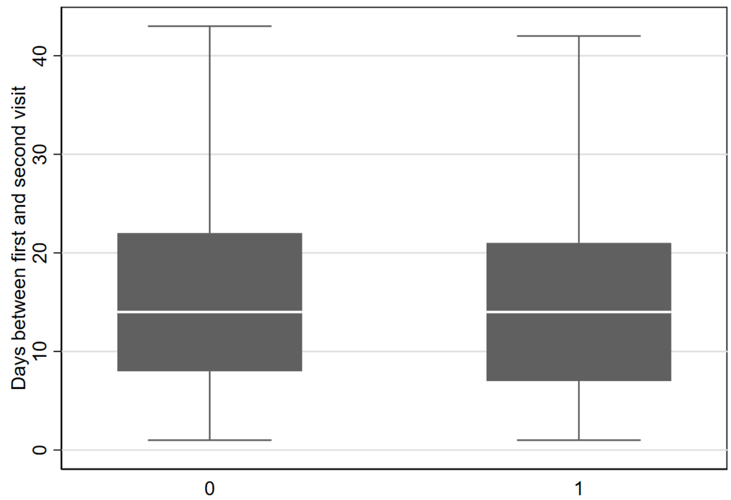
Notes: This figure shows the propensity of enrolled households to redeem the voucher as function of the days passed since the voucher was generated for the sample of enrolled households in the period January 2018 to January 2019.

Figure A2.24: Comparison of subsidy and procedural effect (replacement rate) as function of bandwidth



Notes: This figure shows the effect of the change in subsidy value (+€ 50) (hollow dots, dashed confidence interval markers) and the effect of the change in procedures from automatic enrolment and flexible voucher terms to elective enrolment and rigid voucher terms (filled dot, solid confidence interval markers) for different bandwidth choices for model 9 in Tables 2.2 and 2.3, respectively.

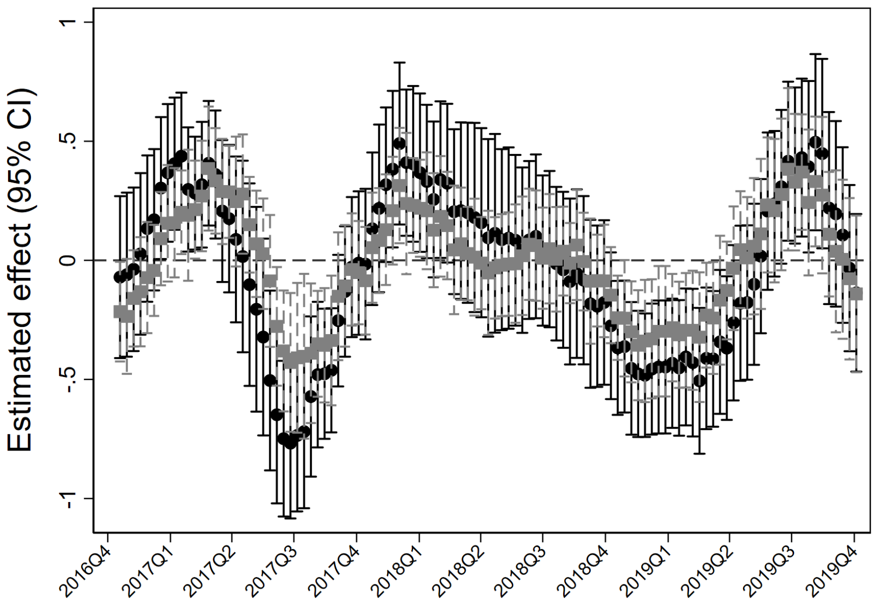
Figure A2.25: Visit patterns in the RRP under different procedures



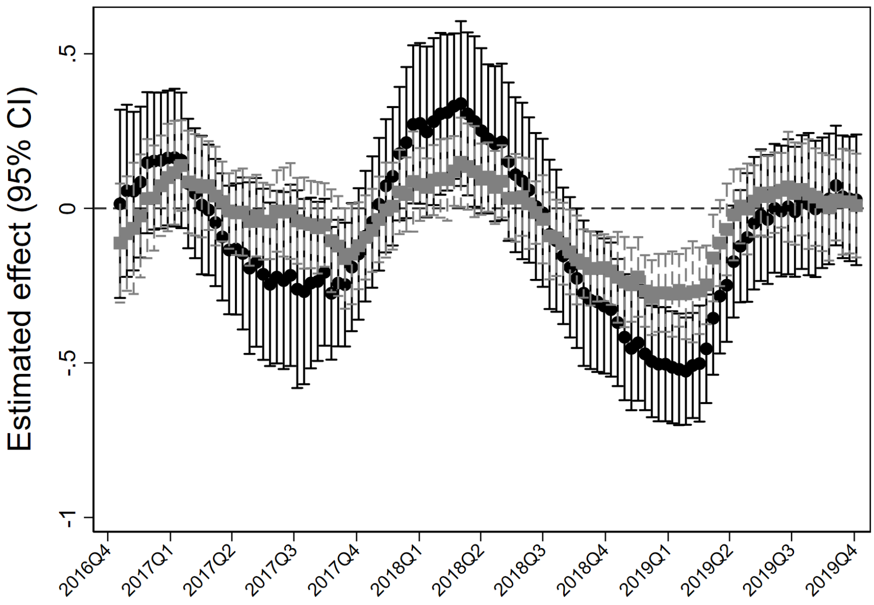
Notes: This figure shows a box plot of days passed from first to second visit. AE-FLEX (0) procedures are on the left, EE-RIG (1) procedures on the right.

A2.3 Robustness checks

Figure A2.26: Placebo test July 2016 – December 2019



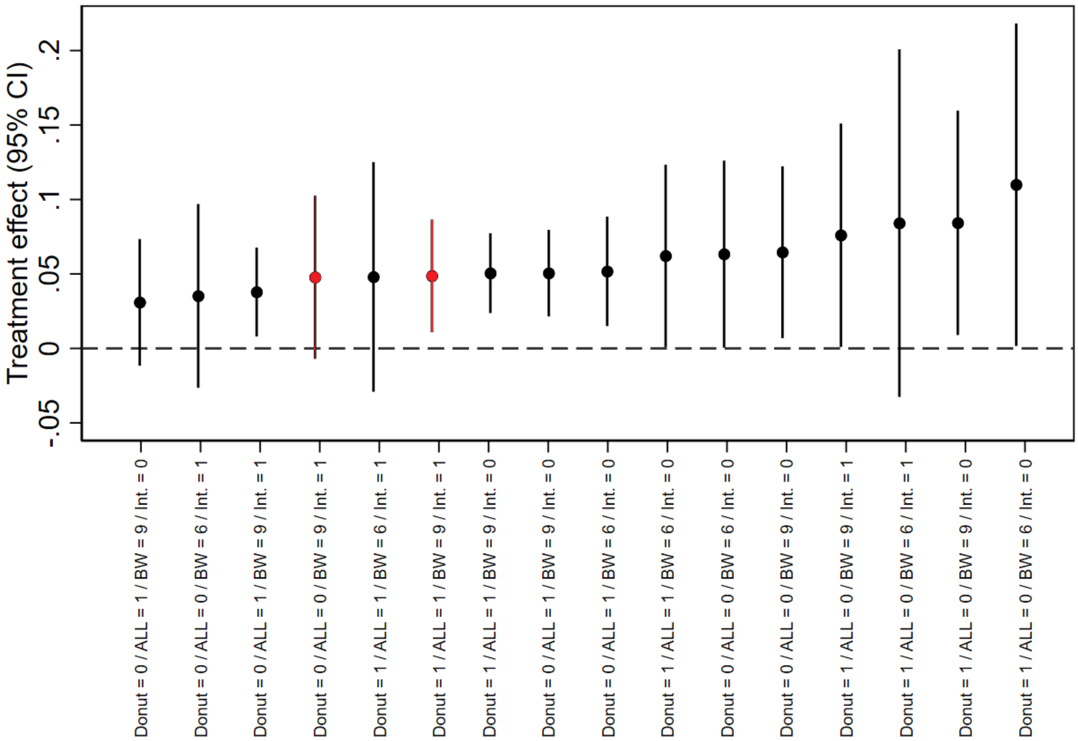
Bandwidth 180 days



Bandwidth 270 days

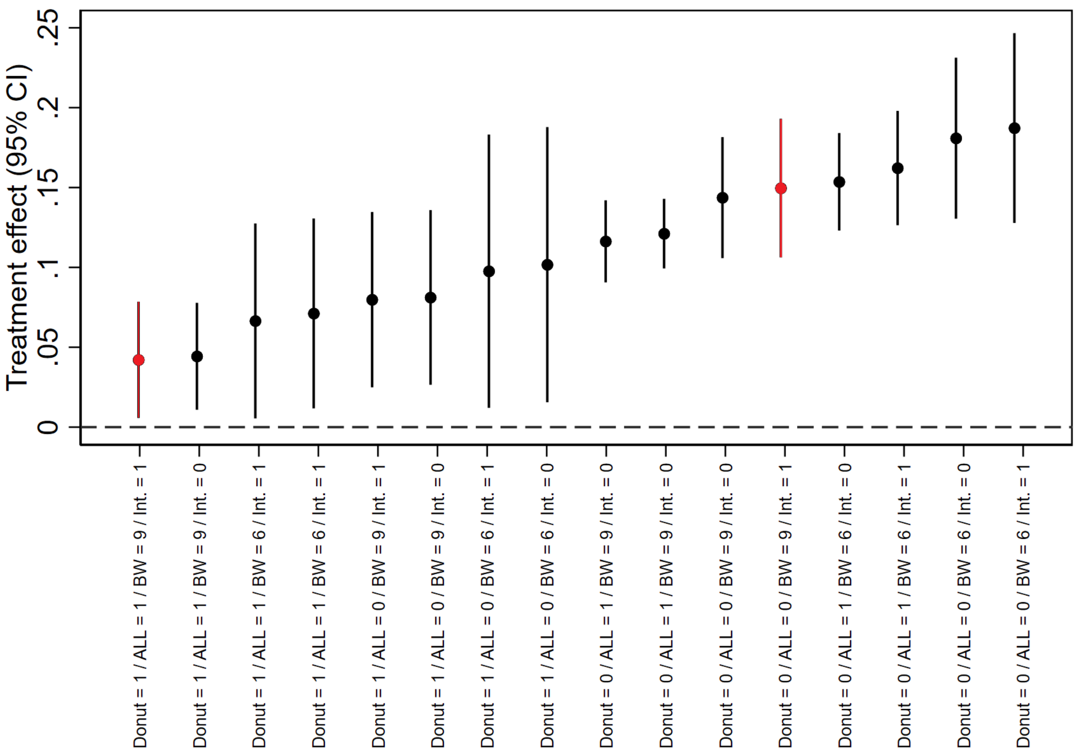
Notes: This figure shows placebo tests for specification 2 (black) and specification 3 (gray), shifting the placebo treatment in a 10-day interval between July 2016 and December 2019. For both bandwidths, a significant effect is visible at the subsidy and procedural change. Other, seemingly cyclical, amplitudes are visible at other points in time.

Figure A2.27: Specification chart for robustness of the treatment effect of the subsidy change



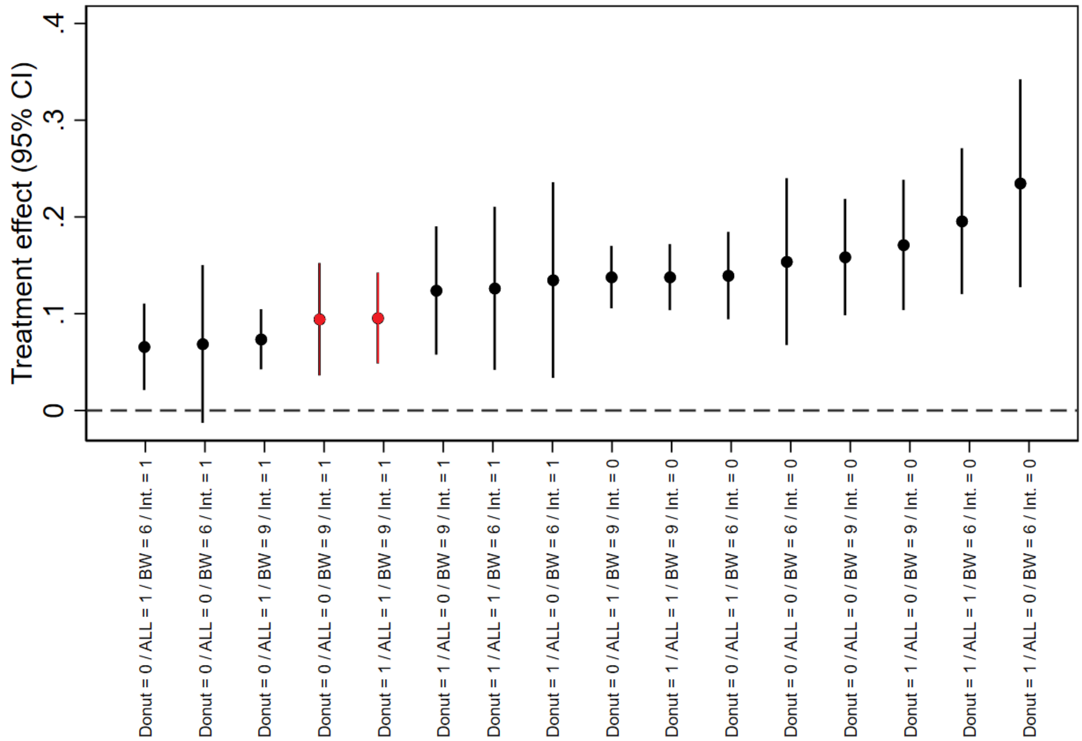
Notes: This figure shows how the treatment effect varies with the choice of specification along four dimensions: including the Donut design (leaving out observations in the two months interim period around the program change; Donut = 0/1), including the ALL approach (controlling for spatial effects in a two-stage apporach using information from the full sample period; ALL = 0/1), choosing a lower or upper bound bandwidth of 6 or nine months (BW = 6/9), and including an interaction term between treatment indicator and running variable (Int. = 0/1). Our two main RDD specifications are marked in red.

Figure A2.28: Specification chart for robustness of the treatment effect of the procedural change



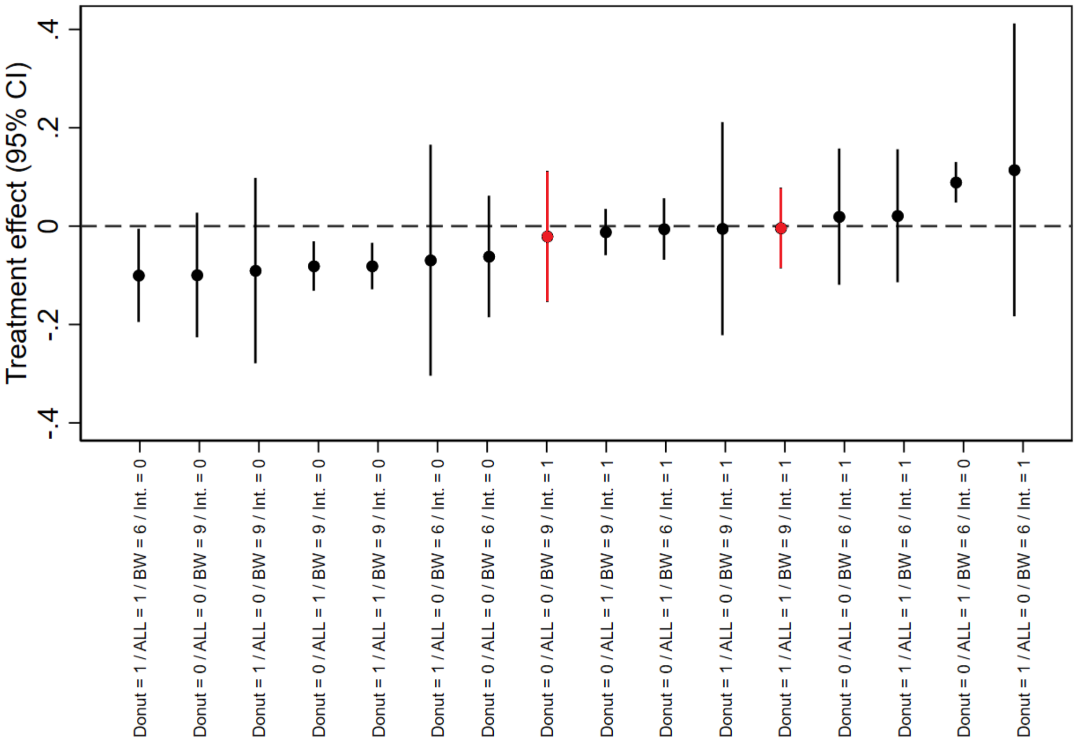
Notes: This figure shows how the treatment effect varies with the choice of specification along four dimensions: including the Donut design (leaving out observations in the two months interim period around the program change; Donut = 0/1), including the ALL approach (controlling for spatial effects in a two-stage apporach using information from the full sample period; ALL = 0/1), choosing a lower or upper bound bandwidth of 6 or nine months (BW = 6/9), and including an interaction term between treatment indicator and running variable (Int. = 0/1). Our two main RDD specifications are marked in red.

Figure A2.29: Specification chart for robustness of the treatment effect of the first stage of the subsidy change



Notes: This figure shows how the treatment effect varies with the choice of specification along four dimensions: including the Donut design (leaving out observations in the two months interim period around the program change; Donut = 0/1), including the ALL approach (controlling for spatial effects in a two-stage apporach using information from the full sample period; ALL = 0/1), choosing a lower or upper bound bandwidth of 6 or nine months (BW = 6/9), and including an interaction term between treatment indicator and running variable (Int. = 0/1). Our two main RDD specifications are marked in red.

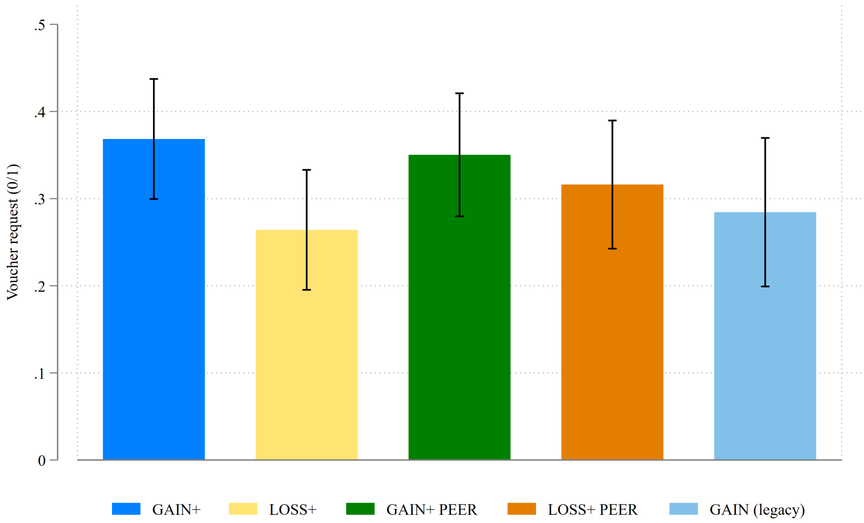
Figure A2.30: Specification chart for robustness of the treatment effect of the second stage of the subsidy change



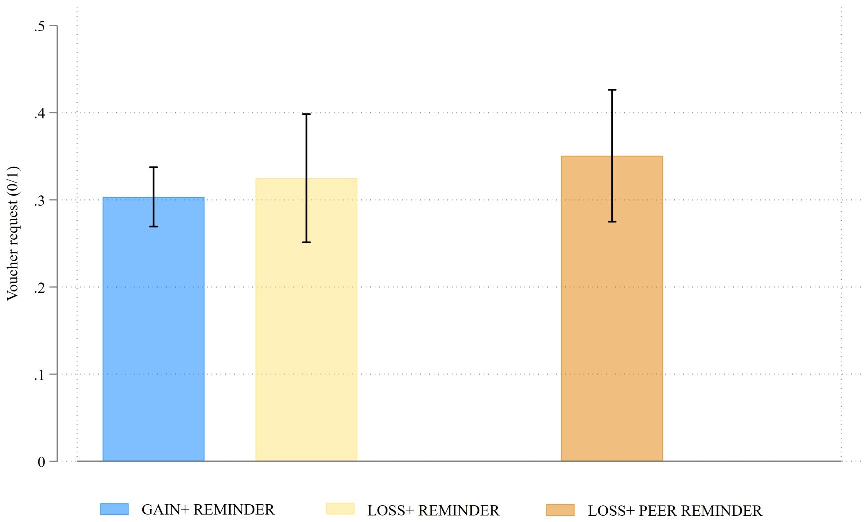
Notes: This figure shows how the treatment effect varies with the choice of specification along four dimensions: including the Donut design (leaving out observations in the two months interim period around the program change; Donut = 0/1), including the ALL approach (controlling for spatial effects in a two-stage apporach using information from the full sample period; ALL = 0/1), choosing a lower or upper bound bandwidth of 6 or nine months (BW = 6/9), and including an interaction term between treatment indicator and running variable (Int. = 0/1). Our two main RDD specifications are marked in red.

Chapter A3

Appendix for Chapter 3



(a) Information letter treatments.



(b) Reminder treatments.

Figure A3.1: Voucher request rates by treatment group.

Table A3.1: Treatment effects on refrigerator replacement conditional on having requested the voucher

	(1)	(2)	(3)	(4)
	Refrigerator replacement (0/1)			
$GAIN^+$	REF	REF	REF	REF
$LOSS^+$	-0.132	-0.167*	-0.144	-0.197*
	(0.095)	(0.094)	(0.108)	(0.111)
$GAIN^+$ PEER	-0.179*	-0.245**	-0.167	-0.229*
	(0.091)	(0.100)	(0.103)	(0.120)
$LOSS^+$ PEER	-0.257***	-0.247**	-0.256**	-0.228*
	(0.097)	(0.104)	(0.121)	(0.126)
GAIN (legacy)	0.043	0.009	-0.052	-0.071
	(0.100)	(0.098)	(0.117)	(0.115)
$GAIN^+$ REMINDER	-0.084	-0.108	-0.132*	-0.135*
	(0.066)	(0.069)	(0.073)	(0.076)
$LOSS^+$ REMINDER	-0.122	-0.136	-0.130	-0.128
	(0.090)	(0.092)	(0.092)	(0.097)
$LOSS^+$ PEER REMINDER	-0.154	-0.196*	-0.210*	-0.249**
	(0.094)	(0.100)	(0.110)	(0.117)
Constant	0.754***	0.423**	0.741***	0.460*
	(0.072)	(0.168)	(0.086)	(0.249)
Savings Info	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Fixed Effects	No	No	Yes	Yes
Request=1	Yes	Yes	Yes	Yes
N	571	555	541	525

Note: Linear probability models of refrigerator replacement (yes/no) on treatments. The treatments are included as indicators for the respective treatment group. The $GAIN^+$ treatment is the omitted reference treatment group. All regressions control for the communicated savings from replacement. Columns (2) and (4) add control variables for household's electricity price, number of persons in the household, past electricity consumption, living space, federal state, the social benefit transfer scheme, and whether the household heats warm water with electricity. Columns (3) and (4) include fixed effects for the respective intervention sites, the month of when the participant is informed about his/her replacement eligibility and a month-site interaction. Robust standard errors are in parenthesis. Significance levels: $^*: p < 0.10$, $^{**}: p < 0.05$, $^{***}: p < 0.01$.

Table A3.2: Treatment effects on voucher request: Allowing for differential reminder effects

	(1)	(2)	(3)	(4)
	Voucher request (0/1)			
$GAIN^+$	REF	REF	REF	REF
$LOSS^+$	-0.105**	-0.114**	-0.108**	-0.112**
	(0.050)	(0.049)	(0.047)	(0.047)
$GAIN^+$ PEER	0.031	0.056	0.000	-0.004
	(0.053)	(0.054)	(0.055)	(0.057)
$LOSS^+$ PEER	-0.004	0.032	0.003	0.014
	(0.054)	(0.055)	(0.053)	(0.053)
GAIN (legacy)	-0.098*	-0.085	-0.078	-0.072
	(0.056)	(0.055)	(0.054)	(0.053)
$GAIN^+$ Letter/SMS	-0.049	-0.042	-0.050	-0.036
	(0.043)	(0.043)	(0.041)	(0.041)
$GAIN^+$ Tag	-0.061	-0.059	-0.068	-0.052
	(0.049)	(0.049)	(0.046)	(0.046)
$GAIN^+$ Letter/SMS Tag	-0.099**	-0.100**	-0.103**	-0.109**
	(0.048)	(0.048)	(0.046)	(0.046)
$LOSS^+$ Tag	-0.044	-0.048	-0.044	-0.042
	(0.051)	(0.050)	(0.047)	(0.047)
$LOSS^+$ PEER Tag	0.031	0.071	-0.033	-0.034
	(0.055)	(0.056)	(0.055)	(0.056)
Constant	0.301***	0.540***	0.310***	0.419***
	(0.043)	(0.085)	(0.045)	(0.101)
Savings Info	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Fixed Effects	No	No	Yes	Yes
N	1802	1761	1785	1745

Note: Linear probability models of voucher request (yes/no) on treatments. The treatments are included as indicators for the respective treatment group. The $GAIN^+$ treatment is the omitted reference treatment group. All regressions control for the communicated savings from replacement. Columns (2) and (4) add control variables for household's electricity price, number of persons in the household, past electricity consumption, living space, federal state, the social benefit transfer scheme, and whether the household heats warm water with electricity. Columns (3) and (4) include fixed effects for the respective intervention sites, the month of when the participant is informed about his/her replacement eligibility and a month-site interaction. Robust standard errors are in parenthesis. Significance levels: $^{*}: p < 0.10$, $^{**}: p < 0.05$, $^{***}: p < 0.01$.

Table A3.3: Treatment effects on refrigerator replacement: Allowing for differential reminder effects

	(1)	(2)	(3)	(4)
	Refrigerator replacement (0/1)			
GAIN^{+}	REF	REF	REF	REF
LOSS^{+}	-0.104**	-0.117***	-0.102**	-0.113***
	(0.042)	(0.041)	(0.040)	(0.040)
GAIN^{+} PEER	-0.044	-0.045	-0.062	-0.075
	(0.046)	(0.046)	(0.046)	(0.046)
LOSS^{+} PEER	-0.088*	-0.064	-0.071	-0.056
	(0.045)	(0.046)	(0.046)	(0.046)
GAIN (legacy)	-0.051	-0.051	-0.046	-0.047
	(0.049)	(0.048)	(0.048)	(0.047)
GAIN^{+} Letter/SMS	-0.081**	-0.081**	-0.082**	-0.072**
	(0.037)	(0.037)	(0.036)	(0.036)
GAIN^{+} Tag	-0.030	-0.037	-0.041	-0.031
	(0.043)	(0.043)	(0.042)	(0.042)
GAIN^{+} Letter/SMS Tag	-0.088**	-0.093**	-0.095**	-0.104***
	(0.041)	(0.041)	(0.040)	(0.040)
LOSS^{+} Tag	-0.070	-0.078*	-0.081*	-0.084**
	(0.043)	(0.044)	(0.042)	(0.042)
LOSS^{+} PEER Tag	-0.035	-0.020	-0.090*	-0.097**
	(0.047)	(0.048)	(0.046)	(0.047)
Constant	0.234***	0.292***	0.233***	0.239**
	(0.037)	(0.072)	(0.038)	(0.095)
Savings Info	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes
Fixed Effects	No	No	Yes	Yes
N	1802	1761	1785	1745

Note: Linear probability models of refrigerator replacement (yes/no) on treatments. The treatments are included as indicators for the respective treatment group. The GAIN^{+} treatment is the omitted reference treatment group. All regressions control for the communicated savings from replacement. Columns (2) and (4) add control variables for household's electricity price, number of persons in the household, past electricity consumption, living space, federal state, the social benefit transfer scheme, and whether the household heats warm water with electricity. Columns (3) and (4) include fixed effects for the respective intervention sites, the month of when the participant is informed about his/her replacement eligibility and a month-site interaction. Robust standard errors are in parenthesis. Significance levels: $^{*}: p < 0.10$, $^{**}: p < 0.05$, $^{***}: p < 0.01$.

Table A3.4: Heterogeneous treatments effects on refrigerator replacement by transfer type

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Refrigerator replacement (0/1)							
$GAIN^+$	REF	REF	REF	REF	REF	REF	REF	REF
$LOSS^+$	-0.010	-0.028	-0.030	-0.035	-0.159***	-0.168***	-0.130***	-0.141***
	(0.079)	(0.078)	(0.081)	(0.081)	(0.047)	(0.047)	(0.046)	(0.047)
$GAIN^+$ PEER	-0.070	-0.024	-0.128	-0.088	-0.013	-0.057	0.024	-0.024
	(0.080)	(0.081)	(0.087)	(0.090)	(0.056)	(0.056)	(0.057)	(0.059)
$LOSS^+$ PEER	-0.010	0.042	0.019	0.062	-0.104**	-0.115**	-0.089*	-0.100*
	(0.088)	(0.088)	(0.093)	(0.097)	(0.052)	(0.053)	(0.054)	(0.054)
GAIN (legacy)	0.071	0.088	0.072	0.083	-0.141***	-0.139**	-0.108**	-0.112**
	(0.089)	(0.087)	(0.103)	(0.104)	(0.055)	(0.055)	(0.055)	(0.055)
$GAIN^+$ REMINDER	-0.014	-0.009	-0.033	-0.021	-0.100**	-0.104**	-0.083**	-0.083**
	(0.060)	(0.059)	(0.062)	(0.065)	(0.042)	(0.043)	(0.040)	(0.040)
$LOSS^+$ REMINDER	-0.063	-0.054	-0.108	-0.100	-0.076	-0.081	-0.064	-0.068
	(0.074)	(0.074)	(0.077)	(0.082)	(0.054)	(0.055)	(0.053)	(0.053)
$LOSS^+$ PEER REMINDER	0.025	0.051	0.012	0.017	-0.047	-0.056	-0.086	-0.104*
	(0.087)	(0.087)	(0.096)	(0.101)	(0.056)	(0.057)	(0.054)	(0.055)
Constant	0.236***	0.474***	0.245***	0.332**	0.207***	0.227***	0.185***	0.249*
	(0.071)	(0.132)	(0.079)	(0.152)	(0.045)	(0.086)	(0.047)	(0.127)
ALG II	No	No	No	No	Yes	Yes	Yes	Yes
Savings Info	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
N	638	620	610	591	1164	1141	1129	1105

Note: Linear probability models of refrigerator replacement (yes/no) on treatments. The treatments are included as indicators for the respective treatment group. The $GAIN^+$ treatment is the omitted reference treatment group. Columns (1)-(4) cover program participants who receive social transfer payments other than long-term unemployment benefits (ALG II=no). Columns (5)-(8) cover program participants receiving long-term unemployment benefits (ALGI II=yes). All regressions control for the communicated savings from replacement. Columns (2), (4), (6) and (8) add control variables for household's electricity price, number of persons in the household, past electricity consumption, living space, federal state and whether the household heats warm water with electricity. Columns (3), (4), (7) and (8) include fixed effects for the respective intervention sites, the month of when the participant is informed about his/her replacement eligibility and a month-site interaction. Robust standard errors are in parenthesis. Significance levels: * : $p < 0.10$, ** : $p < 0.05$, *** : $p < 0.01$.

Chapter A4

Appendix for Chapter 4

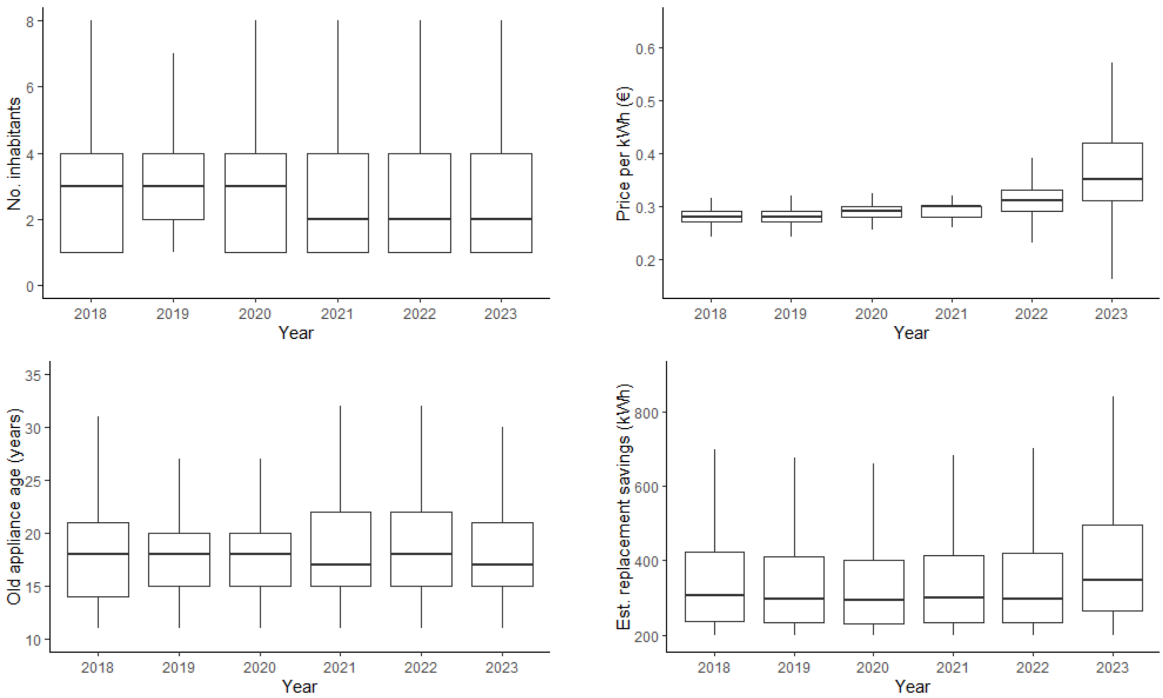


Figure A4.1: Household characteristics at the appliance level over years (wide dataset))
Notes: This figure shows the distribution of household characteristics over time by year.

Table A4.1: Relevance of energy advisors for household decision to request voucher

Dependent Variable:	Voucher request				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Inhabitants	0.0184***	0.0160***	0.0162***	0.0152***	0.0144***
	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0021)
Electricity price per kWh	-0.1920	-0.2304*	-0.2115*	-0.2280*	-0.2795*
	(0.1253)	(0.1280)	(0.1279)	(0.1351)	(0.1466)
Refrigerator age	0.0024***	0.0026***	0.0027***	0.0027***	0.0020***
	(0.0005)	(0.0004)	(0.0005)	(0.0004)	(0.0005)
Transfer: AsylbLG	0.0243	0.0420*	0.0364	0.0411*	0.0108
	(0.0224)	(0.0233)	(0.0243)	(0.0239)	(0.0250)
Transfer: None	0.0662***	0.0717***	0.0597***	0.0673***	0.0704***
	(0.0158)	(0.0151)	(0.0153)	(0.0148)	(0.0164)
Transfer: Child supplements	0.0318	0.0374	0.0510*	0.0527*	0.0546*
	(0.0294)	(0.0304)	(0.0293)	(0.0291)	(0.0320)
Transfer: Pension supplements	0.0390***	0.0432***	0.0431***	0.0439***	0.0430***
	(0.0110)	(0.0120)	(0.0120)	(0.0124)	(0.0139)
Transfer: Basic income	0.0335*	0.0408**	0.0356*	0.0410**	0.0410*
	(0.0193)	(0.0186)	(0.0198)	(0.0195)	(0.0217)
Transfer: Housing allowance	0.0798***	0.0856***	0.0839***	0.0875***	0.0828***
	(0.0169)	(0.0155)	(0.0182)	(0.0166)	(0.0180)
Est. savings from replacement	$8.19 \times 10^{-5***}$	$7.62 \times 10^{-5***}$	$7.25 \times 10^{-5***}$	$6.73 \times 10^{-5***}$	$7.74 \times 10^{-5***}$
	(1.87×10^{-5})	(1.93×10^{-5})	(2.01×10^{-5})	(2.01×10^{-5})	(2.32×10^{-5})
Fixed-effects					
ZIP code	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes
Branch-year	Yes	Yes	Yes	Yes	Yes
Advisor		Yes		Yes	Yes
Co-advisor			Yes	Yes	Yes
Combination of advisors					Yes
Fit statistics					
Observations	48,686	48,686	48,686	48,686	48,686
R^2	0.25666	0.30787	0.31850	0.36006	0.43831

Notes: Standard errors in parentheses are clustered by advisor and co-advisor. p-values ***: 0.01, **: 0.05, *: 0.1.

Table A4.2: Determinants: Effect of advisor characteristics on household investment decisions

Dependent Variable:	Voucher request	Refrigerator replacement
Model:	(1)	(2)
Program controls		
Position: Ref. cat. Management staff		
Position: Other	0.2318	0.1692
	(0.3273)	(0.2754)
Position: Short-term contract	0.3979	0.2003
	(0.3032)	(0.2562)
Position: Long-term contract	0.8317**	0.2202
	(0.4129)	(0.2672)
Joined program: Ref. cat. 2008-2012		
Joined program: 2013-2015	0.3111***	0.2496**
	(0.1110)	(0.0941)
Joined program: 2016-2018	0.3134**	0.2806**
	(0.1293)	(0.1223)
Joined program: 2019-2022	0.1681	0.2180
	(0.1849)	(0.1553)
Socio-demographic characteristics		
Gender: Female	0.0483	0.0401
	(0.0564)	(0.0510)
Age: Ref. cat. < 40 years old		
Age: 40-59 years old	0.1339**	0.0569
	(0.0635)	(0.0441)
Age: ≥ 60 years old	0.1813	0.2188
	(0.1771)	(0.1402)
Education: Ref. cat. Vocational training		
Education: No high school degree	0.0339	0.0474
	(0.0616)	(0.0614)
Education: High school degree	-0.0767	-0.1361**
	(0.0917)	(0.0680)
In relationship	0.0907	-0.0095
	(0.0851)	(0.0658)
Children	0.3069**	0.1666
	(0.1412)	(0.1099)
Native language: German	0.1370*	0.0016
	(0.0716)	(0.0504)
Economic preferences		
Altruism	0.0002	-0.0325**
	(0.0156)	(0.0123)
Risk	0.0053	0.0160
	(0.0168)	(0.0152)
Patience	0.0369**	0.0513***
	(0.0169)	(0.0111)
Own investment by advisor		
Ref. cat. Replacement ≤ 10 years ago		
Replacement > 10 years ago	-0.3821***	-0.2594***
	(0.1131)	(0.0839)
Replacement timing not reported	-0.0224	0.3008
	(0.2148)	(0.2017)
Attitudes towards replacement program		
Program is a good idea	-0.0438	-0.0272
	(0.0548)	(0.0469)
Right households profit	0.0100	-0.0177
	(0.0557)	(0.0395)
Households save energy	-0.0421	-0.0022
	(0.0756)	(0.0599)
Households protect climate	0.0985**	0.1432***
	(0.0484)	(0.0238)
Voucher worth the effort	-0.0661	-0.0967**
	(0.0526)	(0.0442)
Peer score	0.0668**	0.0630**
	(0.0324)	(0.0239)
Household controls	Yes	Yes
Fixed effects	Yes	Yes
Observations	5,747	5,747
R^2	0.43006	0.42028

Notes: Standard errors in parentheses are clustered by advisor. Household characteristics include the number of inhabitants, electricity price per kWh, old refrigerator age, estimated savings from refrigerator replacement and the governmental transfer type received. Fixed effects account for ZIP code, year-month, branch-year and co-advisor. p-values ***: 0.01, **: 0.05, *: 0.1.

Chapter A5

Appendix for Chapter 5

A5.1 Tables

Table A5.1: Relationship between annual energy consumption and purchase price

Dependent Variable:	Purchase price in Euro				
Model:	(1)	(2)	(3)	(4)	(5)
Annual energy consumption (kWh)	-0.5114	-6.528***	-1.207**	-1.126	0.0249
	(0.4387)	(0.7618)	(0.5490)	(1.828)	(1.157)
Estimator	OLS	OLS	OLS	OLS	OLS
Brand	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Type	Yes				
Volume	Yes				
Observations	1,436	103	94	106	75

Notes: The data is aggregated by model at the mean purchase price. Column (1) uses data from all models. Columns (2), (3), (4) and (5) use observations of models located in the volume brackets of 300-304, 320-324, 335-339 and 340-344 liters, being the brackets with the highest density of observations. Standard errors are clustered at the brand level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A5.2: Smoothness of covariates around the federal subsidy reduction

Dependent Variable:	**Total energy consumption**		
Model:	(1)	(2)	(3)
Subsidy +50€	101.8	106.3	110.6
	(182.4)	(186.5)	(184.2)
Day count		Yes	Yes
Rescale × Day count			Yes
Fixed effects	Yes	Yes	Yes
Observations	10,179	10,179	10,179

Dependent Variable:	**No. persons in household**		
Model:	(1)	(2)	(3)
Subsidy +50€	0.1418	0.1434	0.1526
	(0.1394)	(0.1416)	(0.1432)
Day count		Yes	Yes
Rescale × Day count			Yes
Fixed effects	Yes	Yes	Yes
Observations	10,179	10,179	10,179

Dependent Variable:	**Age of old appliance**		
Model:	(1)	(2)	(3)
Subsidy +50€	-0.1552	-0.0985	-0.0985
	(0.3784)	(0.3825)	(0.3847)
Day count		Yes	Yes
Rescale × Day count			Yes
Fixed effects	Yes	Yes	Yes
Observations	10,179	10,179	10,179

Dependent Variable:	**Est. savings from replacement**		
Model:	(1)	(2)	(3)
Subsidy +50€	0.1361	0.3596	0.1497
	(12.14)	(12.46)	(12.36)
Day count		Yes	Yes
Rescale × Day count			Yes
Fixed effects	Yes	Yes	Yes
Observations	10,179	10,179	10,179

Notes: Fixed effects include model, region-year, month, volume-region and manufacturing brand-region and -year fixed effects. Standard errors are clustered at the region level. All regressions use the sample of households located in a bandwidth of two years around the reduction of the federal subsidy by 50 Euro on 1 April 2019. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A5.3: Balance of household characteristics around the federal subsidy reduction

	Pre-treatment		Post-treatment			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
No. persons	3.0	1.8	2.9	1.8	-0.1*	0.0
Age of old appliance	18.4	4.9	18.4	4.9	0.0	0.1
Est. savings from replacement	340.5	157.5	349.9	165.0	9.4***	3.4
Total energy consumption	2853.3	1683.9	2901.5	1686.3	48.1	35.5

Notes: The table shows the balance of household-level covariates in a period of two years before and after the subsidy reduction on 1 April 2019. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

Table A5.4: Robustness check results (Staggered adoption) for the demand response estimation: Financial constraints

Dependent Variable:	N models sold	
Model:	(1)	(2)
Annual energy cost	-0.0013	-0.0027
	(0.0062)	(0.0097)
Net purchase price	0.0002	-0.0001
	(0.0005)	(0.0005)
Annual EC × treat	0.0218**	0.0305***
	(0.0101)	(0.0117)
Net price × treat	-0.0034***	-0.0033***
	(0.0006)	(0.0007)
Estimator	GLM Poisson	GLM Poisson
Fixed effects	Yes	Yes
Observations	1,068,817	556,557
Change in trade-off ratio:	-0.15	-6.70
$\frac{\alpha_1+\alpha_2}{(\theta_1+\theta_2)\rho(\frac{1-\rho^L}{1-\rho})} - \frac{\alpha_1}{\theta_1\rho(\frac{1-\rho^L}{1-\rho})}$	(2.78)	(1.19)
p-value H0: $m = 0$	0.508	0.616

Notes: For column (1), observations in a bandwidth of two years around each subsidy implementation are considered. For column (2), the bandwidth is one year. Fixed effects include model, region, quarter-year, volume-state, manufacturing brand-state and manufacturing brand-year fixed effects. In column (1) 6,002 observations were dropped due to only zero outcomes in fixed effects cells, in column (2) 31,120 observations, respectively. Standard errors are clustered at the model level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A5.5: Robustness check results (Difference-in-Difference) for the demand response estimation: Financial constraints

Dependent Variable:	N models sold		
Model:	(1)	(2)	(3)
Annual energy cost	-0.0116	0.0474***	0.0247**
	(0.0107)	(0.0143)	(0.0110)
Net purchase price	0.0011	-0.0017	-0.0018***
	(0.0008)	(0.0012)	(0.0007)
Annual EC × treat	0.0217	-0.0846**	-0.0069
	(0.0174)	(0.0417)	(0.0670)
Net price × treat	-0.0027***	0.0045*	0.0073
	(0.0011)	(0.0025)	(0.0057)
Annual EC × post	0.0277***	-0.0651***	-0.0217**
	(0.0084)	(0.0160)	(0.0088)
Net price × post	-0.0006	0.0002	0.0012*
	(0.0005)	(0.0015)	(0.0007)
Annual EC × treat × post	-0.0115	0.1377***	-0.5721***
	(0.0159)	(0.0459)	(0.0512)
Net price × treat × post	-0.0008	-0.0048*	0.3661***
	(0.0009)	(0.0028)	(0.0238)
Estimator	GLM Poisson	GLM Poisson	GLM Poisson
Fixed effects	Yes	Yes	Yes
Observations	732,012	302,739	412,537
Change in trade-off ratio:	-10.9	-10.6	-3.31
$\frac{\alpha_1+\alpha_2}{(\theta_1+\theta_2)\rho(\frac{1-\rho^L}{1-\rho})} - \frac{\alpha_1}{\theta_1\rho(\frac{1-\rho^L}{1-\rho})}$	(12.29)	(6.53)	(3.99)
p-value H0: $m = 0$	0.813	0.508	0.797

Notes: Column (1) presents results for the subsidy implementation in North Rhine-Westphalia, column (2) for Berlin and column (3) for Saxony-Anhalt. The sample used for column (1) is the sample Q3 2014 to Q2 2018, for column (2) from Q1 2019 to Q4 2022 excluding observations from Q4 2020 and observations from Saxony-Anhalt and Saarland, and for column(3) from Q2 2018 to Q1 2022, excluding Q2 2020 and excluding observations from Berlin and Saarland. Fixed effects include model, region, quarter-year, volume-state, manufacturing brand-state and manufacturing brand-year fixed effects. In column (1) 4,633 observations were dropped due to only zero outcomes in fixed effects cells, in column (2) 20,071 observations, and in column (3) 28,899 observations, respectively. Standard errors are clustered at the model level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A5.6: Smoothness of covariates around the revision of the EU Energy Label

Dependent Variable:		**Total energy consumption**	
Model:	(1)	(2)	(3)
Revision	40.42	39.04	-48.43
	(229.5)	(230.5)	(271.6)
Day count		Yes	Yes
Rescale × Day count			Yes
Fixed effects	Yes	Yes	Yes
Observations	5,599	5,599	5,599
5,599			

Dependent Variable:		**No. persons in household**	
Model:	(1)	(2)	(3)
Revision	0.1658	0.1649	0.1499
	(0.2336)	(0.2341)	(0.2604)
Day count		Yes	Yes
Rescale × Day count			Yes
Fixed effects	Yes	Yes	Yes
Observations	5,599	5,599	5,599
5,599			

Dependent Variable:		**Age of old appliance**	
Model:	(1)	(2)	(3)
Revision	-0.3680	-0.4001	-0.3814
	(0.5728)	(0.5695)	(0.6308)
Day count		Yes	Yes
Rescale × Day count			Yes
Fixed effects	Yes	Yes	Yes
Observations	5,599	5,599	5,599
5,599			

Dependent Variable:		**Est. savings from replacement**	
Model:	(1)	(2)	(3)
Revision	-29.46	-30.01	-10.31
	(26.87)	(27.08)	(25.82)
Day count		Yes	Yes
Rescale × Day count			Yes
Fixed effects	Yes	Yes	Yes
Observations	5,599	5,599	5,599
5,599			

Notes: Fixed effects include model, region-year, month, volume-region and manufacturing brand-region and -year fixed effects. Standard errors are clustered at the region level. All regressions use the sample of households located in a bandwidth of two years around the revision of the EU Energy Label on 1 March 2021. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A5.7: Balance of household characteristics around the revision of the EU Energy Label

	Pre-treatment		Post-treatment			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
No. persons	3.1	1.9	3.0	1.8	-0.1	0.1
Age of old appliance	18.9	5.5	18.4	4.9	-0.4***	0.1
Est. savings from replacement	359.1	169.7	340.5	158.6	-18.6***	4.5
Total energy consumption	2968.4	1777.4	2854.1	1690.7	-114.3**	47.2

Notes: The table shows the balance of household-level covariates in a period of two years before and after the revision of the EU Energy Label on 1 March 2021. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

A5.2 Figures

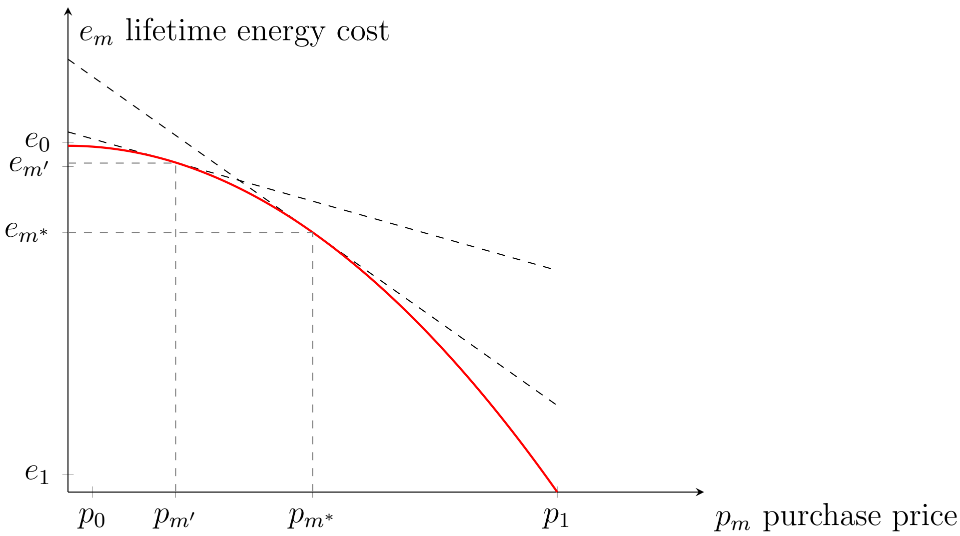


Figure A5.1: Efficiency level m' chosen under constrained budget
Notes: This figure illustrates the choice of efficiency level m under a constrained budget. The purchase price p_m on the x axis increases in m from 0 to 1. The lifetime energy cost on the y axis decreases in m from 1 to 0. The red curve plots the distribution of appliance models $m \in [0, 1]$ in the choice set. The optimal efficiency threshold m^* is at (p^*_m, e^*_m). The efficiency level m' chosen under constrained budget is below m^* at $(p_{m'}, e_{m'})$, the slope of the appliance supply curve at that point being greater than -1.

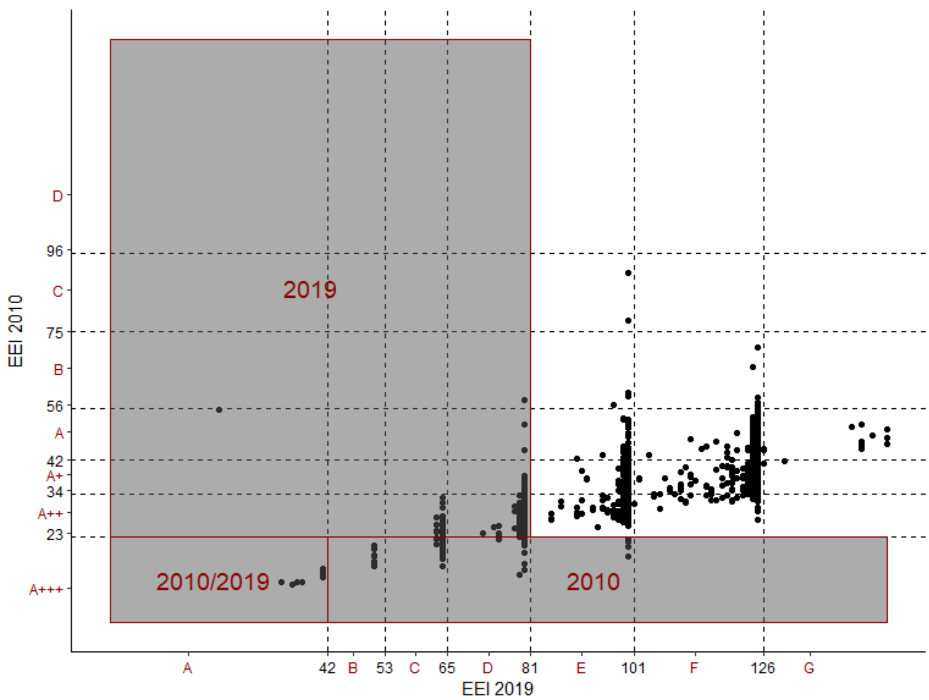


Figure A5.2: Relationship between 2010 and 2019 EEI for EPREL 2019 sample
Notes: This figure shows how the EEI under the 2016/2019 Regulation maps into the EEI under the 1060/2010 Regulation for the sample of models registered under the 2016/2019 Regulation. Observations above a 2010 and 2019 EEI of 150 are respectively removed.

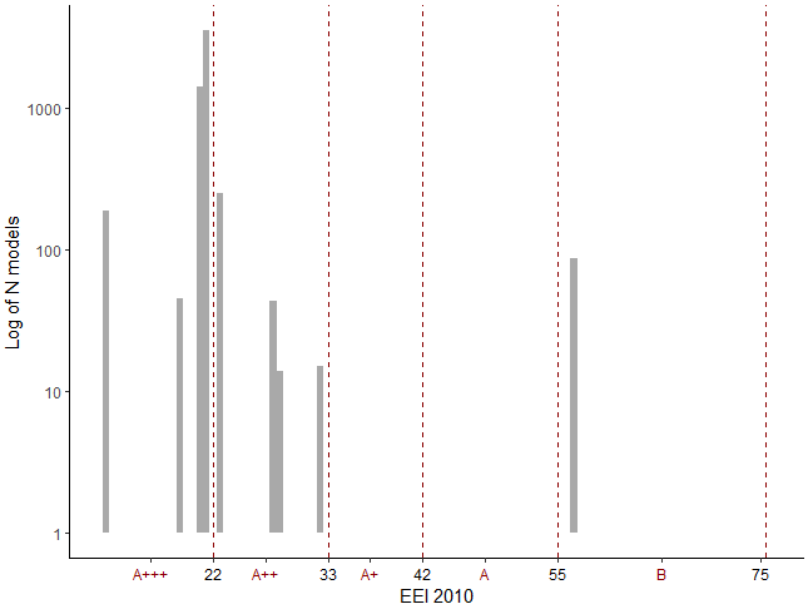


Figure A5.3: Density distribution of household investment by 2010 EEI
Notes: This figure shows the logarithmized distribution of appliance purchases by EEI that were matched to EPREL 2010 appliance characteristics. The red dashed lines mark the EEI thresholds of each energy class of the EU Energy Label. Apparently, a few households choose a model in energy classes A++ and B which are below the SSC program requirements.

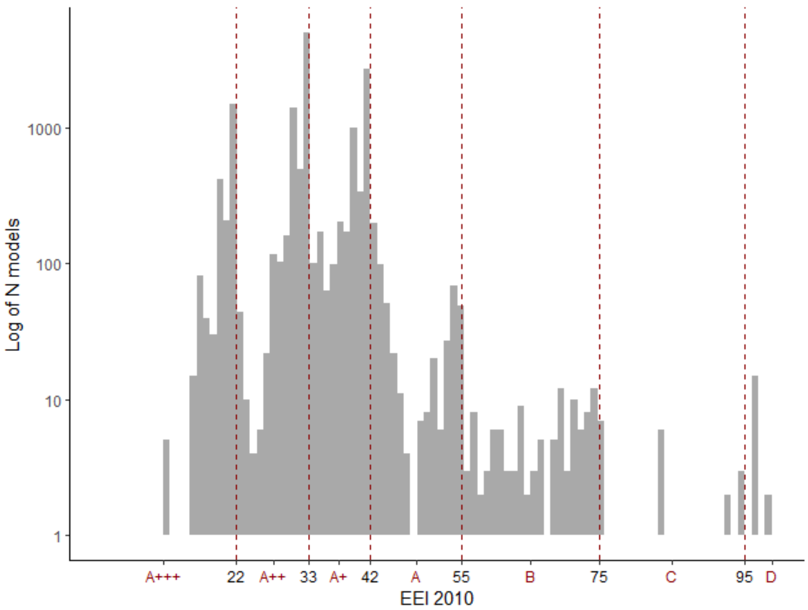


Figure A5.4: Density distribution of models on the market by 2010 EEI
Notes: This figure shows the logarithmized distribution of the models supplied on the market by Energy Efficiency Index that were registered under the 2016/2019 Regulation. The red dashed lines mark the thresholds of each energy class of the EU Energy Label.

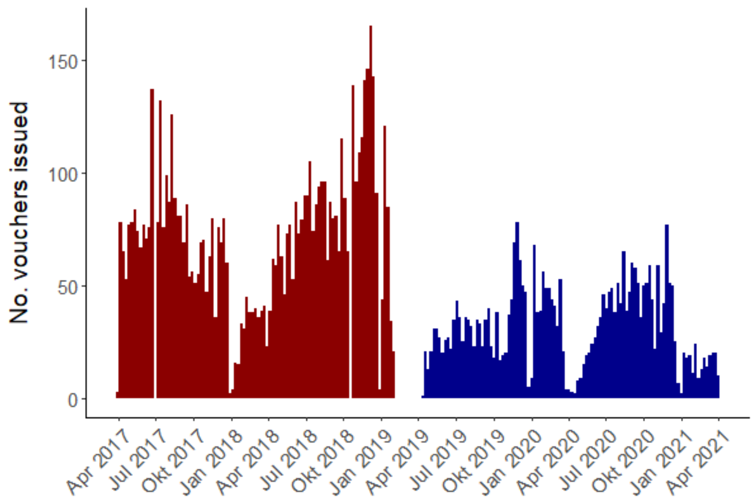


Figure A5.5: Density of voucher issuances, April 2017 to March 2021

Notes: This figure shows the density of voucher issuances in a bandwidth of two years around the reduction of the federal subsidy by 50 Euro on 1 April 2019 by week. Voucher issuances before the reduction are shown in red, issuances after the reduction in blue. Seasonal variation in the density of voucher issuances is strong. Issuances are more frequent towards the end of the year. The gaps in each December mark the end-of-year breaks. The gap in February and March 2019 marks the end of the funding period 2017 to 2019. The new funding period started in April 2019. The strong drop in April 2020 represents the first Covid lockdown and a slow restart of activities after.

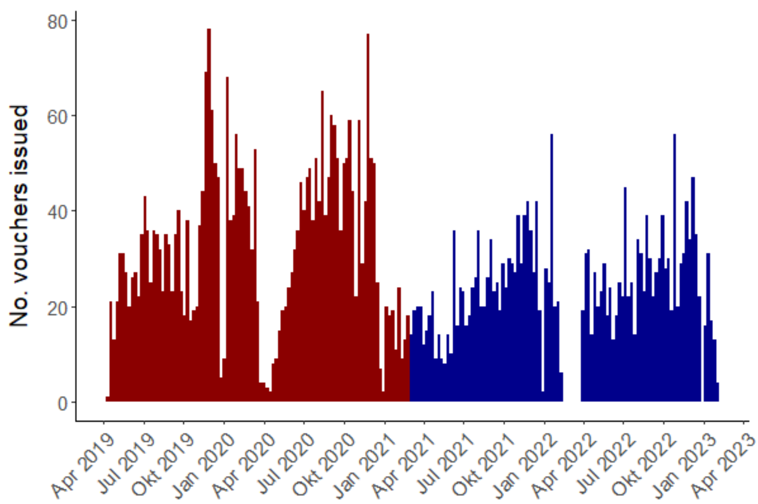


Figure A5.6: Density of voucher issuances, April 2019 to March 2023

Notes: This figure shows the density of voucher issuances in a bandwidth of two years around the revision of the EU Energy Label on March 1, 2021 by week. Voucher issuances before the revision are shown in red, issuances after the revision in blue. Seasonal variation in the density of voucher issuances is strong. Issuances are more frequent towards the end of the year. The strong drop in April 2020 represents the first Covid lockdown. The gaps in each December mark the end-of-year breaks. The gap in February and March 2022 marks the end of the funding period 2019 to 2022. The new funding period started in April 2022.

Chapter A6

Appendix for Chapter 6

A6.1 Background on the E-PRTR

Reporting procedures to the E-PRTR are set in the E-PRTR Regulation (European Parliament and European Council, 2006). Facilities located within the EU that undertake any of the activities specified for reporting must report the amounts of all pollutants that are higher than the capacity and pollutant-specific thresholds to its competent authority, i.e., the national authorities. The national authorities report them to the European Commission.

The stakes of reporting accuracy may vary spatially, as strict enforcement of accuracy, completeness, consistency and credibility of the reported data is the responsibility of national authorities. However, national authorities are liable to the European Commission in following its enforcement rules and can be penalized by the Commission in the case of infringement. Non-compliance of facilities in reporting to the national authorities is penalized via the national justice systems. There are no incentives for purposefully inaccurate and incorrect reporting as no direct consequences follow from pollutant reports, the European Commission collects the data for informational purpose. Since pollutant reports are made at the facility level, it is unlikely that a change in ownership would directly affect reporting behavior.

A6.2 Additional tables

Table A6.1: Predictive power of firm characteristics for treatment timing

	(1)	(2)
Total emissions	0.0001	0.0001
	(0.0001)	(0.0001)
Emissions intensity	0.0213	0.0102
	(0.0383)	(0.0283)
Output	0.0598	-1.94e-12
	(1.25e-11)	(1.32e-11)
Capital input	5.54e-11	4.06e-11
	(4.59e-11)	(3.73e-11)
Labor expenditures	-5.15e-8	1.32e-5
	(2.12e-5)	(1.95e-5)
Labor input	-1.56e-10	-2.67e-10
	(3.58e-10)	(3.28e-10)
Operating profits	2.2e-11	1.7e-11
	(3.53e-11)	(5.76e-11)
Intangible fixed assets	2.15e-10*	-4.18e-10***
	(1.27e-10)	(9.25e-11)
Adjusted R^2	0.00748	0.19498
Observations	3,946	3,946
Country-Year fixed effects		✓
Sector-Year fixed effects		✓

Notes: The outcome variable in both columns is the year of ownership change of firms as continuous variable. Standard errors clustered at the firm level are in parentheses. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A6.2: Event study estimates for total emissions at the facility level

	Main (1)	Imp (0) (2)	Imp (threshold) (3)
time_to_treat = -4	0.1143	0.1510	0.1260
	(0.0851)	(0.0931)	(0.0817)
time_to_treat = -3	0.0469	0.0149	0.0225
	(0.0563)	(0.0635)	(0.0546)
time_to_treat = -2	0.0598	0.0209	0.0197
	(0.0371)	(0.0414)	(0.0358)
time_to_treat = 0	-0.1927***	-0.1705***	-0.1937***
	(0.0350)	(0.0395)	(0.0341)
time_to_treat = 1	-0.3256***	-0.2831***	-0.3244***
	(0.0507)	(0.0549)	(0.0496)
time_to_treat = 2	-0.4134***	-0.3793***	-0.4107***
	(0.0632)	(0.0671)	(0.0610)
time_to_treat = 3	-0.6856***	-0.6322***	-0.6775***
	(0.0916)	(0.0954)	(0.0880)
Adjusted R^2	0.77745	0.74034	0.76831
Observations	10,624	11,479	11,479
Country-Year fixed effects	✓	✓	✓
Sector-Year fixed effects	✓	✓	✓
FacilityID fixed effects	✓	✓	✓

Notes: The first column shows the point estimates and standard errors of the main specification. The second and third column present the results using the data set in which gaps in emission reports are imputed with zero and the pollutant specific threshold, respectively. The fourth column shows results on total emissions scaled by CO2 emissions for the data set without imputation. All results refer to the Sun and Abraham (2021) specification. Standard errors clustered at the facility level are in parentheses. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

Table A6.3: Event study estimates for total emissions at the firm level

	Main (1)	Imp (0) (2)	Imp (threshold) (3)	Intensity (4)
time_to_treat = -4	0.4570***	0.4282***	0.4475***	0.0026
	(0.1360)	(0.1357)	(0.1294)	(0.0312)
time_to_treat = -3	0.2696***	0.1841*	0.2318***	0.0162
	(0.0970)	(0.0961)	(0.0897)	(0.0175)
time_to_treat = -2	0.1863***	0.1254*	0.1505**	-0.0054
	(0.0642)	(0.0652)	(0.0599)	(0.0124)
time_to_treat = 0	-0.3389***	-0.3683***	-0.3518***	-0.0201**
	(0.0509)	(0.0530)	(0.0484)	(0.0090)
time_to_treat = 1	-0.4028***	-0.4244***	-0.4373***	-0.0186
	(0.0690)	(0.0697)	(0.0661)	(0.0148)
time_to_treat = 2	-0.4638***	-0.4282***	-0.4504***	-0.0226
	(0.0924)	(0.0919)	(0.0881)	(0.0215)
time_to_treat = 3	-0.7854***	-0.7389***	-0.7700***	-0.0435
	(0.1373)	(0.1317)	(0.1306)	(0.0372)
Adjusted R^2	0.71605	0.69154	0.71246	0.41212
Observations	6,272	6,737	6,737	5,175
Country-Year fixed effects	✓	✓	✓	✓
Sector-Year fixed effects	✓	✓	✓	✓
BVDID fixed effects	✓	✓	✓	✓

Notes: The first column shows the point estimates and standard errors of the main specification. The second and third column present the results using the data set in which gaps in emission reports are imputed with zero and the pollutant specific threshold, respectively. All results refer to the Sun and Abraham (2021) specification. Standard errors clustered at the firm level are in parentheses. $^{***}p <$ 1%, $^{**}p <$ 5%, $^{*}p <$ 10%.

Table A6.4: Event study estimates for total emissions at the parent company level

	Main (1)	Imp (0) (2)	Imp (threshold) (3)	Intensity (4)
time_to_treat = -4	0.0396	0.0396	0.0753	0.0214
	(0.3364)	(0.3364)	(0.3312)	(0.0335)
time_to_treat = -3	0.0649	0.0649	0.0859	-0.0028
	(0.2217)	(0.2217)	(0.2212)	(0.0172)
time_to_treat = -2	0.2039	0.2039	0.1870	0.0155
	(0.1334)	(0.1334)	(0.1305)	(0.0153)
time_to_treat = 0	0.9832***	0.9832***	0.9847***	0.0398
	(0.1670)	(0.1670)	(0.1629)	(0.0372)
time_to_treat = 1	0.8008***	0.8008***	0.8426***	0.0368
	(0.1989)	(0.1989)	(0.1944)	(0.0399)
time_to_treat = 2	0.7257***	0.7257***	0.7548***	0.0335
	(0.2499)	(0.2499)	(0.2432)	(0.0405)
time_to_treat = 3	0.4986	0.4986	0.5113	-0.0128
	(0.3463)	(0.3463)	(0.3362)	(0.0196)
Adjusted R^2	0.74651	0.74651	0.75650	0.45075
Observations	2,274	2,274	2,274	1,858
Year fixed effects	✓	✓	✓	✓
ParentCompany fixed effects	✓	✓	✓	✓

Notes: The first column shows the point estimates and standard errors of the main specification. The second and third column present the results using the data set in which gaps in emission reports are imputed with zero and the pollutant specific threshold, respectively. Note that the first and second columns are the same since gaps in individual facilities or firms do not contribute to overall emissions of the parent company as in the case in which gaps are imputed by zero. All results refer to the Sun and Abraham (2021) specification. Standard errors clustered at the firm level are in parentheses. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

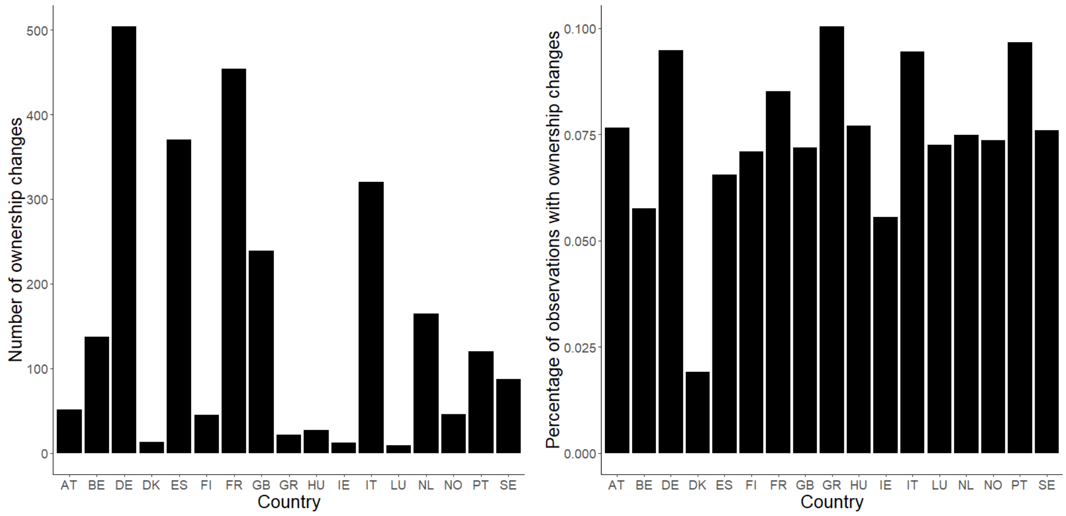
Table A6.5: Aggregate effects on firms: 50% largest firms

Dependent variables	**ATT**	**SE**	**N**
Total emissions	-0.284	(0.182)	1,654
Output	-1.754**	(0.618)	1,656
Emissions intensity	-0.010	(0.020)	1,542
Total factor productivity	0.015*	(0.009)	990
Operating profits	-0.394	(0.713)	1,312
Labor input	-0.627**	(0.236)	1,555
Capital input	-1.450*	(0.610)	1,646
Labor expenditures	-0.853	(0.583)	1,532
Intangible fixed assets	0.097	(0.094)	1,562

Notes: The first column denotes the respective dependent variables each with an inverse hyperbolic sine transformation. Each line represents a separate event study regression on the sample of 50 percent largest firms according to operating revenues in 2007. For each regression, we report the point estimate of the aggregated effect of the event study following Sun and Abraham (2021) (ATT), the standard error (SE) and the number of observations (N). Standard errors clustered at the firm level are in parentheses. Output refers to deflated operating revenues, emissions intensity to total emissions divided by output, labor input to number of employees, capital input to deflated tangible fixed assets, respectively. $^{***}p < 1\%$, $^{**}p < 5\%$, $^{*}p < 10\%$.

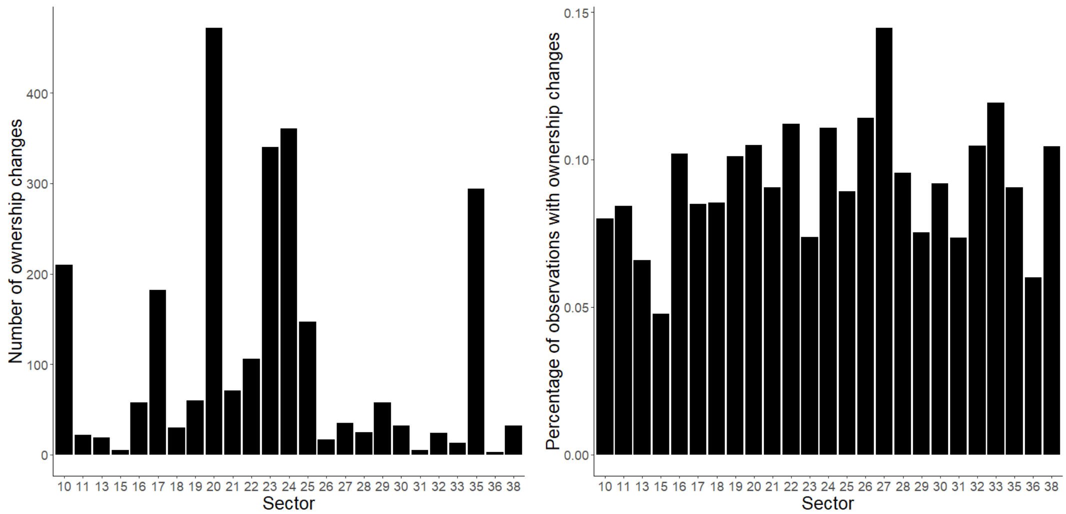
A6.3 Additional figures

Figure A6.1: Distribution of ownership changes over countries



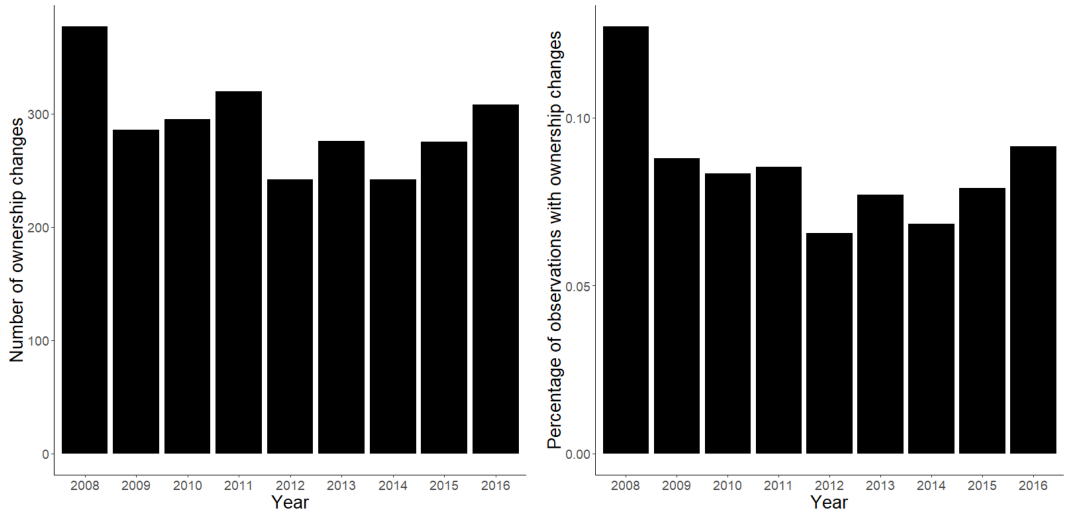
Notes: This figure shows the distribution of ownership changes over countries. The left panel shows the absolute number of changes and the right panel shows the relative share of changes out of all observations for the respective country.

Figure A6.2: Distribution of ownership changes over sectors



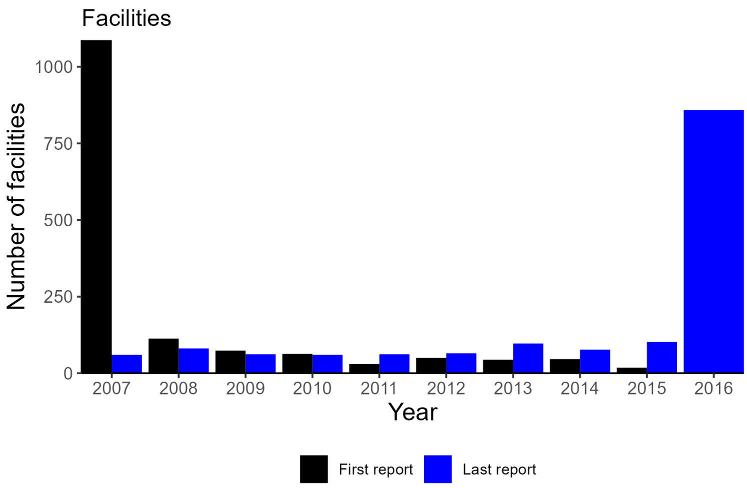
Notes: This figure shows the distribution of ownership changes over the sectors. The left panel shows the absolute number of changes and the right panel shows the relative share of changes out of all observations for the respective sector.

Figure A6.3: Distribution of ownership changes over years



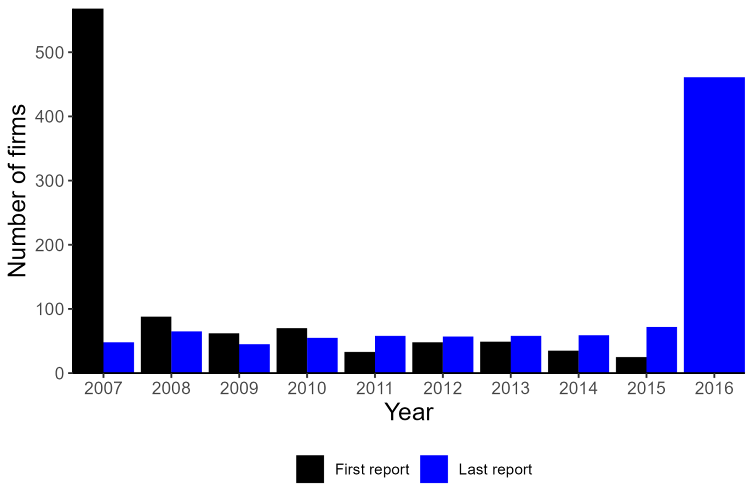
Notes: This figure shows the distribution of ownership changes over the years. The left panel shows the absolute number of changes and the right panel shows the relative share of changes out of all observations for the respective year.

Figure A6.4: First and last reporting year of facilities



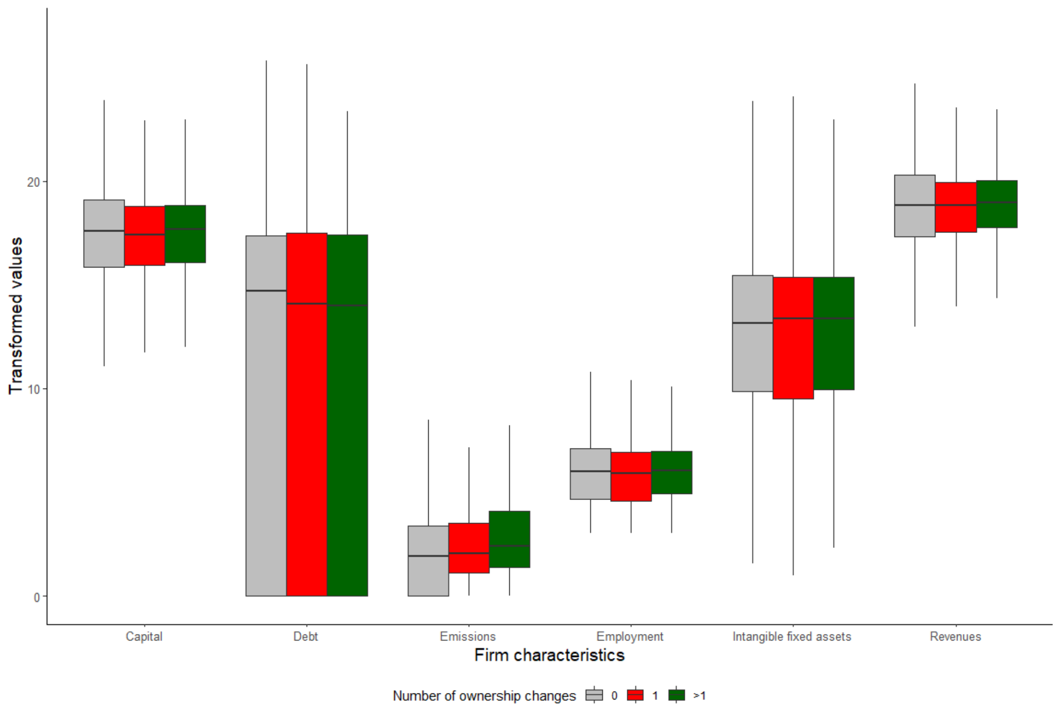
Notes: This figure shows the number of facilities that have their first reporting year and last reporting year, respectively, in the particular year. The sample is restricted to facilities that experience only one change in ownership.

Figure A6.5: First and last reporting year of firms



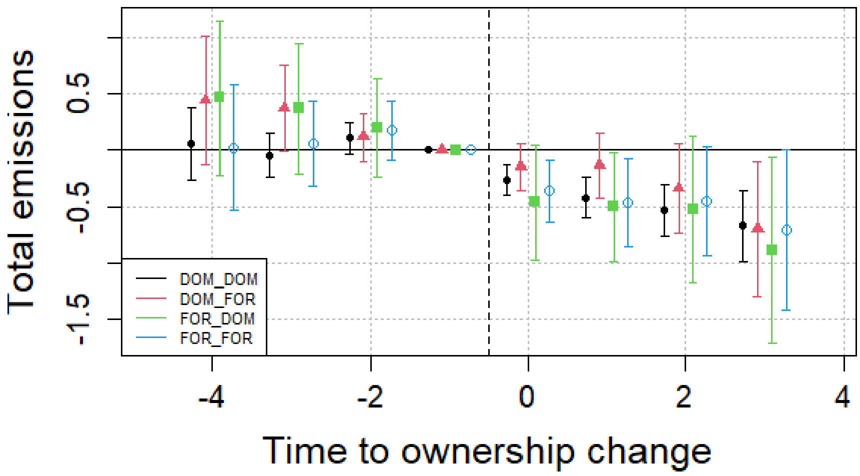
Notes: This figure shows the number of firms that have their first reporting year and last reporting year, respectively, in the particular year. The sample is restricted to firms that experience only one change in ownership.

Figure A6.6: Differences in firm characteristics for firms with and without ownership change



Notes: This figure shows a comparison of several firm characteristics for groups of firms with no, one and more than one ownership change event during our sample period. The values for capital, long-term debt, total emissions, employment, intangible fixed assets and operating revenues are inverse hyperbolic sine transformed to facilitate the comparison.

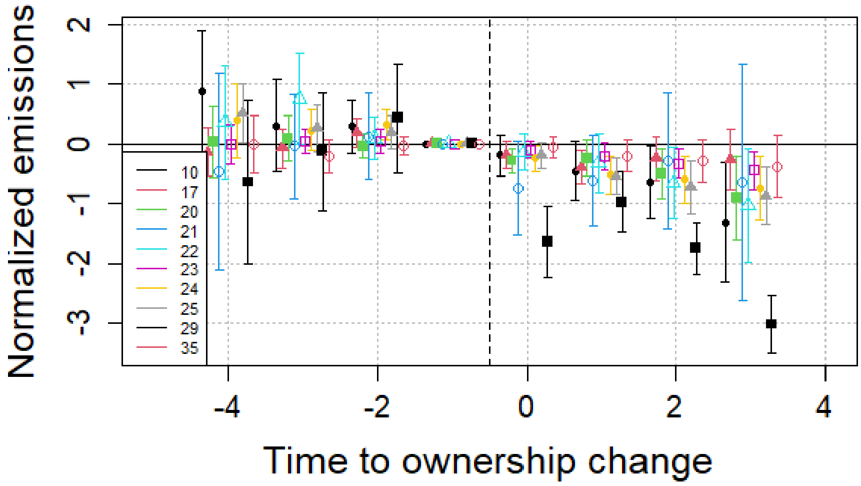
Figure A6.7: Effect on total emissions by type of ownership change at the facility level



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions for different subsamples of ownership change types for the Sun and Abraham (2021) estimator. The inverse hyperbolic sine transformation is applied to the independent variable. "DOM_DOM", "DOM_FOR", "FOR_DOM", and "FOR_FOR" are the changes from a domestic to a domestic owner, from a domestic to a foreign owner, from a foreign to a domestic owner, and from a foreign to a foreign owner, respectively.

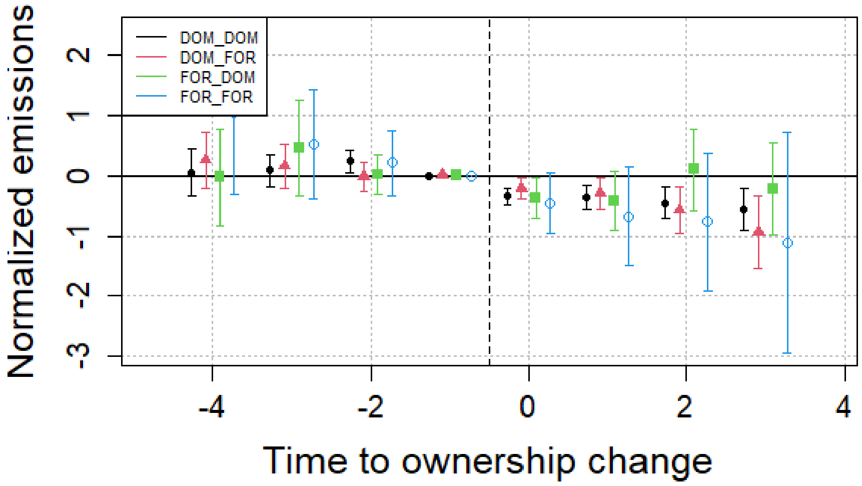
We investigate differences in the effects of ownership changes among foreign or domestic parent companies. We define a foreign (domestic) parent company when the global ultimate owner is based in a different (the same) country as facility or firm. Based on this definition, we distinguish four different cases of ownership changes: first, from a domestic to another domestic owner; second, from a domestic to a foreign owner; third, from a foreign to a domestic owner; fourth, from a foreign to a foreign owner. Dividing samples by these four different categories, we estimate the event study regression for each of the samples. The effects are similar to the overall sample and confidence intervals for all point estimates overlap for all groups, suggesting no large differences across different owner types.

Figure A6.8: Effect on total emissions by sector at the facility level



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions for different subsamples based on the facility's main sector for the Sun and Abraham (2021) estimator. The inverse hyperbolic sine transformation is applied to the independent variable. We include only sectors with a minimum of 1,000 observations. Sector 10 refers to manufacture of food products, sector 17 to manufacture of paper and paper products, sector 20 to manufacture of chemicals and chemical products, sector 21 to manufacture of basic pharmaceutical products and pharmaceutical preparations, sector 22 to manufacture of rubber and plastic products, sector 23 to manufacture of other non-metallic mineral products, sector 24 to manufacture of basic metals, sector 25 to manufacture of fabricated metal products, except machinery and equipment, sector 29 to manufacture of motor vehicles, trailers and semi-trailers, and sector 35 to electricity, gas, steam and air conditioning supply.

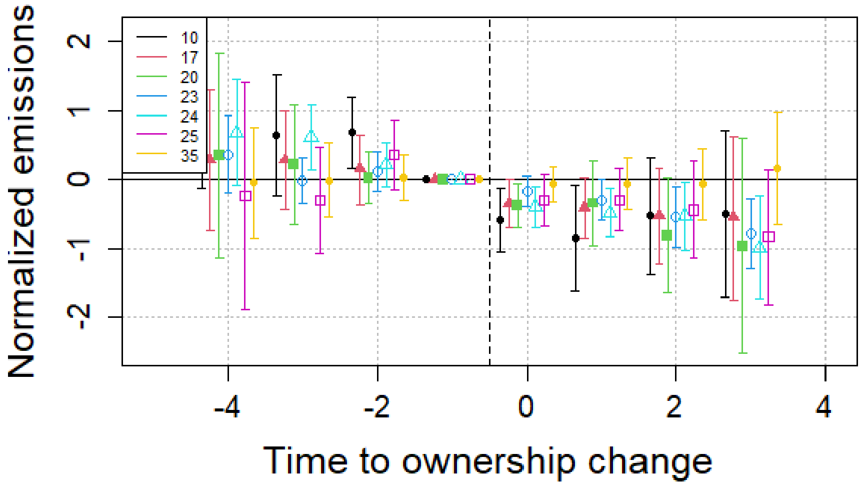
Figure A6.9: Effect on total emissions by type of ownership change at the firm level



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions for different subsamples of ownership change types for the Sun and Abraham (2021) estimator. The inverse hyperbolic sine transformation is applied to the independent variable. "DOM_DOM", "DOM_FOR", "FOR_DOM", and "FOR_FOR" are the changes from a domestic to a domestic owner, from a domestic to a foreign owner, from a foreign to a domestic owner, and from a foreign to a foreign owner, respectively.

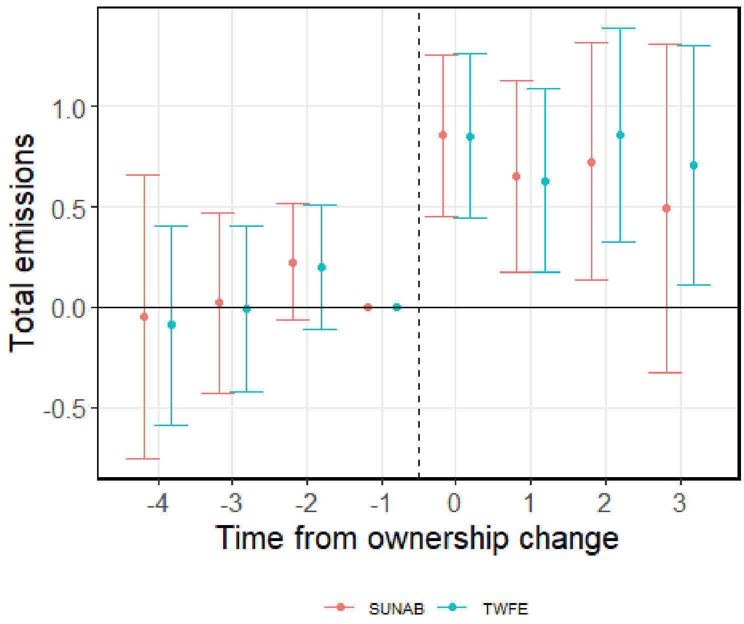
We investigate differences in the effects of ownership changes among foreign or domestic parent companies. We define a foreign (domestic) parent company when the global ultimate owner is based in a different (the same) country as facility or firm. Based on this definition, we distinguish four different cases of ownership changes: first, from a domestic to another domestic owner; second, from a domestic to a foreign owner; third, from a foreign to a domestic owner; fourth, from a foreign to a foreign owner. Dividing samples by these four different categories, we estimate the event study regression for each of the samples. The effects are similar to the overall sample and confidence intervals for all point estimates overlap for all groups, suggesting no large differences across different owner types.

Figure A6.10: Effect on total emissions by sector at the firm level



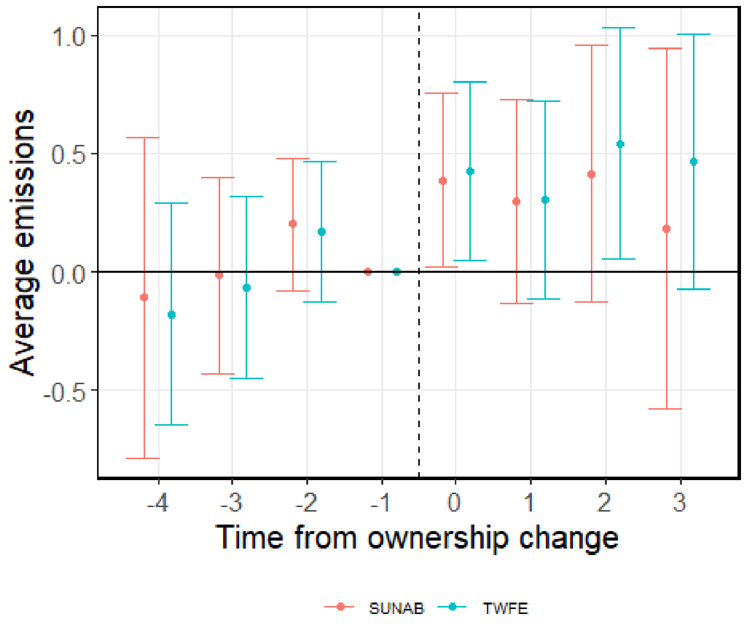
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions for different subsamples based on the firm's main sector for the Sun and Abraham (2021) estimator. The inverse hyperbolic sine transformation is applied to the independent variable. We include only sectors with a minimum of 1,000 observations. Sector 10 refers to manufacture of food products, sector 17 to manufacture of paper and paper products, sector 20 to manufacture of chemicals and chemical products, sector 23 to manufacture of other non-metallic mineral products, sector 24 to manufacture of basic metals, sector 25 to manufacture of fabricated metal products, except machinery and equipment, and sector 35 to electricity, gas, steam and air conditioning supply.

Figure A6.11: Effect on total emissions at the parent company level without shutdown of targets



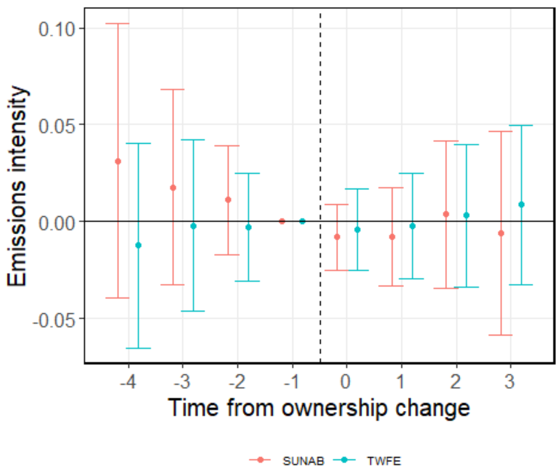
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.12: Effect on total emissions per industrial facility at the parent company level without shutdown of targets

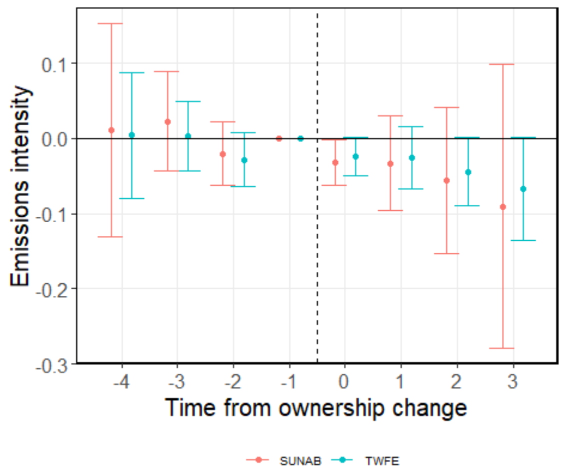


Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total emissions per industrial facility at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.13: Effect on emissions intensity at the firm level by subsample



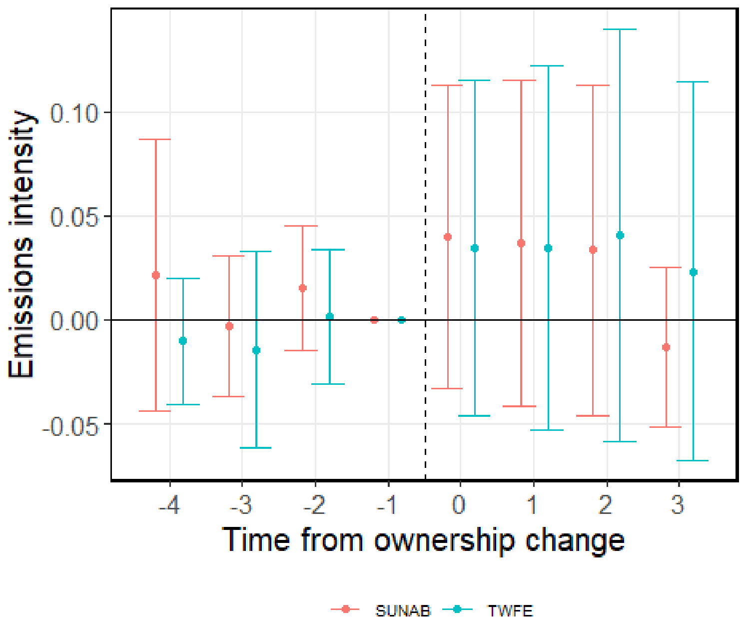
(a) Remaining firms



(b) Exiting firms

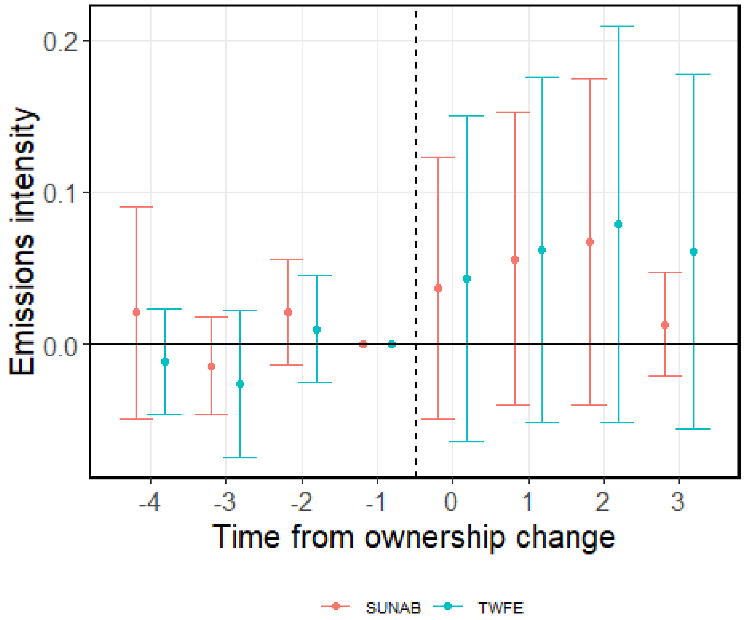
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on the emissions intensity, i.e., total emissions scaled by deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.14: Effect on emissions intensity at the parent company level



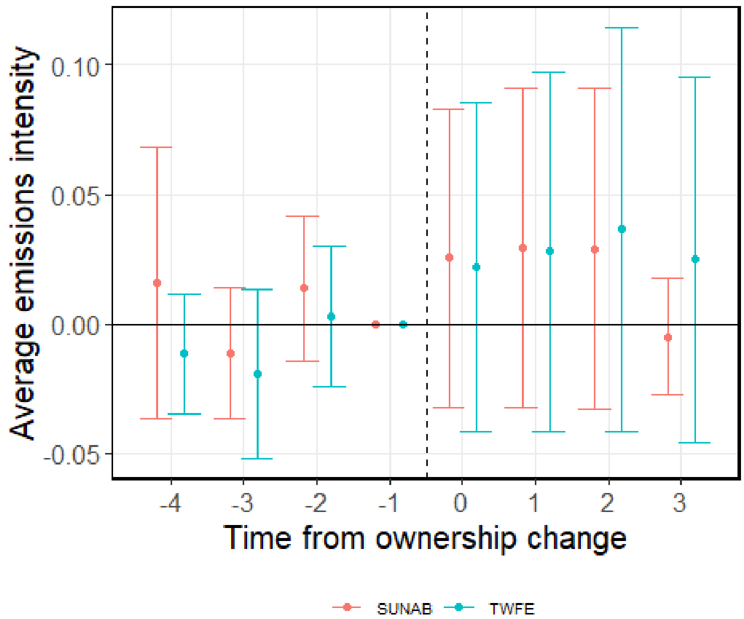
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on the emissions intensity, i.e., total emissions scaled by deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.15: Effect on emissions intensity at the parent company level without shut-down of targets



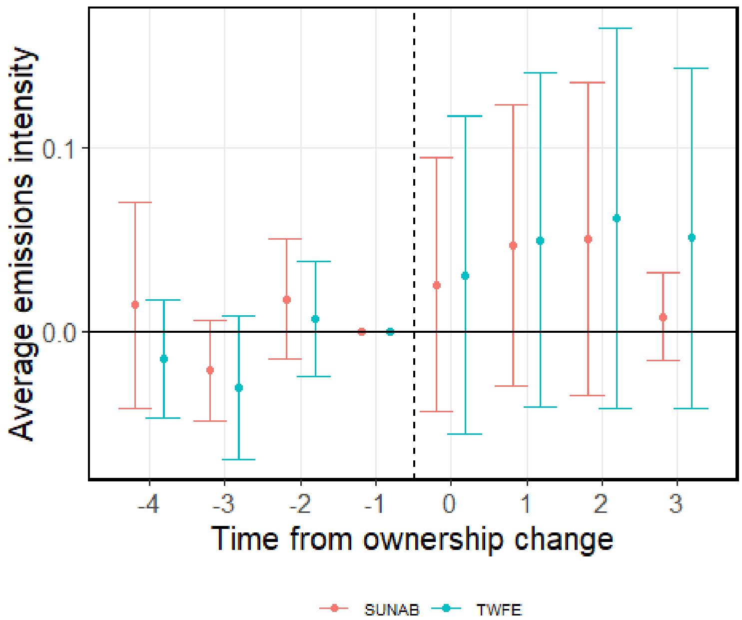
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on the emissions intensity, i.e., total emissions scaled by deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.16: Effect on emissions intensity per industrial facility at the parent company level



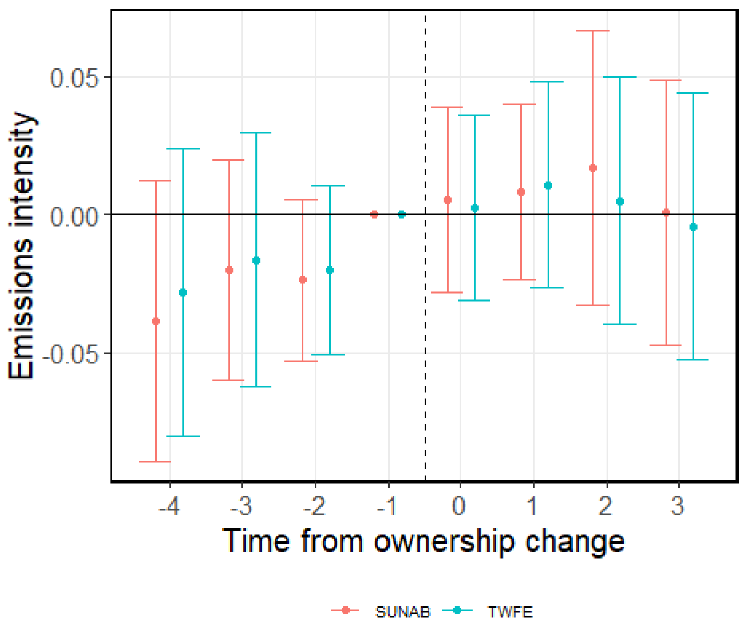
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on the emissions intensity, i.e., total emissions scaled by deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.17: Effect on emissions intensity per industrial facility at the parent company level without shutdown of targets



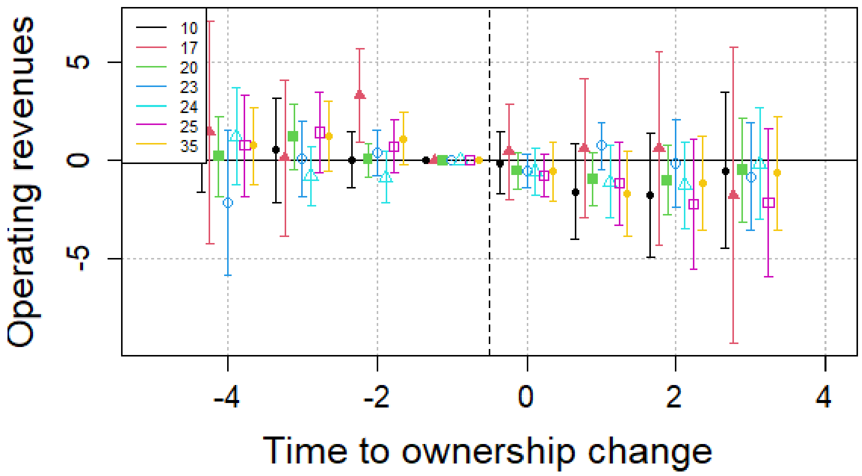
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on the emissions intensity, i.e., total emissions scaled by deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.18: Effect on emissions intensity in other facilities of the acquiring parent company



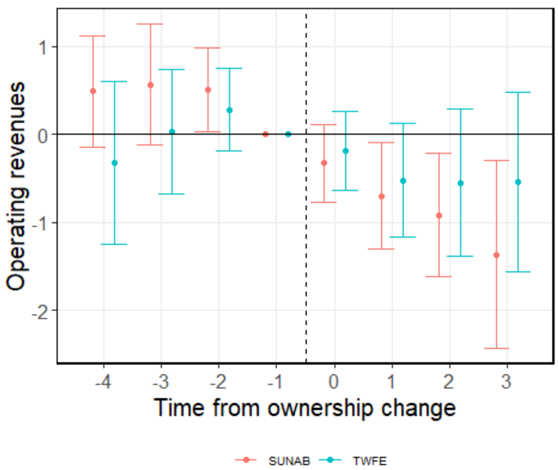
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on the emissions intensity, i.e., total emissions scaled by deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.19: Effect on output by sector at the firm level

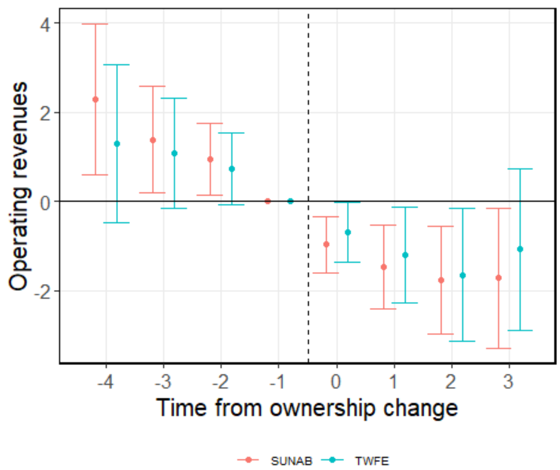


Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on operating revenues for different subsamples based on the firm's main sector for the Sun and Abraham (2021) estimator. The inverse hyperbolic sine transformation is applied to the independent variable. We include only sectors with a minimum of 1,000 observations. Sector 10 refers to manufacture of food products, sector 17 to manufacture of paper and paper products, sector 20 to manufacture of chemicals and chemical products, sector 23 to manufacture of other non-metallic mineral products, sector 24 to manufacture of basic metals, sector 25 to manufacture of fabricated metal products, except machinery and equipment, and sector 35 to electricity, gas, steam and air conditioning supply.

Figure A6.20: Effect on output at the firm level by subsample



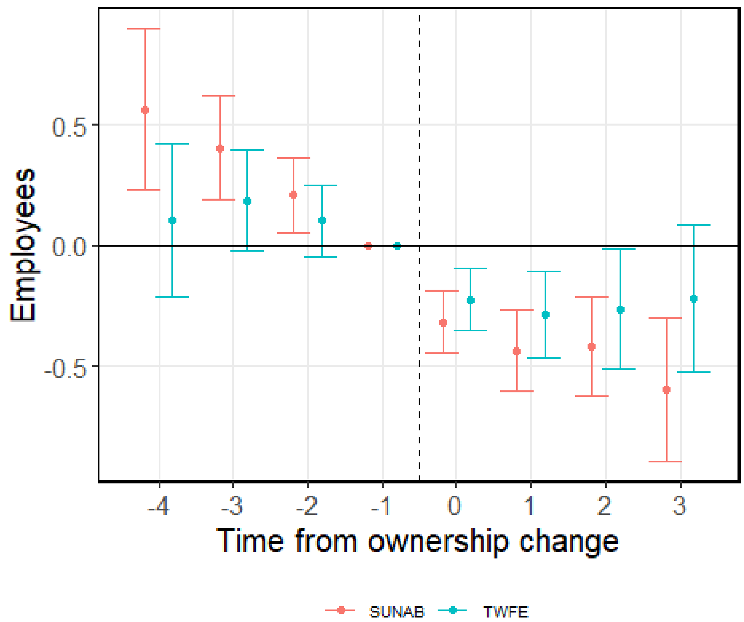
(a) Remaining firms



(b) Exiting firms

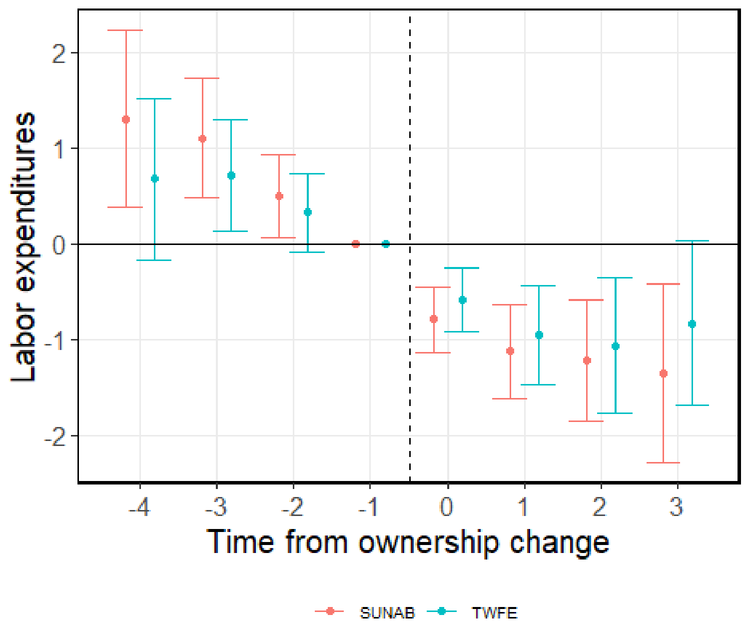
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.21: Effect on labour input at the firm level



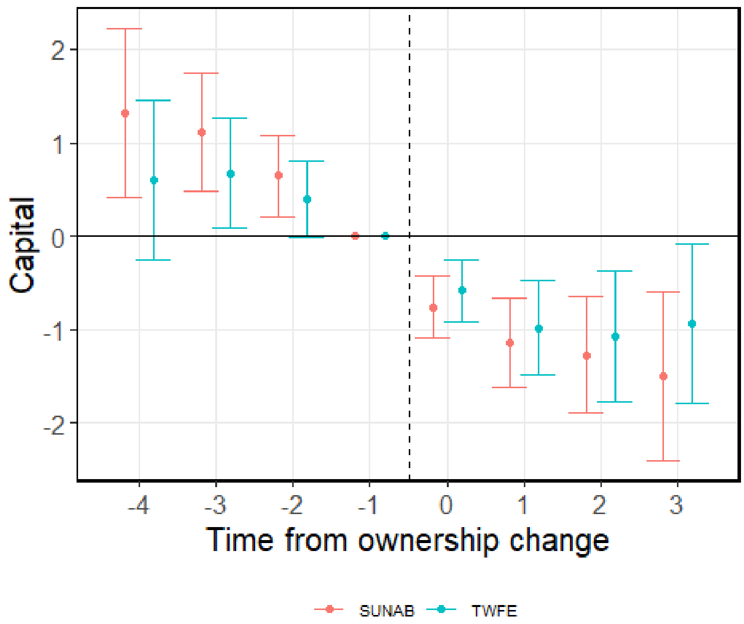
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on labour input. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.22: Effect on labour expenses at the firm level



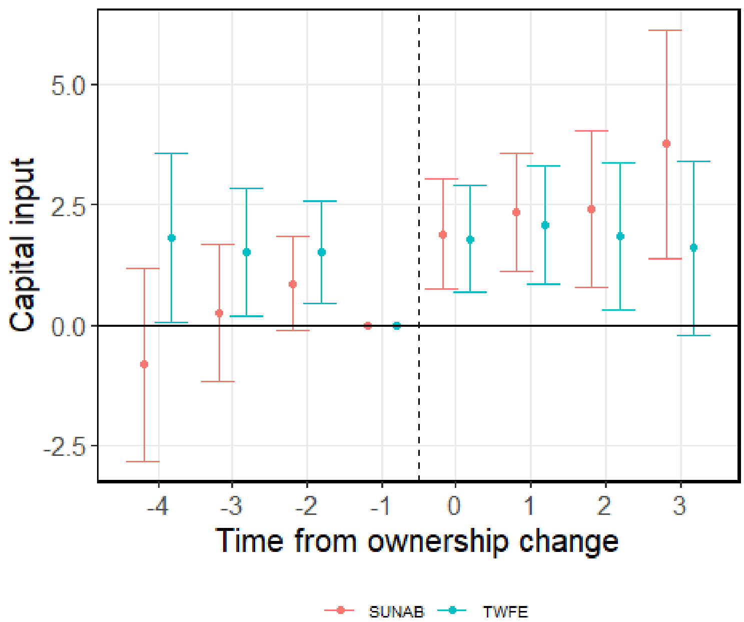
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on labour expenses. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.23: Effect on capital input at the firm level



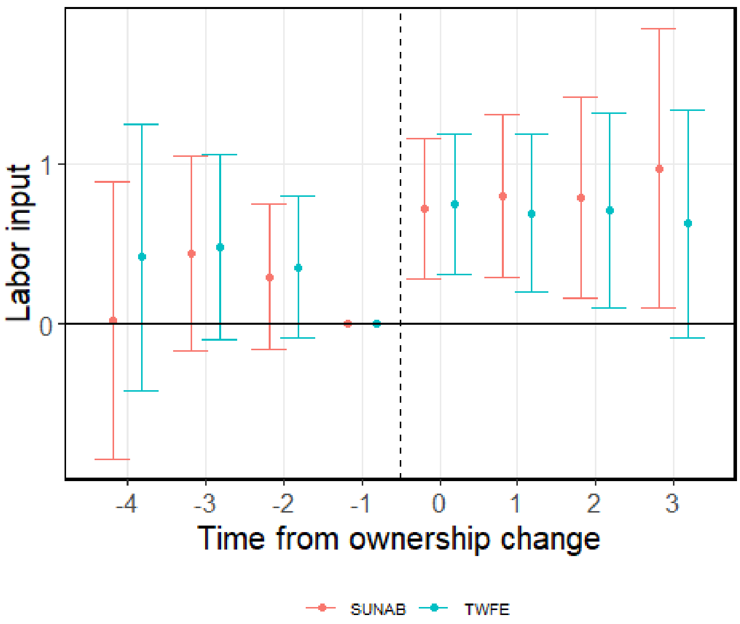
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on capital input. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.24: Effect on capital input at the parent company level



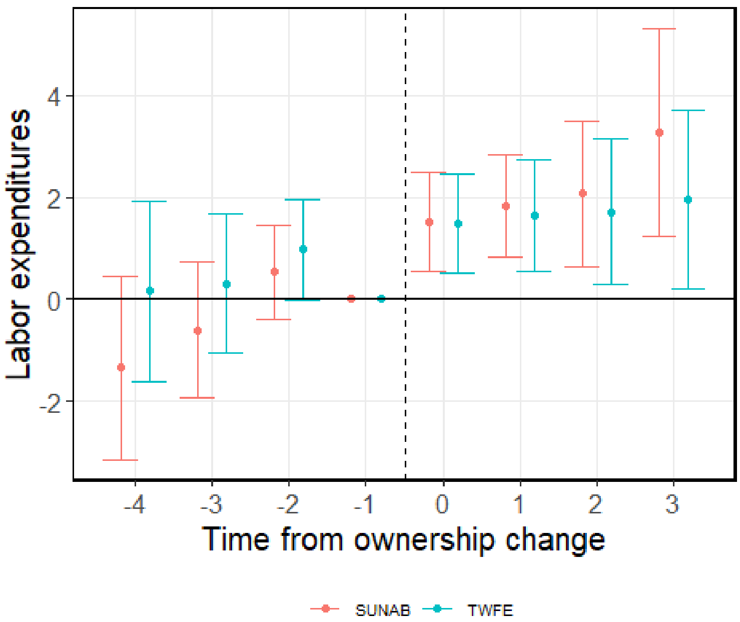
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on capital input. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.25: Effect on labor input at the parent company level



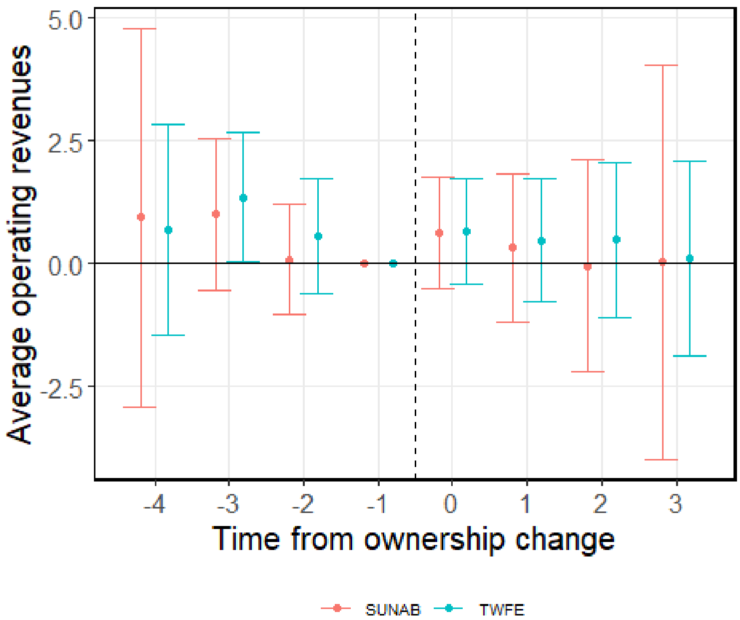
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on labor input. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.26: Effect on labor expenses at the parent company level



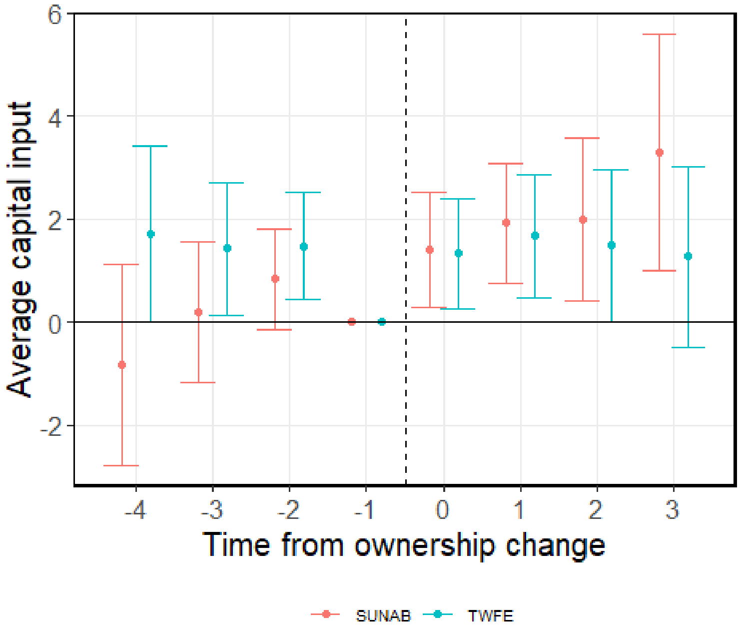
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on labor expenses. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.27: Effect on output per industrial facility at the parent company level



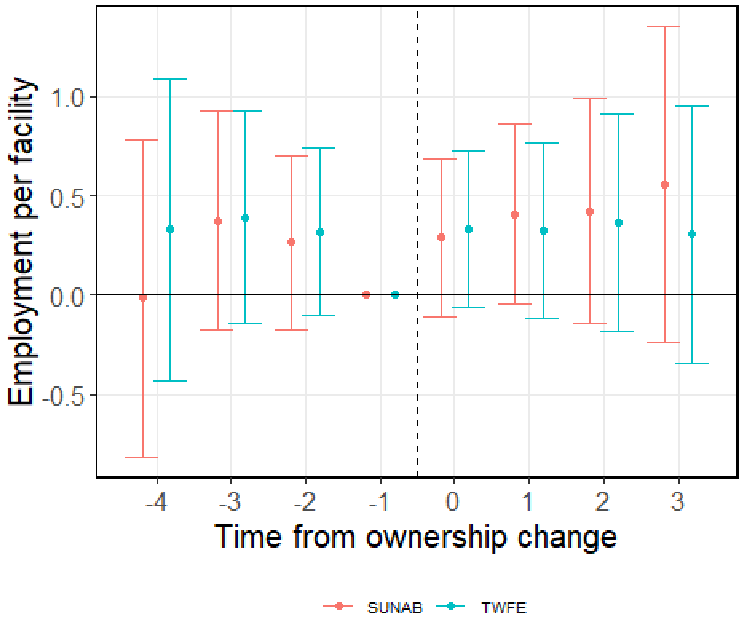
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.28: Effect on capital input per industrial facility at the parent company level



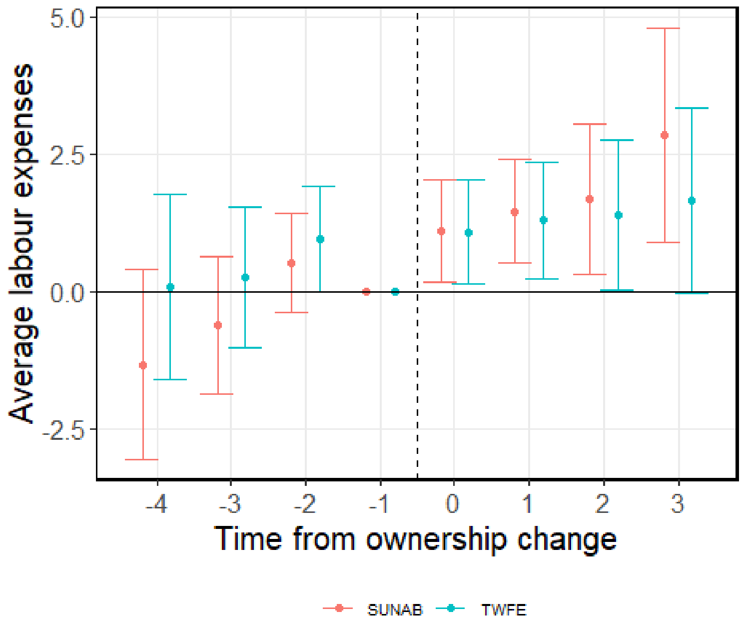
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on capital input. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.29: Effect on labor input per industrial facility at the parent company level



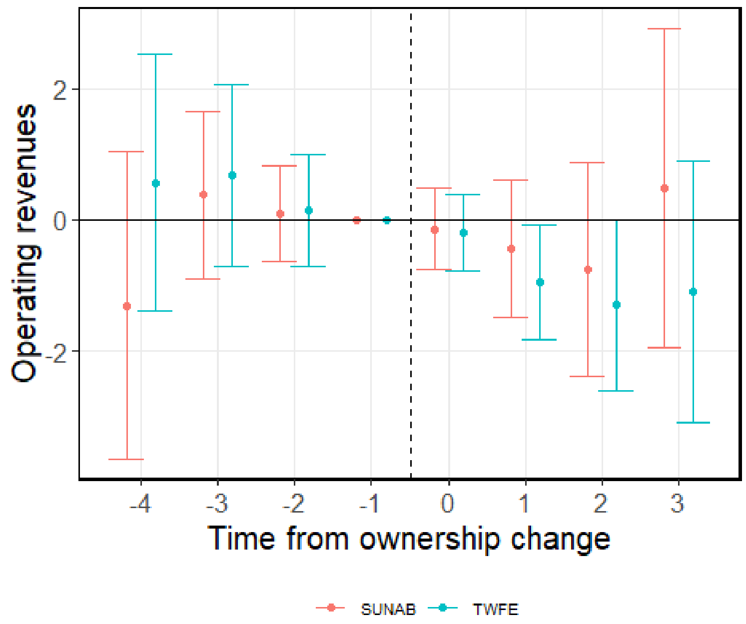
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on labor input. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.30: Effect on labor expenses per industrial facility at the parent company level



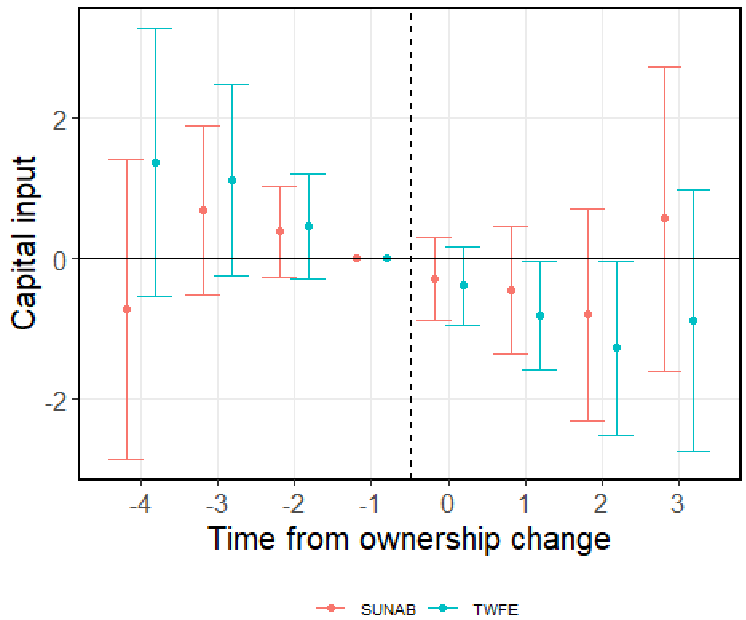
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on labor expenses. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.31: Effect on output of other facilities of the acquiring parent company



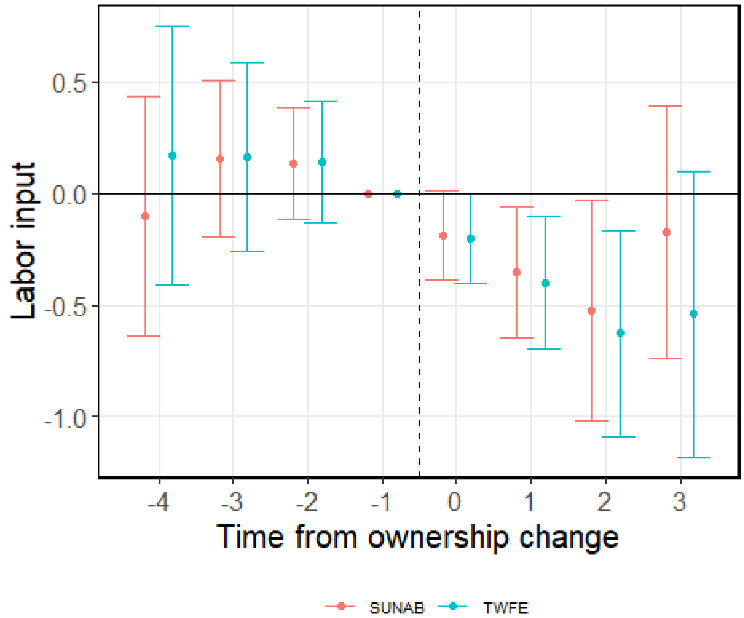
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated operating revenues. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.32: Effect on capital input of other facilities of the acquiring parent company



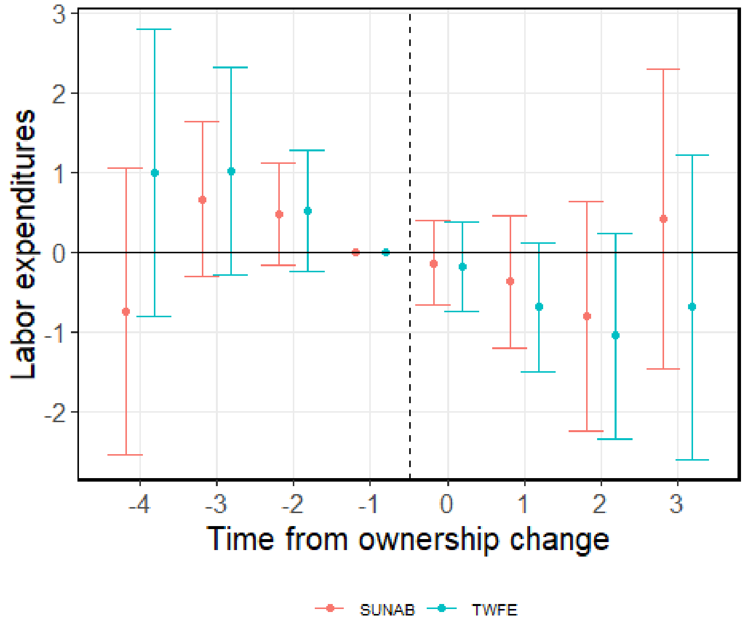
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on capital input. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.33: Effect on labor input of other facilities of the acquiring parent company



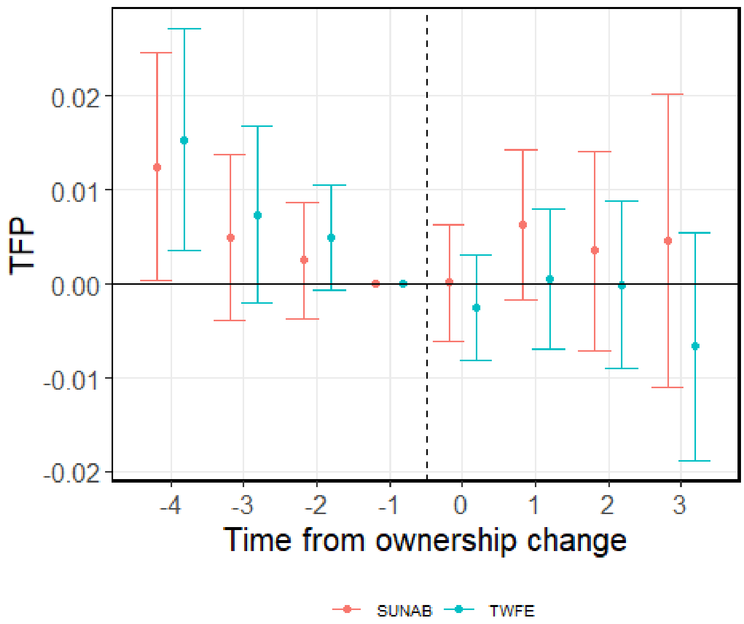
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on labor input. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.34: Effect on labor expenses of other facilities of the acquiring parent company



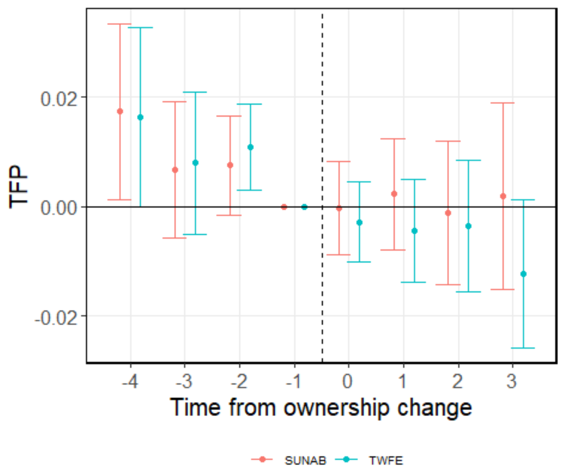
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on labor expenses. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.35: Effect on total factor productivity at the firm level

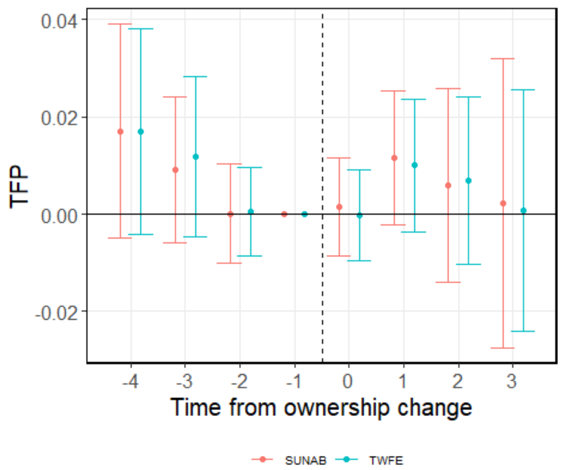


Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.36: Effect on total factor productivity at the firm level by subsample



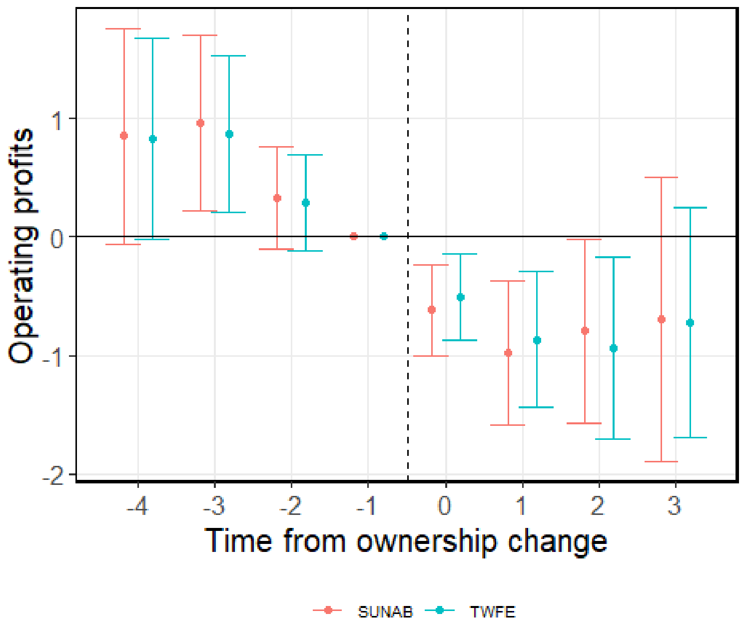
(a) Remaining firms



(b) Exiting firms

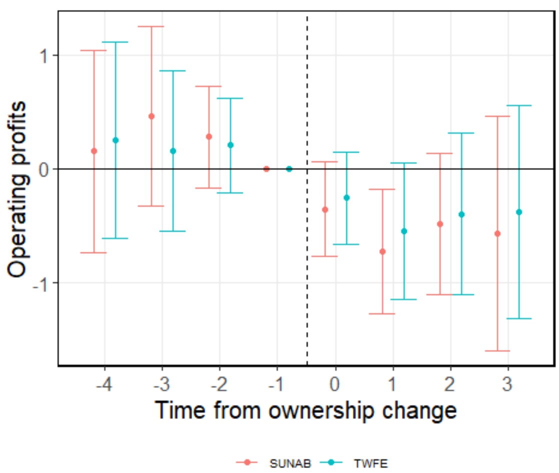
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.37: Effect on operating profits at the firm level

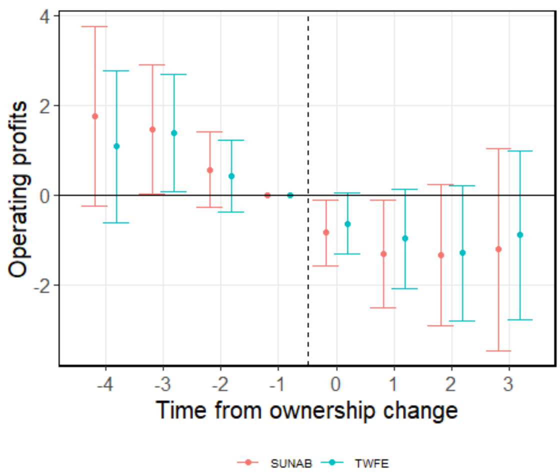


Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.38: Effect on operating profits at the firm level by subsample



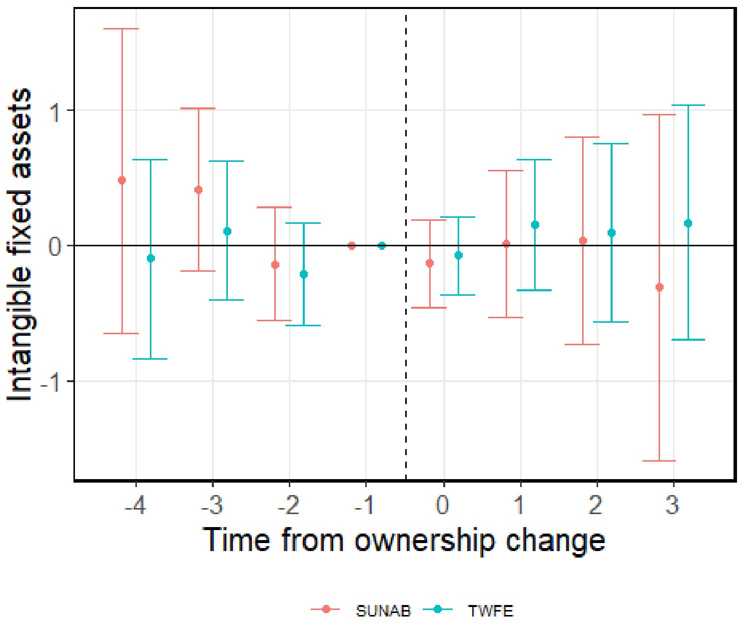
(a) Remaining firms



(b) Exiting firms

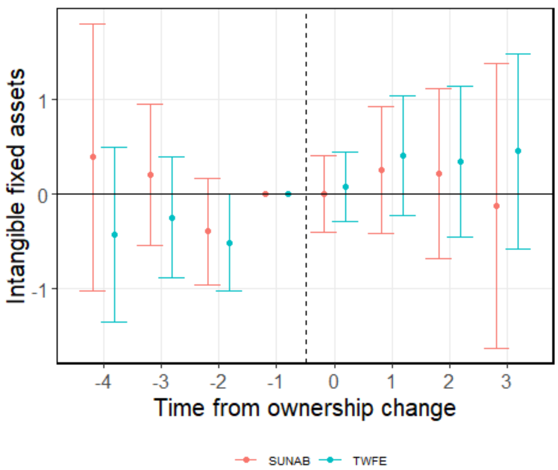
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.39: Effect on intangible fixed assets at the firm level

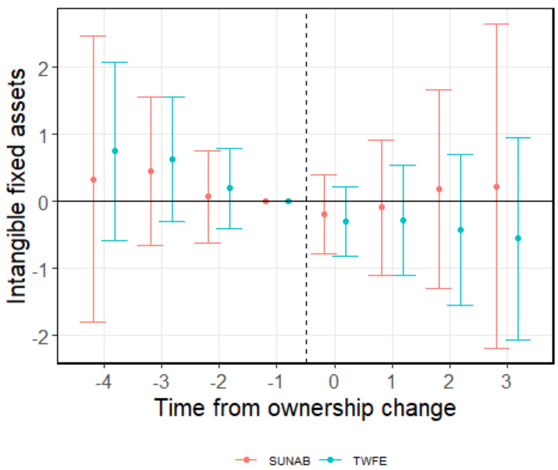


Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on intangible fixed assets. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.40: Effect on intangible fixed assets at the firm level by subsample



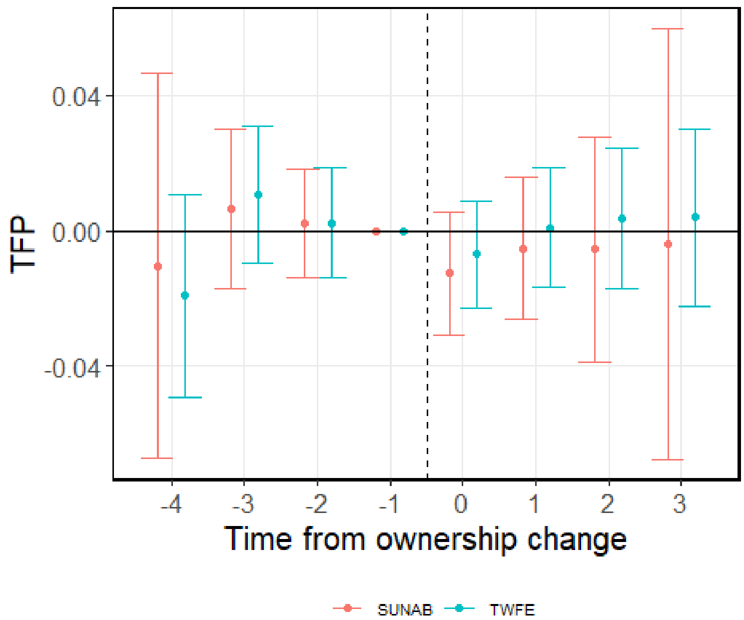
(a) Remaining firms



(b) Exiting firms

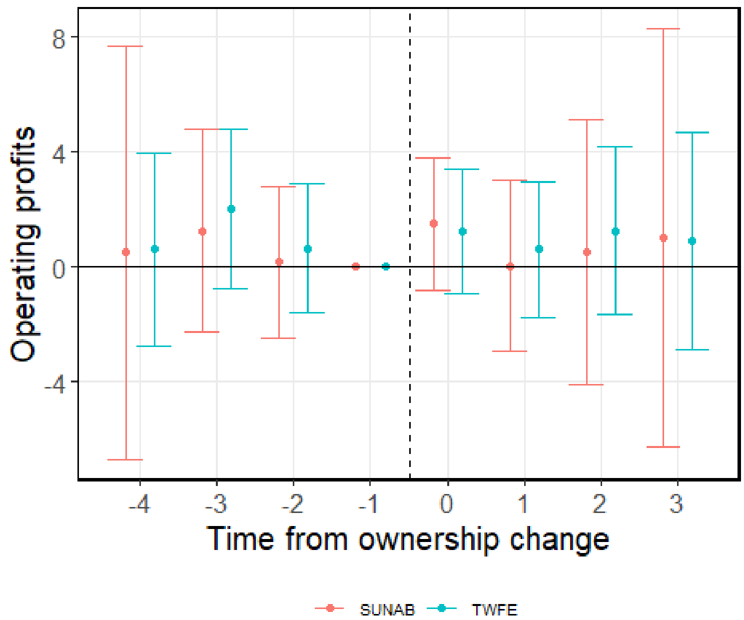
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on intangible fixed assets. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.41: Effect on total factor productivity at the parent company level



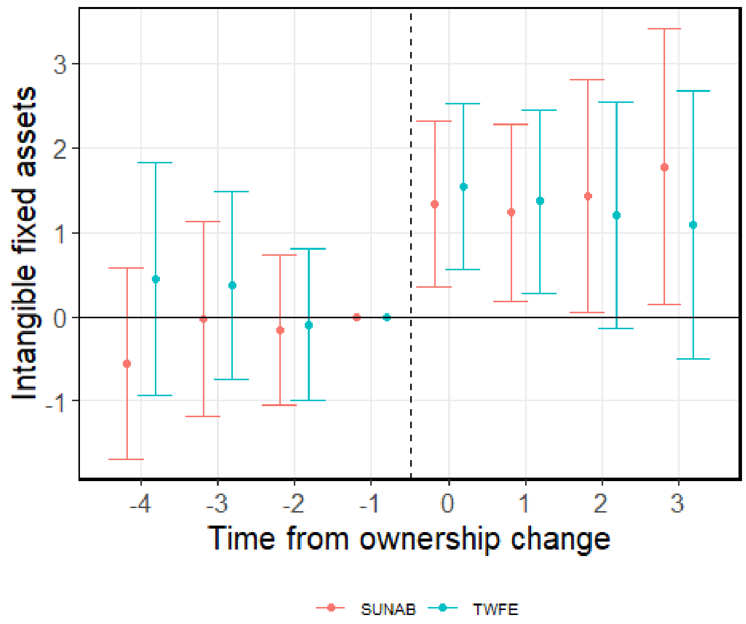
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.42: Effect on operating profits at the parent company level



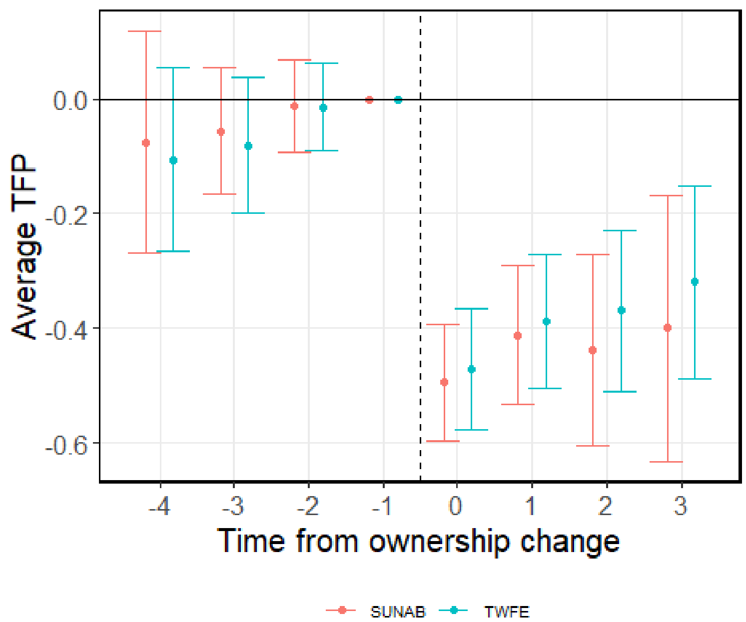
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on operating profits. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.43: Effect on intangible fixed assets at the parent company level



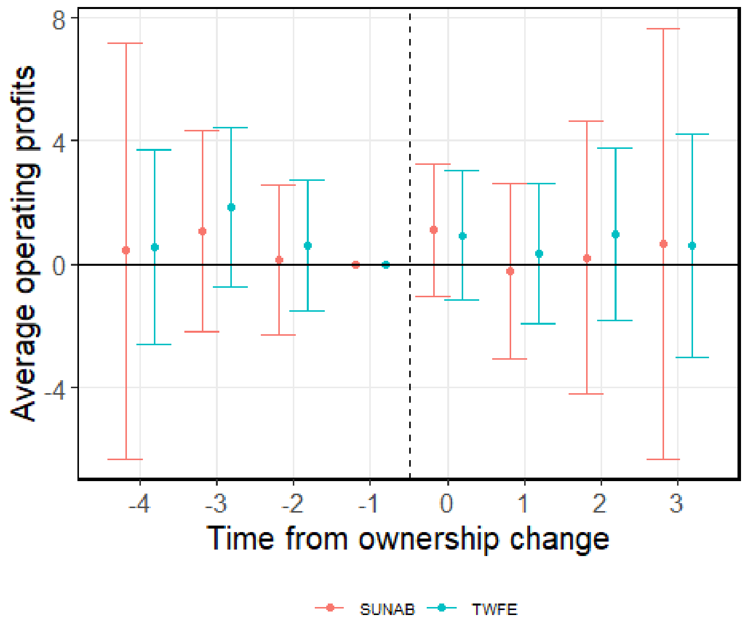
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on intangible fixed assets. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.44: Effect on total factor productivity per industrial facility in the parent company



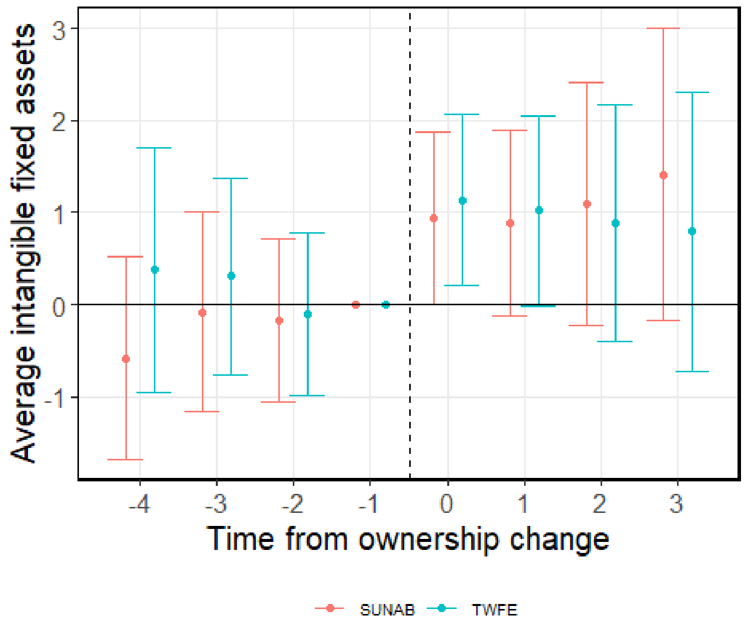
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.45: Effect on operating profits per industrial facility in the parent company



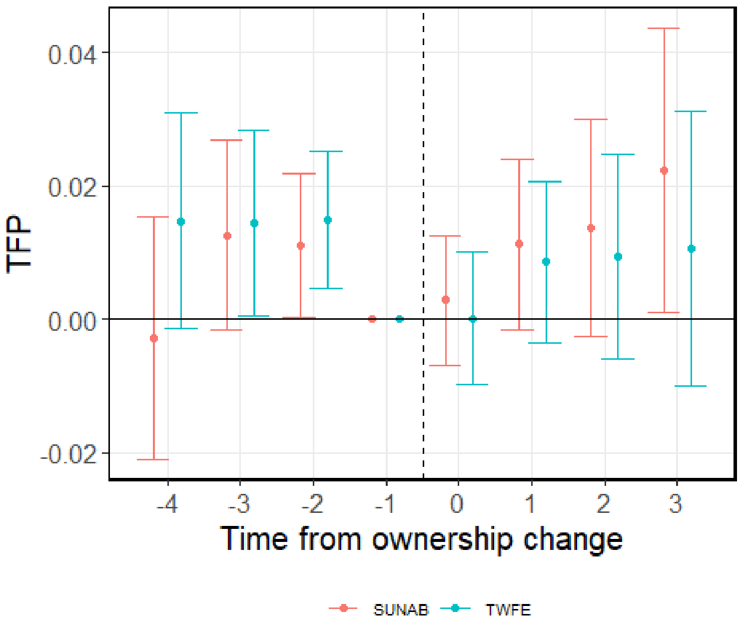
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on operating profits. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.46: Effect on intangible fixed assets per industrial facility in the parent company



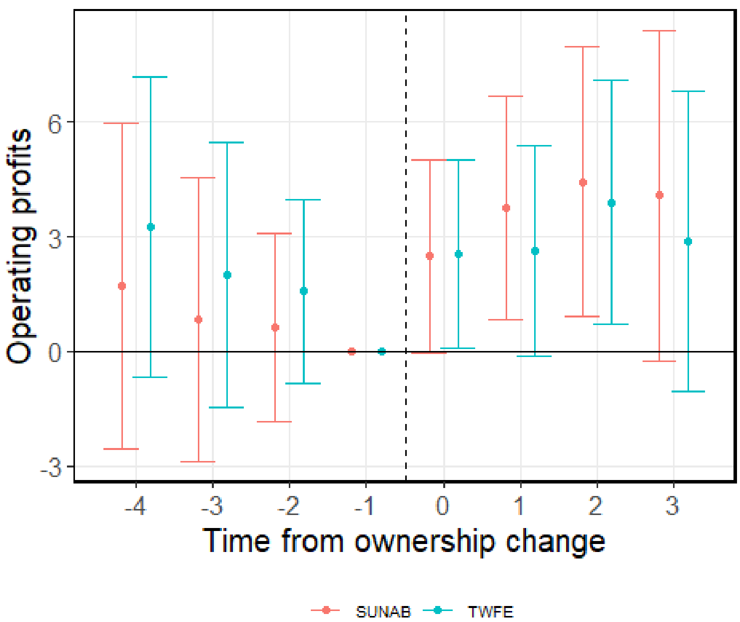
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on intangible fixed assets. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.47: Effect on total factor productivity in other industrial facilities of the acquiring parent company



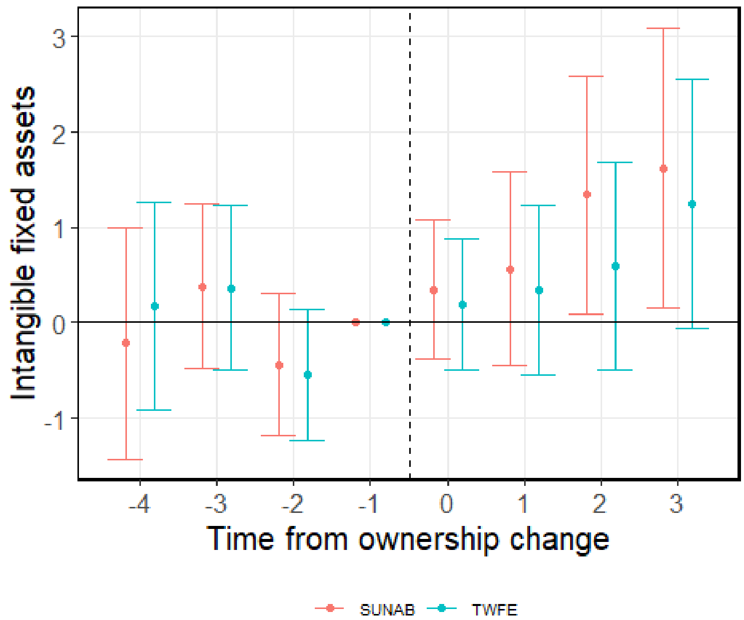
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.48: Effect on operating profits in other industrial facilities of the acquiring parent company



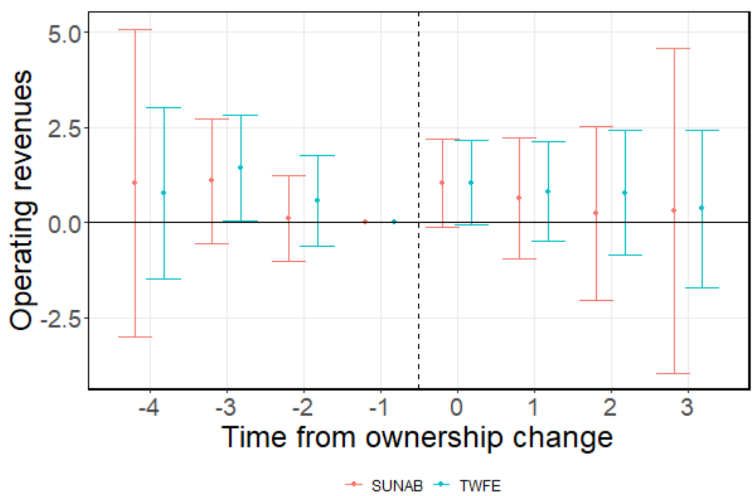
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on operating profits. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.49: Effect on intangible fixed assets in other industrial facilities of the acquiring parent company



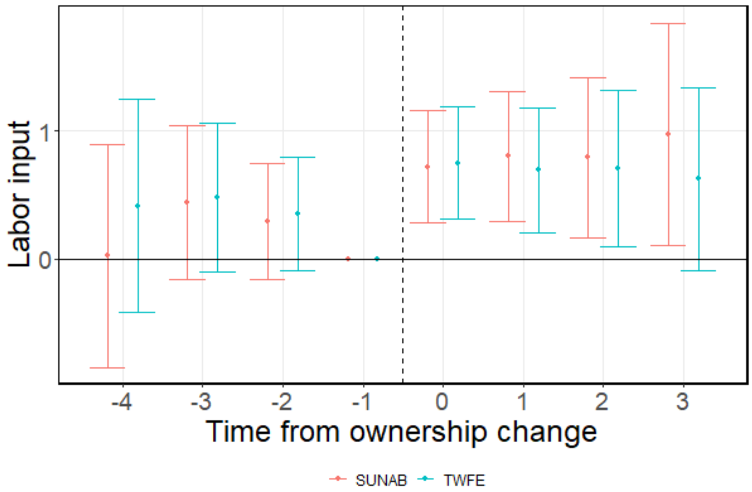
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on intangible fixed assets. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.50: Effect on output at the parent company level



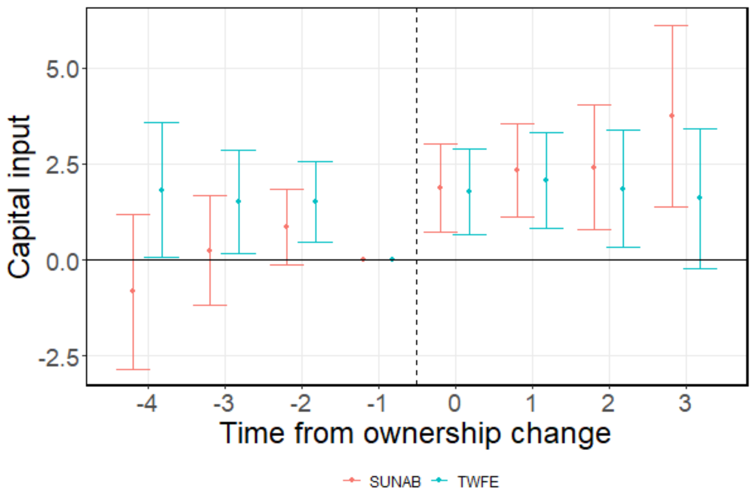
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated operating revenues at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.51: Effect on labor input at the parent company level



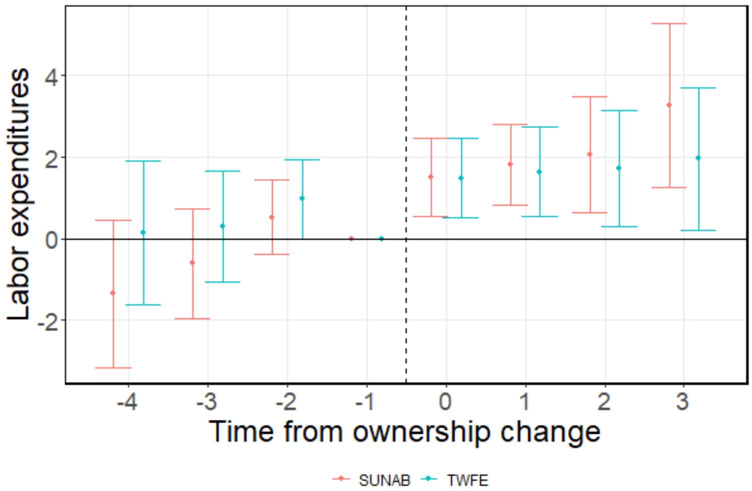
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on number of employees at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.52: Effect on capital input at the parent company level



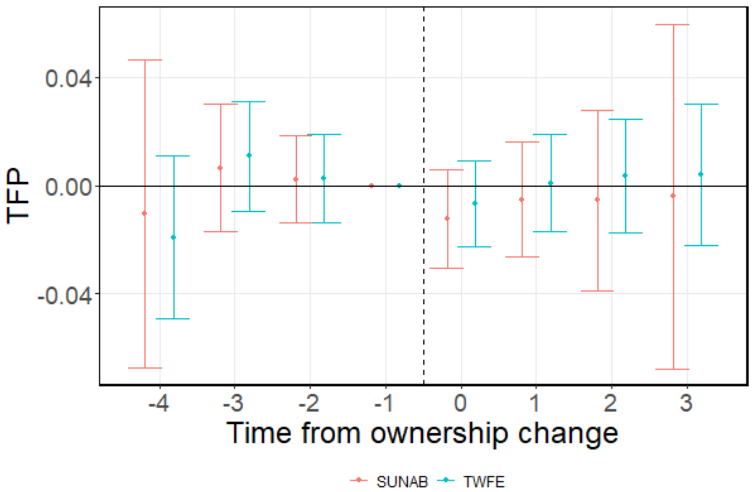
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated tangible fixed assets at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.53: Effect on labor expenditures at the parent company level



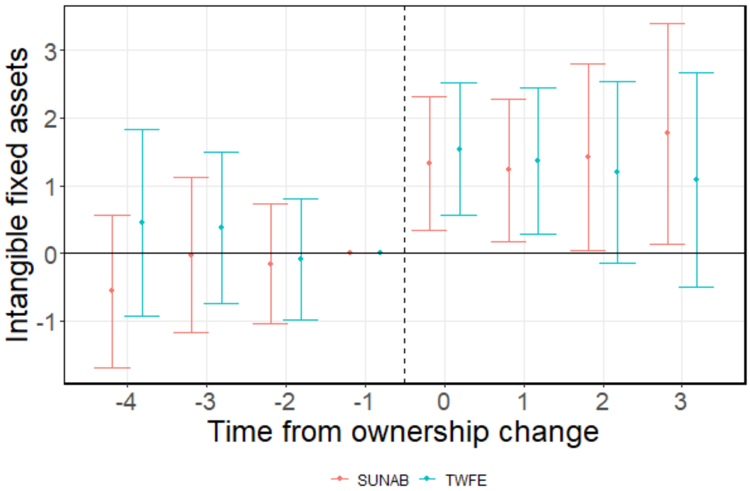
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated labor expenditures at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.54: Effect on total factor productivity (TFP) at the parent company level



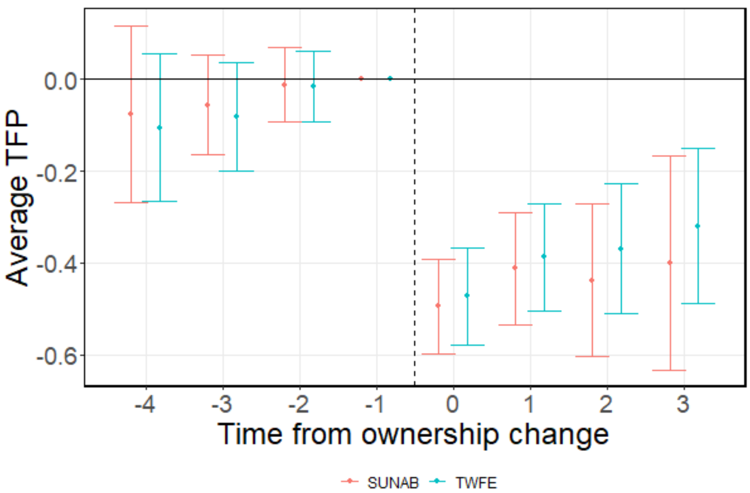
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.55: Effect on intangible fixed assets at the parent company level



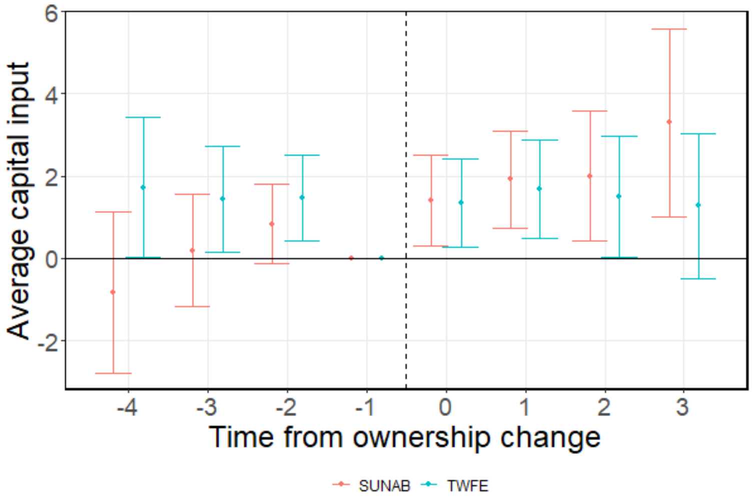
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated intangible fixed assets at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.56: Effect on average total factor productivity (TFP) per industrial facility at the parent company level



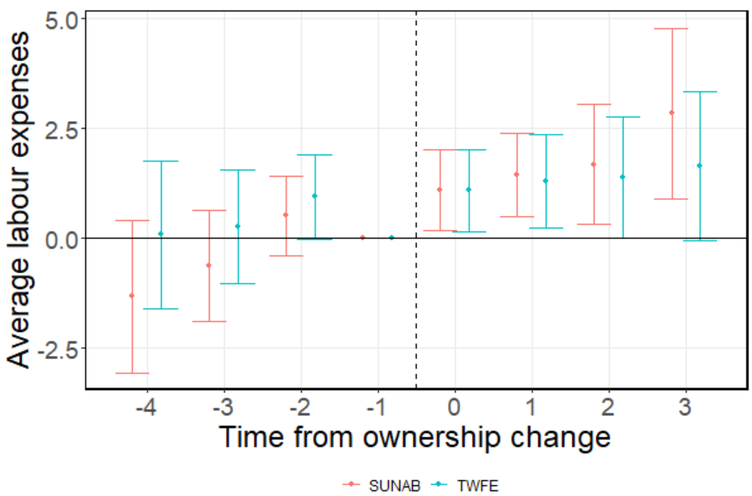
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity per industrial facility at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.57: Effect on average capital input per industrial facility at the parent company level



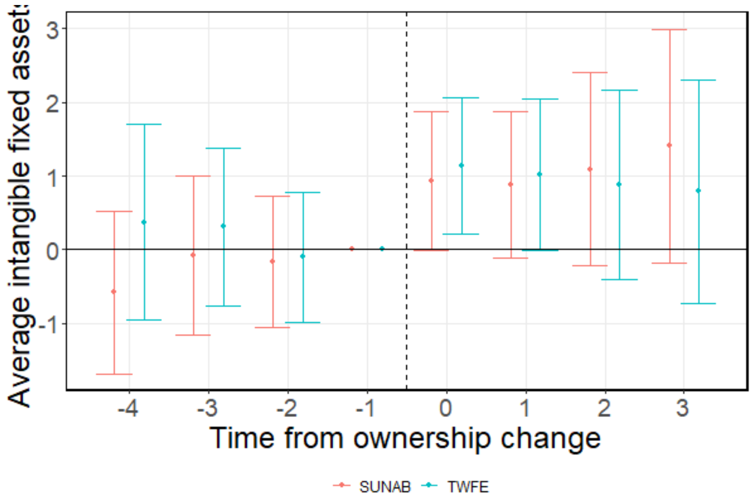
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated total fixed assets per industrial facility at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.58: Effect on average labour expenditures per industrial facility at the parent company level



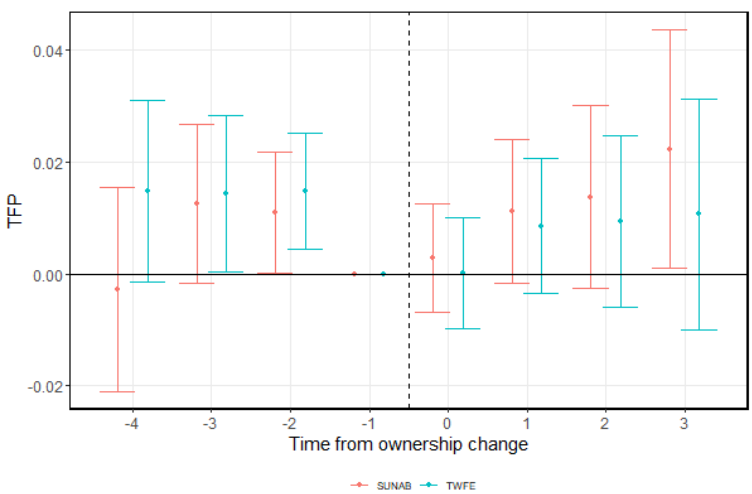
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on labor expenditures per industrial facility at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.59: Effect on average intangible fixed assets per industrial facility at the parent company level



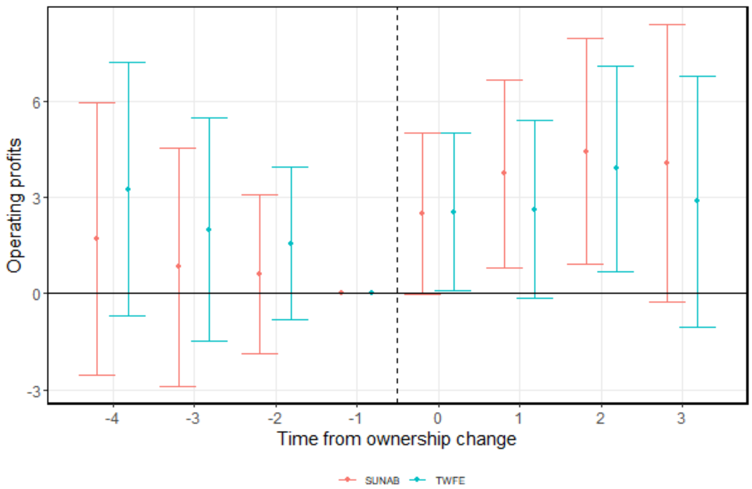
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on deflated intangible fixed assets per industrial facility at the parent company level. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.60: Effect on total factor productivity (TFP) for the other firms of the parent company acquiring a new facility



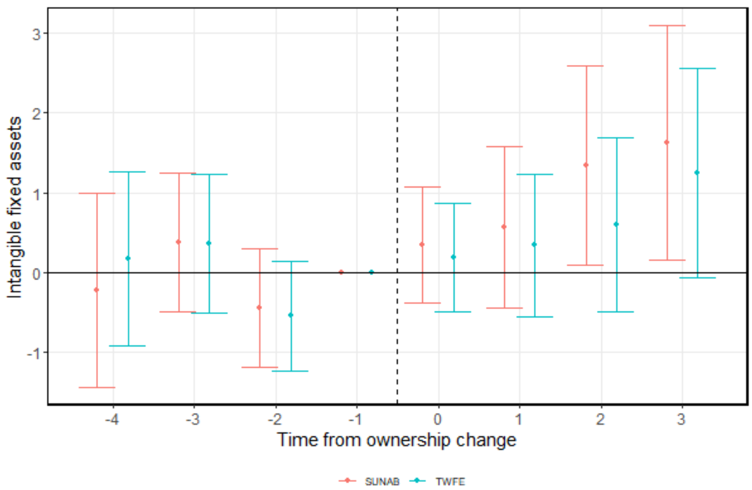
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on total factor productivity for the other firms of the parent company acquiring a new facility. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.61: Effect on operating profits for the other firms of the parent company acquiring a new facility



Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on operating profits for the other firms of the parent company acquiring a new facility. The inverse hyperbolic sine transformation is applied to the independent variable.

Figure A6.62: Effect on intangible fixed assets for the other firms of the parent company acquiring a new facility



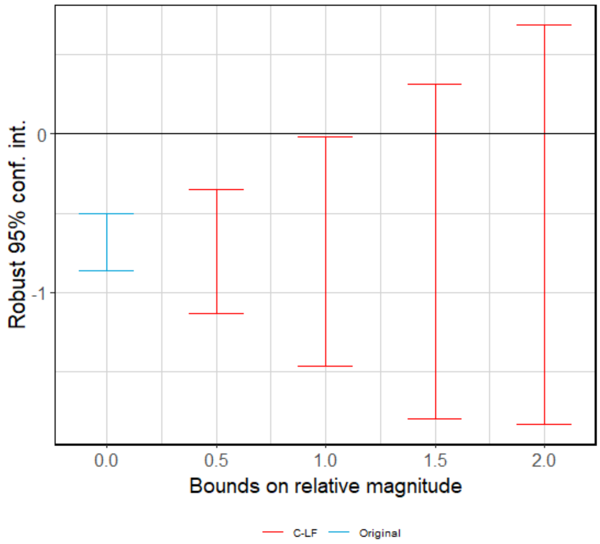
Notes: This figure shows the estimates and 95% confidence intervals for the event study coefficients for the effect of an ownership change on intangible fixed assets for the other firms of the parent company acquiring a new facility. The inverse hyperbolic sine transformation is applied to the independent variable.

A6.4 Sensitivity analysis on pre-trends

Our identification in the event study rests on the parallel trends assumption that facilities and firms that experienced an ownership change would have developed the same way in absence of treatment as entities that do not (yet) experience an ownership change. A common check for this assumption is to look at pre-trends before treatment. In this setting, pre-treatment coefficients for some outcomes at the firm level are significantly different from zero and even show a falling trend in the years before acquisition. Therefore, we cannot readily interpret the estimated effect after treatment as causal impact of the ownership change. To get an idea about how significant pre-trends could have affected the robustness of our findings, we apply a method by Rambachan and Roth (2023). Their approach estimates the magnitude of the post-treatment violations of parallel trends, relative to the observed maximum pre-treatment violation, and provides the bounds of relative magnitude in post-treatment violation at which the estimated coefficient would turn insignificant. The assumption behind it is that the violation of parallel trends in the post-treatment period may be similar to that in the pre-treatment period. Bounds of relative magnitude equal to 1 would impose that the post-treatment violation is not stronger than the strongest pre-treatment violation between consecutive periods. The results from the test provide a check on how sensitive estimates are to violation of the parallel trends assumption.

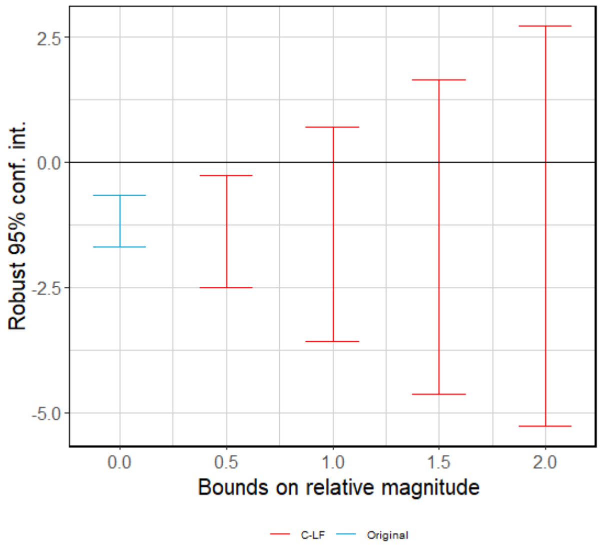
In our setting, we conduct the test on the bounds of relative magnitude for the fourth post-treatment coefficient (the third year after the ownership change). The standard approach is to test sensitivity for the first coefficient after treatment. However, most of our treatment effects increase over time and are largest at the end of the post-treatment period.

Figure A6.63: Test on the bounds of relative magnitude for the effect on emissions at the facility level



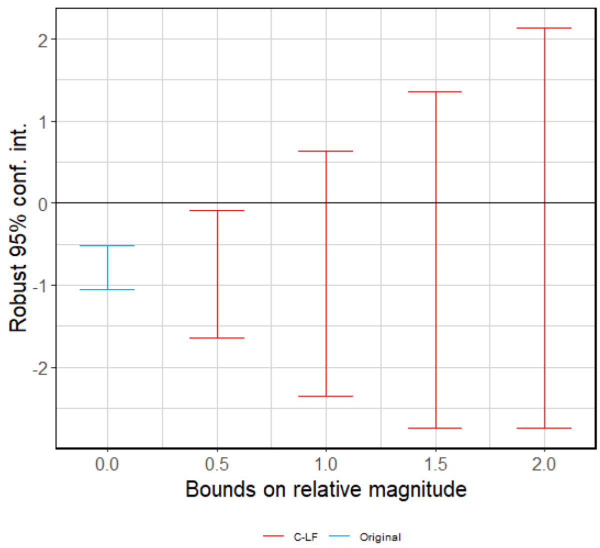
Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is up to the maximum violation oberved in the pre-treatment period.

Figure A6.64: Test on the bounds of relative magnitude for the effect on emissions for exiting facilities



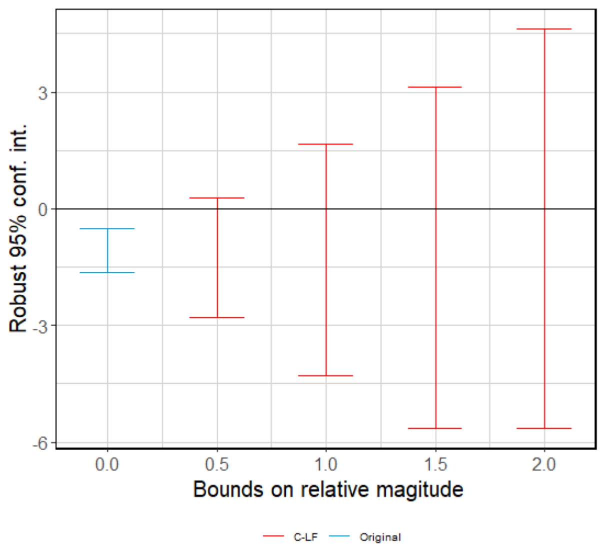
Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is half or less of the maximum violation observed in the pre-treatment period.

Figure A6.65: Test on the bounds of relative magnitude for the effect on emissions at the firm level



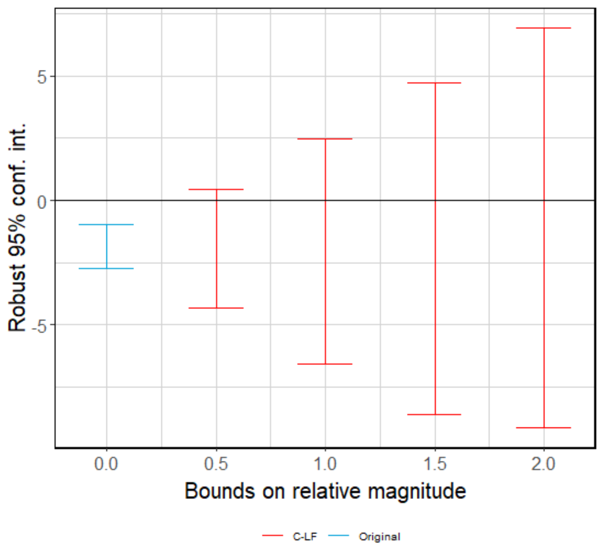
Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is half or less of the maximum violation observed in the pre-treatment period.

Figure A6.66: Test on the bounds of relative magnitude for the effect on emissions for exiting firms



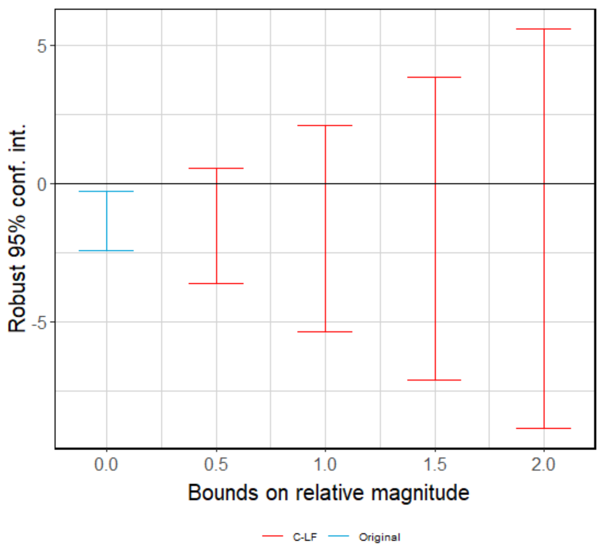
Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is less than half of the maximum violation observed in the pre-treatment period.

Figure A6.67: Test on the bounds of relative magnitude for the effect on output at the firm level



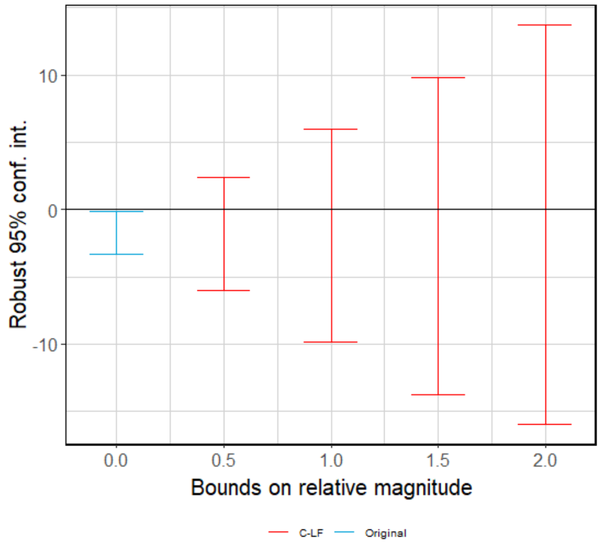
Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is less than half of the maximum violation observed in the pre-treatment period.

Figure A6.68: Test on the bounds of relative magnitude for the effect on output for firms remaining in operation



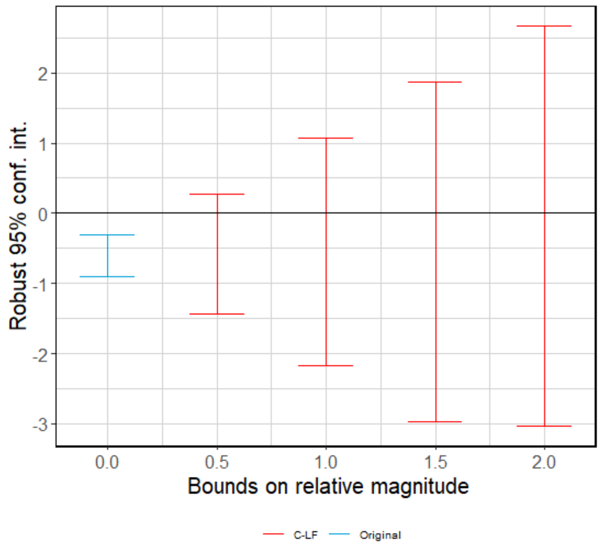
Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is less than half of the maximum violation observed in the pre-treatment period.

Figure A6.69: Test on the bounds of relative magnitude for the effect on output for exiting firms



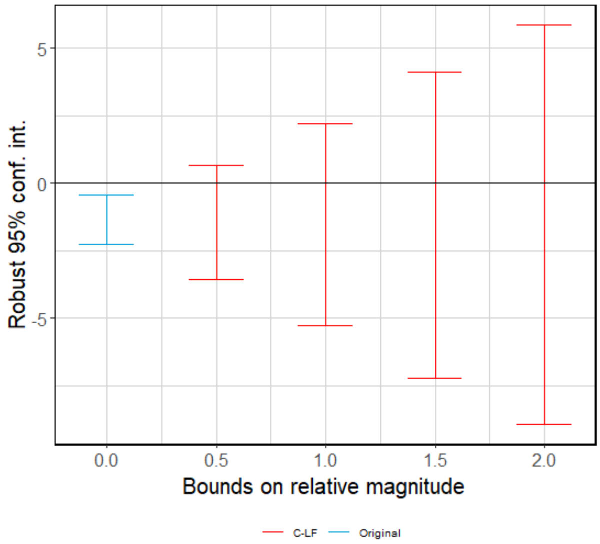
Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is less than half of the maximum violation observed in the pre-treatment period.

Figure A6.70: Test on the bounds of relative magnitude for the effect on labor input at the firm level



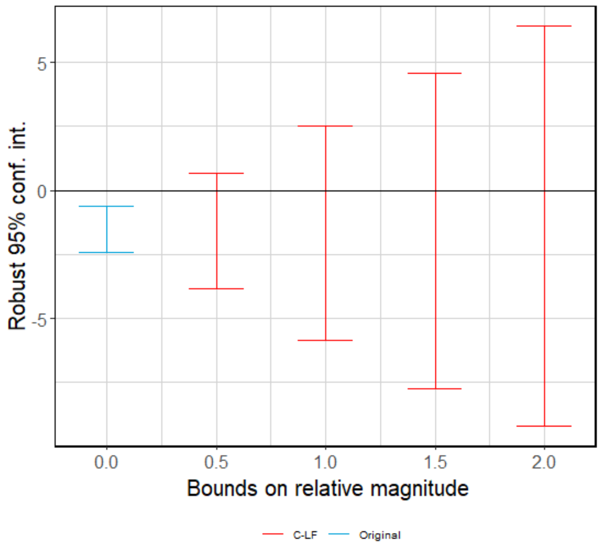
Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is less than half of the maximum violation observed in the pre-treatment period.

Figure A6.71: Test on the bounds of relative magnitude for the effect on labor expenses at the firm level



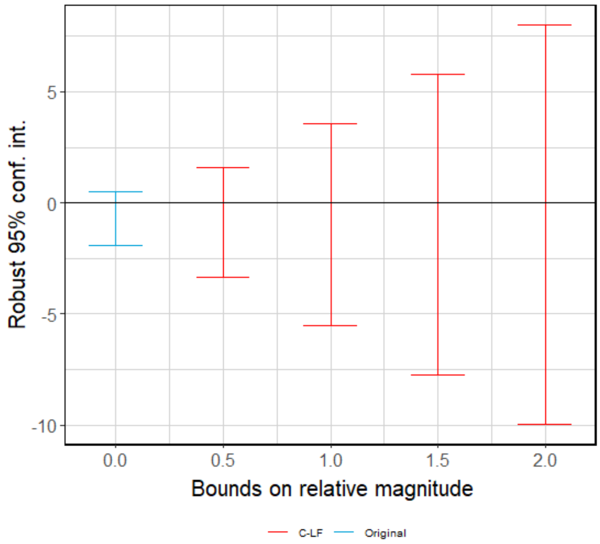
Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is less than half of the maximum violation observed in the pre-treatment period.

Figure A6.72: Test on the bounds of relative magnitude for the effect on capital input at the firm level



Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The coefficient is significant as long as the deviation from parallel trends is less than half of the maximum violation observed in the pre-treatment period.

Figure A6.73: Test on the bounds of relative magnitude for the effect on profits at the firm level



Notes: This figure shows the event study coefficient with 95% confidence interval in the fourth year after the ownership change in blue and robust coefficients with 95% interval in red that adjust for different degrees of violation in parallel trends in the post-treatment period relative to violation in the pre-treatment period. The original coefficient is not significant in the fourth year after the ownership change. When adjusting for deviations from parallel trends, adjusted coefficients are even less significant.