Downloaded from orbit.dtu.dk on: Apr 18, 2024

DTU Library

Encouraging Residential Energy

Petersen, Sebastian Christoph

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Petersen, S. C. (2017). *Encouraging Residential Energy*. DTU Management.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Technical University of Denmark

Department of Management Engineering

Ph.D. Thesis

Encouraging Residential Energy Efficiency Improvements

Sebastian Christoph Petersen

sepe@dtu.dk

Supervisor: Professor Frits Møller Andersen

Co-Supervisors: Assoc. Professor Geraldine Henningsen

Assoc. Professor Simon Bolwig

October 2, 2017

Summary in English

This cumulative PhD thesis explores the potential of a more contextual approach for energy policies targeted at private households. The five chapters of the thesis consist of an introduction, which motivates the thesis and gives a brief summary of the chapters with a joint conclusion, three paper-based chapters intended for publication, and a methodological chapter for an extension of the existing sampleSelection package in the statistics programme R. The thesis mainly contributes to the energy policy and economics literature, but it is also relevant to related disciplines, such as other social sciences, engineering, and psychology.

Energy savings in private households play an important role in achieving reduction targets for the European Union and its member states, as the residential sector constitutes about 25 percent of total final energy consumption. As only 1 percent is added to the building stock in new housing per year in the EU, the majority of savings must be realized through retrofitting and changes in consumption habits. Evidence from policy programs in the last couple of decades has been mixed and it appears that households are harder to reach than previously thought. This has led to the perception that there is an Energy Efficiency Gap, describing a situation in which households are not aware of the financial savings they could achieve by investing in more energy efficient technologies. However, more recent evidence indicates that the financial potentials at the household level may have been overstated. While there is still no consensus on whether there is an Energy Efficiency Gap, a pragmatic way forward suggests a more contextual approach that pays closer attention to barriers and promoters of energy efficiency in private homes.

Against this background, the thesis applies different methodologies to investigate the role of context in energy policies targeted at private households: a meta-analysis drawing on existing evidence in the literature, an ex-post policy analysis based on data from an energy audit programme in Southern Denmark, and a novel experimental economic

framework.

The first analysis provides indicative evidence that income, age, and education positively influence a household's propensity to invest, while household size has a negative effect. However, these findings are only partly significant due to a limited base of comparable evidence, highlighting a need for compatibility in empirical studies, and more repetition studies. The second analysis shows that changes in certain life situations increase the propensity for households to have an audit and make energy investments; specifically we find that moving and retiring make households more likely to join a free energy audit programme, while getting married and moving correlates with higher investments. This indicates opportunities to develop more efficient policy programmes that depend on a household response, by reaching out to households at a time when they are more likely to be encouraged. The final analysis investigates potential spillover effects induced by behavioural policy interventions, i.e. when an intervention aimed at one behaviour may also affect another behaviour positively or negatively. In an economic experiment we find evidence for a positive spillover induced by a social norms based intervention, but the primary contribution of this part of the thesis is the novel experimental framework developed for the purpose of a systematic analysis of spillovers.

Overall, the results of this thesis indicate that a more contextual policy approach holds promise to encourage energy efficiency improvements and savings by incorporating the context in which decisions are made, be it the socio-economic circumstances, changes in life situation, or the decision environment. Although the insights generated by the thesis are also relevant for policy design, its main contributions are to future policy research, as it highlights the need for energy policy research to be put on a more robust and evidence-based foundation, mainly through increased application of experimental methods.

Summary in Danish

Denne ph.d.- afhandling undersøger potentialet for en kontekstuel tilgang til energi politiker målrettet private husholdninger. De fem kapitler af afhandlingen består af en introduktion der motiverer afhandlingen samt danner et overblik over de inkluderede kapitler med en fælles konklusion, tre artikel-baserede kapitler tiltænkt publicering, og et metodeafsnit omhandlende udvidelser til den eksisterende sampleSelection pakke til det statistiske programme R. Afhandlingens hovedbidrag er primært til energipolitikker og –økonomisk litteratur, men dens bidrag er også relevant for beslægtede discipliner som; socialvidenskab, ingeniørkunst og psykologi.

Da husholdninger udgør 25 procent af EU's samlede energiforbrug spiller energieffektivisering af disse en vital rolle i opfyldelsen af EU's klimamål. Nybyggeri i EU udgør kun 1 procent, hvorved hovedparten af energieffektivisering skal realiseres energirenovationer samt adfærdsændringer. Resultater af incitamentsprogrammer fra de sidste årtier har været blandet, og vist at husholdninger er svære at influere end hidtil antaget. Dette har ledt til en opfattelse af et "Energy Efficiency Gap", som beskriver en situation hvori husholdninger er uvidende om de økonomiske besparelser som kan opnås via investering i energirenovering. Nyere studier indikerer at det økonomiske potentiale i husholdninger er overvurderet. Uagtet manglende konsensus om eksistensen af "Energy Efficiency Gap", kan en pragmatisk kontekstuel tilgang med fokus på barriere ved- og fortalere for energirenovering være en farbar retning.

Denne afhandling anvender forskellige metoder til at undersøge kontekst som indflydelse på succes af energiincitamenter målrettet private husholdninger. En meta-analyse baseret på eksisterende litteratur, samt data analyse af afsluttede incitamenter, inklusive data fra afsluttet energirevisionsprogram fra Region Syddanmark, samt et nyt eksperimentelt økonomisk design har dannet baggrund for resultaterne.

Den første analyse indikerer en positiv korrelation af indkomst, alder samt uddan-

nelsesniveau på en husholdnings tilbøjelighed til at udføre energirenoveringer, mens husstandstørrelsen har en negativ indflydelse. Disse indikatorer er kun delvist signifikante grundet begrænset sammenlignelig viden, hvilket understreger nødvendigheden af empiriske studier. Anden analyse viser at visse ændringer i livssituation/levebetingelser øger husholdningers villighed til energirådgivning og investering. Særligt finder vi at flytning og påbegyndelse af pension øger husholdningers villighed til at deltage i gratis energirådgivning, mens at blive gift og flytning korrelerer med husholdningers villighed til at investere. Dette indikerer muligheder for at målrette incitamentsprogrammer til specifikke husholdninger ved rettidig kontakt til husholdninger. Den afsluttende analyse undersøger potentiallet for "spillover" effekter fremkaldt af indsatser specielt rettet mod adfærd, eksempelvis ved interventioner rettet mod specifik adfærd kan afføde anden positiv og negativ adfærd. I et økonomisk eksperiment finder vi evidens for positiv "spillover" fra sociale normer. Hovedbidragene fra denne analyse skal findes i nyt eksperimentelt design udviklet til systematisk analyse af "spillover" effekter.

Overordnet set indikere resultaterne i denne afhandling at en mere kontekstuel tilgang til incitamenter, åbner for hvor husstandens aktuelle forhold som; socio-økonomi, signifikante ændringer livssituation eller andet, kan bruges som motivation for energirenovering. Afhandlingens resultater er relevante for "policy design" men dens hovedbidrag omhandler "future policy" da den fremhæver nødvendigheden mere strukturerede energi incitaments forskning med større forankring i eksperimentelle metoder.

Preface

This thesis has been submitted to the Department of Management Engineering at the Technical University of Denmark (DTU) in partial fulfilment of the requirements for a PhD degree. The work was supervised by Professor Frits Møller Andersen (DTU), Associate Professor Géraldine Henningsen (DTU), and Associate Professor Simon Bolwig (DTU). The PhD project was funded internally by the Department of Management Engineering, with support for individual papers from Innovation Fund Denmark as part of the SAVE-E and SusTrans research projects.

Kongens Lyngby, October 2017,

Sebastian Christoph Petersen

Acknowledgements

I am deeply grateful to my supervisors Frits Møller Andersen, Geraldine Henningsen, and Simon Bolwig. Throughout the more than three years of this PhD project they have given me support and guidance. Without their efforts and patience this thesis would not have been possible.

I would also like to thank Geraldine Henningsen as my co-author on two of my papers, to which she contributed greatly with her expertise and commitment. I am also grateful to her for many inspiring discussions about academic topics and beyond, in which she has both been a great mentor and a friend. I also extend my thanks to Arne Henningsen, for contributing with his tremendous knowledge and expertise in economics, econometrics, and programming in R. Last but not least among my co-authors, I thank Helene Willadsen for hours spent together learning to program experiments in Z-Tree, and being a great partner in research.

I also offer my sincere thanks to: ProjectZero and especially Peter Rathje, for sharing data and expertise that I employed in one of my papers; Marco Piovesan at University of Copenhagen for providing supervision on my experimental work and his support for my stay abroad in the US; Lars Gårn Hansen and the Institute for Food and Resource Economics at University of Copenhagen for a part-time external stay in Spring 2015; and Francesca Gino and the Negotiation, Organizations & Markets Unit at Harvard University for hosting me during an external stay in Fall 2015.

Furthermore, I would like to thank the following for discussions, feedback, and camaraderie: Alexander Tureczek, Anders Rhiger Hansen, Marco Islam, the Climate Change and Sustainable Development Group, the Systems Analysis Division, and the TRIBE seminar group.

Finally, I also thank my friends and family, who may not have contributed to the PhD project directly, but to my personal well-being.

Contents

Chapter 1

Introduction

1.1 Motivation

About 25% of final energy consumption in the European Union occurs in private households (see Figure 1.1), of which more than two thirds are due to space heating (IEA, 2008). Therefore, the European Commission and national governments place emphasis on energy savings in private homes, either through efficiency improvements or change in consumer behaviour (EC, 2013). Although today's building codes place stronger demands on the physical properties of new construction, most developed countries have a mature building stock with a small replacement rate (Lucon et al., 2014; Artola et al., 2016), and estimates suggest that at least 70% of the building stock that will be used in 2050 already exists today (Power, 2008). Furthermore, new construction often expands the living space, because the total number of apartment units increases, and units increase in size (Serrano et al., 2017). Hence, even with a lower consumption of energy per square meter, a reduction in overall consumption is not guaranteed. This points to the crucial role of household behaviour in the realization of energy savings through energy saving investments or adoption of more energy conscious consumption habits.

Energy policy is a key factor in the realization of energy savings in developed countries, as energy systems and energy markets are usually not left to free market forces. Two welfare-based arguments for state intervention in energy markets are consideration of the social dimension, in that a lot of energy services are considered to provide basic human amenities (heat in cold seasons, cooking amenities, telecommunication, etc.), and externalities that arise in the provision of energy services (Kerr et al., 2017). Hence, energy policy aimed at households is a matter of balancing the demand for affordable energy services with the externalities that arise through the provision of such services.

There is some academic debate about the potential of energy saving investments at the household level. The so-called Energy Efficiency Gap (EEG) describes the apparent discrepancy between estimated high savings from the adoption of a wide range of

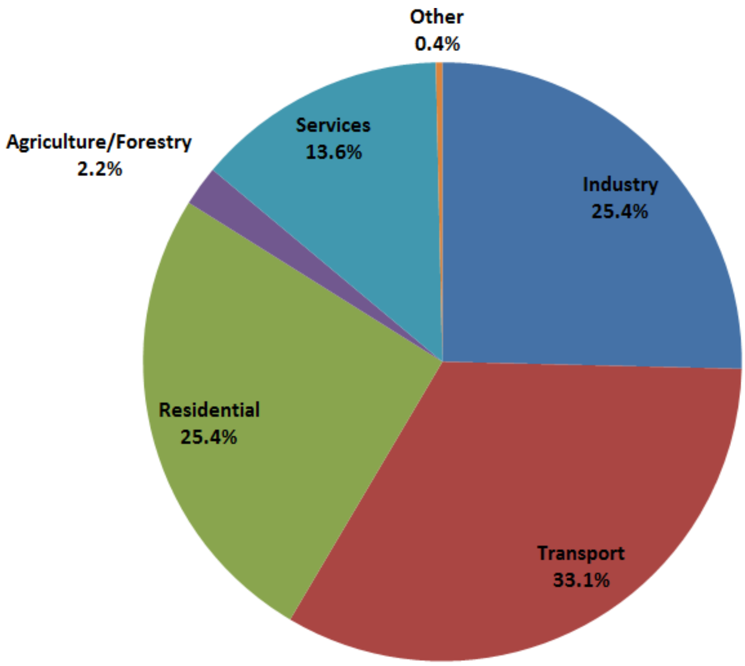


Figure 1.1: Final energy consumption in the European Union in 2015 (% of total, based on all energy sources in TJ; source: Eurostat, online data code: nrg_100a)

technologies, but low uptake by households. A typical way to measure the EEG is to calculate an implied discount rate based on household adoption decisions (given technology costs and estimated savings), and compare them to a common market discount rate, typically some measure of interest rate from capital markets (Gillingham et al., 2009). Results from previous studies suggest implied discount rates ranging between 25 percent to over 100 percent (Sanstad et al., 2006). This appears to be a win-win situation, where households would gain in terms of their welfare (a gain in the social dimension), while externalities are reduced at the same time.

Based on the assumption of the EEG, it can be argued that it is enough to inform and educate households. A fully-informed and rational household should recognize the potential offered by energy saving investments, and implement measures based on a positive net present value. Programmes that subsidize energy audits, sometimes coupled with subsidy programmes based on favourable loans or grants, have been employed frequently on such a premise (Lutzenhiser, 1993).

However, recent empirical studies have cast some doubt on the EEG narrative, and while the existence of an EEG still seems plausible for some contexts (e.g. LED lighting), it does not seem as broad a phenomenon as previously assumed (Allcott and Greenstone, 2012). A typical pattern is that a priori estimates based on engineering analyses and test cases find large potential for energy savings that are even financially beneficial for the households themselves. In ex post evaluation studies, however, savings fall short of expectations and households are generally hard to motivate into becoming active. Explanations for this discrepancy can be broken down into three categories: unrealistic assumptions or flaws in both engineering and economic models; market failures, such as imperfect information or capital market failures; and failure to account for the many behavioural hurdles and possible biases that affect households when making an investment decision (Gerarden et al., 2015).

The lack of evidence for the EEG as a pervasive phenomenon has implications for energy policies targeted at private households. In the absence of a pervasive EEG, the state cannot justify the promotion of energy investments to a vast part of the population purely based on financial savings, and risks losing credibility if it promises financial savings that do not materialize. This necessitates a re-evaluation of the objectives for energy policies targeted at households, as the argument that a large fraction of the population can save money is weakened. Policies need to be more carefully designed and evaluated with emphasis on identifying and reaching segments of the population that are affected by investment inefficiencies. Audit-based interventions in particular need to ensure that they deliver valuable information to those households that need it. Although there is still an argument for energy investments in the form of reduced externalities, this might not motivate all segments of the population.

This shift in the arguments for energy policies targeted at private households requires new approaches in energy policy research and design. Questions about which segments of the population react to energy efficiency programmes, and how households are motivated to become more energy efficient – besides pure economic rationality – are becoming crucial for policy success. Investigating these questions requires a more contextual approach in energy policy research, which emphasizes household heterogeneity as a way to identify investment potentials and receptiveness to certain policy interventions. The ultimate goal is to develop and test new policy solutions using established information- and market-based instruments, as well as insights from behavioural policy.

This thesis explores the potential for a more contextual policy design, and points out new directions for policy evaluation and research. It generates new knowledge and methodological contributions that are mainly relevant for future research-policy collaborations, e.g. in field-based research projects. The remainder of this chapter provides the analytical framework underlying the thesis, outlines the contribution of the different thesis chapters, and gives a joint conclusion and outlook on further research. This is followed by four chapters based on the academic work carried out during the PhD.

1.2 Analytical Framework

This thesis focuses on household energy consumption in a developed country context, i.e. countries that have nearly ubiquitous energy coverage and high per capita CO_2 emissions. At its core this thesis deals with household behaviour and the trade-off between people's demand for energy services and externality related issues, mainly climate change. It takes a pragmatic-descriptive approach, in contrast to a theoretic-normative approach. In other words, the thesis aims to improve knowledge in an applied policy context, and is less concerned with finding optimal solutions in a theoretical framework.

The thesis takes departure in economic theory with regard to private households and their decision making processes. Hence, to generate hypotheses, the dominant theoretical perspective regarding individual decision making is that of utility maximising selfinterested, rational actors with well-defined preferences. However, the discussion and interpretation of results will also draw on recent insights from behavioural economics, a stream of economics that relaxes assumptions made in the rational actor model. Although behavioural economics has become very popular and influential in recent years, it has not yet produced a generalized framework as parsimonious as the neo-classical rational actor theory. This is due to the fact that relaxing assumptions in the rational actor model easily makes models intractable (Wilson and Dowlatabadi, 2007).

The different chapters of the thesis are best understood in an evidence-based policy framework, which aims to understand how policy affects households, and how policy can nudge societal outcomes in a desired direction (see Figure 1.2). This framework is presented using the lens of experimentation as the most robust source for (quantitative) evidence (Burtless, 1995). Researchers use established theoretical frameworks and policy results as feedback to formulate hypotheses, and design experiments to test these. Experiments in the lab are generally good to establish qualitative results and basic behavioural patterns, but it is not certain that results generated in the abstract lab environment translate well to the real world. Therefore, the results from a lab experiment are ideally verified in the field, before the evidence should be applied to new policy programmes (Levitt and List, 2007). The evidence generated through testing starts a new cycle of feedback, which is used to design new experiments in the lab or the field.

Each chapter in this thesis can be located in a different part of Figure 1.2. Chapter 2 is a meta-analysis investigating evidence on heterogeneity in investment decisions related to socio-economic factors. The majority of the studies included in the analysis were not conducted as (field) experiments, but as an ex-post evaluation of existing policy and market data. As a meta-analysis, it draws on multiple sources of evidence and tests the accumulated knowledge. In this way it takes stock of the whole process conceptualized in Figure 1.2.

Chapter 3 mathematically derives an interval regression estimation technique which

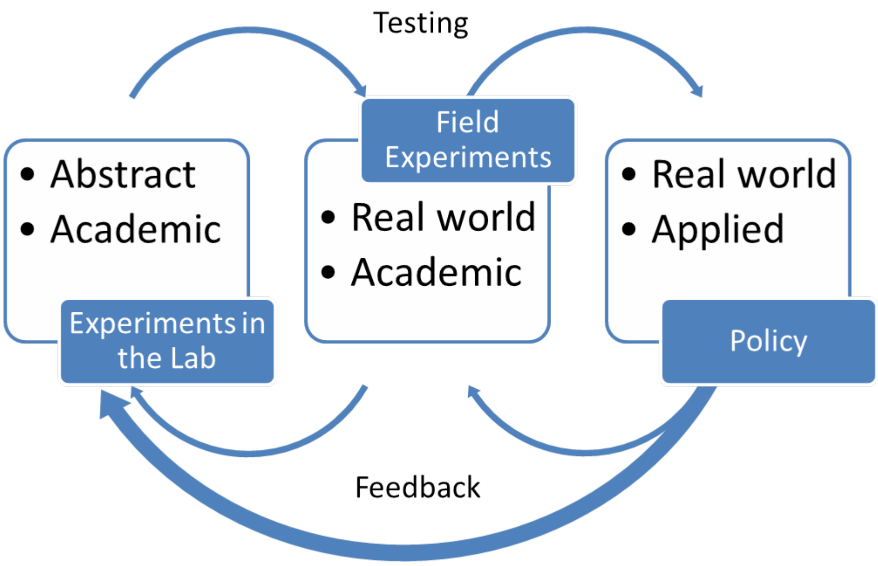


Figure 1.2: Accumulation of Knowledge in Evidence-based Policy Design

accounts for sample selection. This estimator is applied in Chapter 4, which draws on a policy programme run in Southern Denmark that offered private households free energy audits. The two chapters are most appropriately classified as generating feedback from policy in Figure 1.2, as well as making a methodological contribution on how such feedback can be interpreted.

Chapter 5 is a novel framework for investigating spillovers that may arise in the context of behavioural policies. While the work was inspired by existing policy evidence, its contribution is in the domain of developing a framework for testing hypotheses in the lab, which should then be validated in field experiments to provide policy relevant insights. It may thus be seen as a contribution in the left hand side of Figure 1.2.

1.3 Chapter Summaries

The papers contained in this thesis are quite heterogeneous in terms of their focus and the applied methodologies, with a meta-analysis drawing on empirical studies in the literature, an empirical analysis based on a unique dataset in the Danish context, and

Table 1.1: Overview of Thesis Contributions

Chapter	Title	Authors	Publication Status
2	Can Household Characteristics Consistently Explain the Heterogeneity in Households' Energy Efficiency Investments? A Meta-Analysis	Géraldine Henningsen and Sebastian Petersen	Submitted to Energy Policy in May 2017
3	Interval Regression with Sample Selection	Arne Henningsen, Sebastian Petersen, and Géraldine Henningsen	Documentation for sampleSelection package
4	Can Changes in Households' Life Situations Predict Participation in Energy Audits and Investments in Energy Savings?	Sebastian Petersen, Géraldine Henningsen, and Arne Henningsen	Working Paper
5	Detecting Behavioral Spillovers in a Real Effort Public Good Experiment	Sebastian Petersen and Helene Willadsen	Working Paper

an experimental economic study. The following sections give a brief summary of each chapter, with a focus on describing the applied methodologies for a general audience to facilitate the understanding of the later chapters, as they are written for journals in different areas of the economic landscape. Table 1.1 gives an overview of the contributions included in the thesis.

1.3.1 Can Household Characteristics Consistently Explain the Heterogeneity in Households' Energy Efficiency Investments? A Meta-Analysis

Chapter 2 conducts a meta-analysis that investigates patterns of heterogeneity in household investment behaviour related to four socio-economic characteristics: income, age, education, and household size. It draws on evidence from empirical studies in the existing literature which include one or more of these socio-economic characteristics. Identifying consistent patterns of correlation between socio-economic characteristics and investment behaviour may indicate contextual factors in household decision making that could be useful in designing policy, as it can show what characteristics make households more likely to make energy investments. The analysis takes departure in a micro-economic model to generate hypotheses about how different socio-economic characteristics may influence energy efficiency investments.

Methodology

A meta-analysis draws on multiple studies in a given field and aims to extract statistically robust findings at a meta level, and is therefore also referred to as an analysis of analyses (Glass, 1976). Meta analyses are based on the assumption that studies asking similar research questions are investigating the same underlying issues, even though they may use different methodologies and data. In the context of the accumulation of knowledge in science, it can be understood as a way to jointly evaluate evidence from multiple studies to answer a certain research question. In terms of its aim – extracting and refining existing knowledge – it is closely related to the literature review, but with a focus on quantitative results and methods.

In principle, if studies were very homogeneous in the way they have been carried out, a meta-analysis could be conducted by pooling observations and controlling for the study they have come from. However, this is rarely the case for policy studies that have not been conducted with the explicit purpose of including them in a common estimation.

Furthermore, it is often not possible to obtain the full data upon which analyses are based, which in the policy context is often due to privacy concerns. An alternative with limited information is to apply a two-stage approach, which uses the point estimates of a set of studies as observations and weight these based on their standard errors in a fixed-effects estimator (Sutton and Higgins, 2008).

Since the meta-analysis in Chapter 2 is only based on information provided in output tables of different studies in the literature, we apply a fixed-effects estimator under the assumption that estimates from the different studies are distributed:

$$\theta_i \sim N(\tau_i, v_i), \tag{1.1}$$

where i indicates the study the estimate comes from, θ_i the effect estimates, τ_i the (unknown) true effect, and v_i the sampling variance (Hedges and Vevea, 1998). To account for differences in the sampling variances due to, e.g. different sample size and data, the fixed-effects estimation weights each study estimate by the inverse of its sampling variance:

$$\bar{\theta} = \frac{\sum w_i \theta_i}{\sum w_i}, \tag{1.2}$$

where $\bar{\theta}$ is the weighted average of the effect that is estimated, and w_i weights applied to the study estimates calculated as $\frac{1}{v_i}$(Viechtbauer, 2010).

The analysis proved to be challenging, possibly due to differences in the set of variables and the estimation techniques used in our sample of studies, but also because of a limited body of evidence we were able to draw on. These challenges are further discussed in Chapter 2.

Results

Findings from the meta-analysis provide indications for positive correlation between income, age, and education with energy investments, and negative correlation between household size and energy investments. However, these results are only partially significant, as the body of comparable studies is too small. The fact that the meta-study was hard to carry out and only provides indicative results is an important finding in itself, as it documents a severe limitation in the potential to assess accumulated knowledge in this area of energy policy. It stands to reason that the indicative evidence from the meta-analysis warrants more attention to heterogeneity in future studies, but it also calls for more repetition studies, to ensure a broader empirical basis. In the context of project applications this raises the of question how the need for replication studies can be communicated to funding agencies and governments, as funding agencies often look for novelty that may sacrifice comparability. One answer could be an increased use of randomized controlled trials that allow for the testing of different treatments. While one of the most robust techniques available in the scientific toolbox, randomized controlled trials will also allow for testing new and innovative ideas in conjunction with more conventional interventions as a baseline for comparison.

1.3.2 Interval Regression with Sample Selection

Chapter 3 is a methodological chapter that formally develops an interval regression estimator with sample selection. The approach used to develop the estimator is closely related to the selection model formulated by Heckman (1979), which is often applied in policy analysis. Heckman (1979)'s approach, however, only applies to cases in which a linear outcome variable of interest is estimated by OLS and not when the outcome variable of interest is non-linear. This chapter mathematically derives an econometric approach for an interval censored outcome variable estimated by joint maximum likelihood under the assumption of a bivariate normal distribution of the error terms in the selection and outcome stage.

Methodology

Sample selection is a common problem for all kinds of policy evaluations. It is likely to be an issue for policy programmes in which individuals in the target population have to actively join the policy, i.e. participation is voluntary. Self-selection into policy treatment is often related to characteristics that also influence the final outcome of interest, which leads to bias in econometric estimations of the outcome (Reed, 2000). While it is still possible to evaluate the policy outcome by comparing with a control group, this limits the generalizability of findings, as they only capture the effect on the biased sample. Additionally, policy data is often subject to limited observability, i.e. it often only includes data on the treated population, which makes a direct comparison between treated and untreated individuals impossible (Meng and Schmidt, 1985).

These issues can be addressed ex-post by estimating a sample selection model specification. In this class of models, the researcher attempts to separate the decision to join the policy from the outcome of interest. However, this approach requires additional data, especially on the untreated population. Furthermore, to separate the two decisions, an exclusion restriction is needed to ensure identification. An exclusion restriction is an instrumental variable that predicts the selection into policy, but not the actual outcome decision. This is a major obstacle for most ex-post policy evaluations, and although there are a large number of studies that apply sample selection models, it is not always certain if the inference is valid, due to doubts about the chosen instrument (Smith and Sweetman, 2016).

Randomized controlled trials are a way to avoid some self-selection issues by, for example, randomizing treatment among the interested part of the population. However, this does not address issues of bias in the sample, as households that are willing to join the policy have already self-selected. The crucial advantage of a randomization in treatment assignment is that it provides a perfect instrument that can be used as exclusion restriction. By design, being selected into treatment is perfectly correlated with being subject to the policy treatment, but not with the final outcome of interest. Therefore, even when a study is well-designed, sample selection models play an important role in policy evaluation with econometric methods (Smith and Sweetman, 2016).

Results

The main result of Chapter 3 is an econometric estimator suitable to provide unbiased estimates for an interval coded outcome variable that accounts for sample selection. The mathematical derivations of the estimator and the gradients in this chapter were used to program the estimator as an extension of the existing R package "sampleSelection" (Toomet and Henningsen, 2008) in the statistical software R (R Core Team, 2017). The estimator is applied to the case of household energy investments in Petersen et al. (2017), which is presented in Chapter 4.

1.3.3 Can Changes in Households' Life Situations Predict Participation in Energy Audits and Investments in Energy Savings?

Chapter 4 investigates changes in the life situation of private home owners as potential drivers of the decision to have an energy audit and make an investment in their home. The chapter is based on a comprehensive data set on participants of a free energy audit programme and a representative control group of home owners from Sønderborg municipality in Southern Denmark, it investigates the impact of four life situational factors – marriage, retirement, unemployment, and relocation into a new home – on a household's decision to have a free energy audit and make an energy investment.

Methodology

Sample selection is a common issue in ex-post policy evaluation, as discussed in the previous section. This becomes apparent when looking at the decision process for households participating in the ProjectZero energy audit programme, which is analysed in Chapter 4 (see Figure 1.3). Households decide to have an energy audit in the first stage, and then make the decision if and how much to invest in a second stage. However, the investment decision is only observed for households that had an energy audit. Selection issues may arise, when there are factors that determine both the decision to join the policy and the decision to make an energy investment. This results in correlation in the error terms that, if unaccounted for, leads to a biased estimator (Smith and Sweetman, 2016).

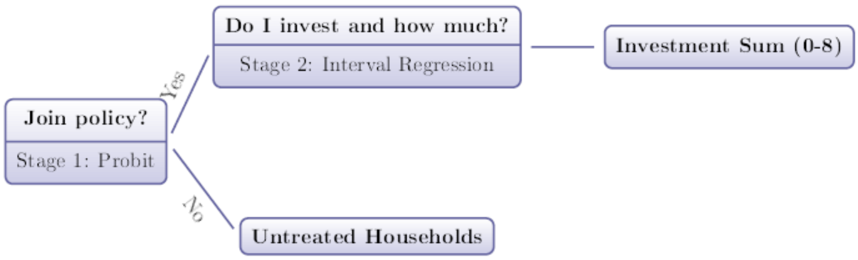


Figure 1.3: Household Decision Process for Participation in the ProjectZero Energy Audit Programme

Although issues with sample selection complicate the estimation of the second stage, it does not impede the analysis of the first stage decision if this is of interest. Probit and logit are common estimators for estimating dichotomous choices, e.g. the decision to join a policy programme. In our case we use a probit estimation, which is a bit more common in economic applications where self-selection is involved. The probit estimator is based on the assumption that the outcome variable, which only takes on values of zero or one is a representation of a latent variable that reflects the propensity of joining the policy (Wooldridge, 2010). In some cases this propensity may have an explicit meaning, e.g. in economics it may conceptually be considered the utility a household derives from joining the treatment group.

Sample selection only becomes an issue when estimating the investment into energy savings in the second stage. As described in Section 1.3.2 this issue is commonly addressed through joint estimation. In the present case the issue is complicated by the outcome of interest coded in intervals. Hence, an appropriate estimator has to address: (1) the self-selection of households into the audit programme and (2) the censored outcome variable for investments, which is coded in intervals. This estimator is described in more detail in Section 1.3.2 and developed in Chapter 3.

Results

Results suggest that relocation and entering retirement are significant predictors of joining the audit programme, while including the current life situation variables does not significantly improve the fit of our estimation model. This confirms that the change in these life situational factors is driving the effect, and not the difference in the life situation by itself. This insight supports the notion that timing is important for reaching private home owners with energy efficiency programmes.

The joint estimation results show that getting married and relocation are strongly associated with the amount that is invested into energy saving measures. This finding, however, is preliminary at the moment, as there are issues with the data that still need to be addressed, before publication of the paper. While the findings are suitable to inspire policies designed to reach out to people in these specific life circumstances (e.g. offering audit programmes that target households that recently moved into a new house), it is highly desirable to implement this in a randomized field experiment. This would allow for a better evaluation of the specific new policy, but also provide a better general understanding, as issues in ex-post evaluation, which could only be partly addressed in this thesis, could be alleviated with an appropriate experimental design.

1.3.4 Detecting Behavioral Spillovers in a Real Effort Public Good Experiment

Chapter 5 develops a novel experimental framework to investigate the conditions determining occurrence and direction of spillovers induced by behavioural policies. The experiment applies a Social Norm intervention, in which the communication of other people's behaviour is used to influence decision making. Among other areas of application, Social Norm interventions are increasingly used as a policy tool to influence people's everyday consumption behaviour. They promise a significant demand response at low financial costs. Often overlooked in the application of such behavioural interventions, however, is that they can have unintended spillover effects. The experimental framework developed in this chapter recreates a decision context in the lab, in which spillover effects of behavioural interventions can be systematically analysed.

Methodology

Experimental economics has become quite popular in recent years, and its rise in use is closely related with the ascent of behavioural economics, which is a sub-field of economics strongly influenced by psychology. Applications of behavioural economics generally relax assumptions commonly made in neo-classical economics, mainly the assumption of perfectly rational actors. As such it moves towards greater behavioural realism, and adoption of experimental methods that allow for more rigorous testing of hypotheses generated by theory (Smith, 1989). While economic experiments are also run in the field, the focus here is on lab experiments.

One of the main advantages of running experiments in the laboratory is the tight control the experimenter has over the decision environment in which actors make decisions. Furthermore, accurate description of the control settings allows for replication of experiments. This gives researchers the ability to validate their own and others' results and increases the explanatory power by running multiple sessions of the same experiment. However, the application of experimental economics is not without critique. The tightly controlled environment often leads to an abstract decision making context, which participants do not find in the real world. This raises the question if participants behave the same way as they do in the real world. Another issue often voiced is that the populations from which experimental economists draw their samples are not representative of the rest of the population. It is common to use university students, because they are available in great number, always offer fresh subjects as new students enroll every year, and are easy to recruit with a low to medium hourly compensation.

It is within the nature of economic experiments that these points of critique can be tested. If there is worry that an experimental outcome is biased, because the subject pool consists of students, the experiment can be replicated with subjects from the general population (as done in e.g. Exadaktylos et al. 2013). The question if results found in an abstract lab environment hold up in the real world is commonly tested by running field experiments to confirm results, before concrete policy recommendations are made. Hence, lab and field experiments should be seen as complementary tools (Levitt and List, 2007).

Economic experiments can take many forms, but nowadays the majority are run in computer-based laboratory environments or online. A typical computer-based laboratory experiment invites a group of participants, which are assigned a computer running a software environment, in which decisions are made. Within the experiment subject decisions are incentivized, i.e. the decisions they make result in monetary payments, depending on the decisions they make. In this way, a subject's decision in the laboratory mimics real-world consequences of decisions and aims to elicit their true preferences (Smith, 1976).

The experimental framework presented in Chapter 5 is a variation of a public good game, which captures problems of collective action described early on in Olsen (1965). In the standard public good game subjects receive an endowment, which they can either keep for themselves, or contribute into a common fund shared with a group of people in the same experiment. Each unit of the endowment kept for themselves is put into their private account, which is converted into real money at the end of the experiment. Each unit of the endowment allocated to the common fund provides a pay-off smaller than one, but to everyone in the group. The total group earnings from a unit in the group fund is usually larger than one, which results in a pay-off function for an individual i as follows:

$$\pi_i = E - c_i + 0.5 \cdot c_i + 0.5 \sum_{j=1}^{3} c_j, \tag{1.3}$$

in which π_i is the earnings (in points that are converted into real money at the end of the experiment), E the endowment, c_i the points invested in the public good, and c_j the amount of points invested into the public good by other player in the game (in this game a group is composed of 4 players). The best course of the group would be for everyone to place their endowment in the common fund, but the individual pay-off is maximized by keeping the endowment. Hence, the game theoretical Nash-equilibrium is for everybody to keep their endowment for themselves. In contrast to this prediction, studies commonly find that subjects do invest a good part of their endowment, though far from the social optimum of full cooperation. This cooperation usually decreases in repeated games, highlighting the conditional nature of contributing to the public good (Ledyard, 1995).

In the public game variation we develop in Chapter 5 we investigate behavioural spillovers. This is inspired by findings from policy and experiments in the field about possible side-effects of behavioural policies. Though, the spillover term describes a somewhat new stream of investigation in experimental economics, it is closely related to the rebound effect in energy economics. Truelove et al. (2014) defines a spillover as "an effect of an intervention on subsequent behaviours not targeted by the intervention" (p. 128), while the rebound effect goes back to Jevons (1866)'s observation that increases in the efficiency in the utilization of coal lead to an increased demand of coal-related services, increasing demand for coal. Therefore, the rebound effect is a special case of a spillover, in which inducing an efficiency improvement in the provision of a good or service changes subsequent consumption behaviour of the same good or service.

In the experiment, subjects complete real effort tasks in a public good environment, in what we call a Real Effort Public Good game. Earnings are generated through a private task which benefits the individual only, and two public tasks, which benefit a group. A social reference intervention is applied in the treatment group in which participants receive information about session-level contributions to only one public good. The experimental framework aims to create an abstract version of a decision environment in which subjects make low-stake decisions about contributing to two different public goods. In the version we tested in the lab the two public goods are only separated through their appearance in the game and in the instructions, while the underlying incentives are identical. This provides a baseline, from which it is possible to design additional treatments that vary and mimic different decision environments in future extensions of this work.

Results

In the experiment we find evidence of a positive spillover between two tasks that were only differentiated through framing, a setup which we deem favourable for the occurrence of spillovers. Nevertheless, we see merit in this result as a baseline from which to generate a systematic analysis framework by varying, e.g.: the similarity of the motive of the two public goods; the public good tasks; group composition in the public goods; or the type of intervention that is applied to the treated public good. Such experiments have great potential to increase our knowledge of the occurrence of currently unaccounted spillovers. While we advise caution when translating lab results into advice for policy makers, our results support general notions from the literature that spillover effects should receive more attention when designing and evaluating policy. Our result of a positive spillover in closely related tasks with a similar motive may open up ways to design policy that generates positive spillovers, by making these linkages between activities salient to consumers.

1.4 Conclusion and Further Research

This thesis presents a collection of studies exploring new approaches to get a better understanding of household decision making in the context of energy policy. It is founded in a contextual approach to policy design influenced by the recent influx of behavioural economic theory and application of experimental methods. While the results are partly applicable to the political discourse, the main contributions are in terms of opening up new pathways in future evidence-based policy research.

Overall, the thesis contributes to the field of energy policy targeted at households, by applying new methodologies and exploring previously neglected aspects of household decision making. A general point the results of this thesis emphasize is that energy policy needs to be put on a more robust evidence-based footing, mainly through increased application of experimental methods. Three broader insights have been produced by this thesis: (1) the current empirical evidence in the literature is limited in terms of the comparability and robustness of studies, making it hard to generalize insights beyond the context of single studies; (2) ex-post evaluations of policy data (i.e. policy-based evidence) are inherently limited in their ability to generate generalizable evidence, due to confounding factors, such as selection effects, and issues with identifying proper instruments to use in sophisticated econometric models to address these issues; (3) a contextual policy approach holds promise to improve energy policy and induce energy efficiency improvements, e.g. by reaching out to households that are in a phase of life where they are more susceptible to policy, or a better understanding of how different consumption domains are related.

All in all, this thesis does not provide direct policy prescriptions on how households can be induced to save energy, but it points out questions that are worthwhile to ask, and how to investigate them. Specifically, it highlights the importance of understanding the different factors that play a role in making people more likely to participate in policy programmes, and when policy programmes might have unintended side-effects. How this information can be used to achieve additional savings is an empirical question that remains to be answered through future testing in the field.

Bibliography

Allcott, H., Greenstone, M., 2012. Is there an energy efficiency gap? Journal of Economic Perspectives 26 (1), 3–28.

Artola, I., Rademaekers, K., Williams, R., Yearwood, J., 2016. Boosting building renovation: What potential and value for europe. Tech. rep., European Parliament's Committee on Industry, Research and Energy.

Burtless, G., 1995. The case for randomized field trials in economic and policy research. The Journal of Economic Perspectives 9 (2), 63–84.

EC, 2013. Guidance for national energy efficiency action plans. Tech. rep., European Commission.

Exadaktylos, F., Espin, A. M., Branas-Garza, P., 2013. Experimental subjects are not different. Scientific Reports 3.

Gerarden, T., Newell, R. G., Stavins, R. N., 2015. Assessing the energy-efficiency gap. Tech. rep., Cambridge, Mass.: Harvard Environmental Economics Program, January 2015.

Gillingham, K., Newell, R., Palmer, K., 2009. Energy efficiency economics and policy. Tech. rep., Resources for the Future.

Glass, G. V., 1976. Primary, secondary and meta-analysis of research. Educational Researcher 5 (10), 3–8.

Heckman, J. J., 1979. Sample selection bias as a specification error. Econometrica: Journal of the econometric society 47 (1), 153–161.

Hedges, L. V., Vevea, J. L., 1998. Fixed- and random-effects models in meta-analysis. Psychological Methods 3 (4), 486 – 504.

IEA, 2008. Promoting energy efficiency investments: Case studies in the residential sectors. Tech. rep., International Energy Agency.

Jevons, W. S., 1866. The Coal Question: An Inquiry Concerning the Progress of the Nation, and the Probably Exhaustion of our Coal Mines. Macmillian and Co., London.

Kerr, N., Gouldson, A., Barrett, J., 2017. The rationale for energy efficiency policy: Assessing the recognition of the multiple benefits of energy efficiency retrofit policy. Energy Policy 106, 212 – 221.

Ledyard, J. O., 1995. Public goods: A survey of experimental research. In: Kagel, J. H., Roth, A. E. (Eds.), The handbook of experimental economics. Princeton university press Princeton, NJ.

Levitt, S. D., List, J. A., 2007. What do laboratory experiments measuring social preferences reveal about the real world? The Journal of Economic Perspectives 21 (2), 153 – 174.

Lucon, O., Ürge Vorsatz, D., Ahmed, A. Z., Akbari, H., Bertoldi, P., Cabeza, L. F., Eyre, N., Gadgil, A., Harvey, L. D. D., Jiang, Y., Liphoto, E., Mirasgedis, S., Murakami, S., Parikh, J., Pyke, C., Vilariño, M. V., 2014. Buildings. In: Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., Kriemann, B., Savolainen, J., Schlömer, S., von Stechow, C., Zwickel, T., Minx, J. (Eds.), Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Lutzenhiser, L., 1993. Social and behavioral aspects of energy use. Annual Review of Energy and the Environment 18, 247–289.

Meng, C.-L., Schmidt, P., 1985. On the cost of partial observability in the bivariate probit model. International Economic Review 26 (1), 71 – 85.

Olsen, M., 1965. The logic of collective action: public goods and the theory of groups. Cambridge: Harvard University Press.

Petersen, S., Henningsen, G., Henningsen, A., 2017. The impact of situational factors on the decision to participate in energy efficiency programs, unpublished Manuscript. Department of Management Engineering, Technical University of Denmark.

Power, A., 2008. Does demolition or refurbishment of old and inefficient homes help to increase our environmental, social and economic viability? Energy Policy 36, 4487–4501.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL `http://www.R-project.org/`

Reed, W., 2000. A unified statistical model of conflict onset and escalation. American Journal of Political Science 44 (1), 84 – 93.

Sanstad, A. H., Hanemann, W.-M., Aufhammer, M., 2006. Chapter 6: End-use energy efficiency in a ”post-carbon” california economy: Policy issues and research frontiers. In: Hanemann, W. M., Farrell, A. E. (Eds.), Managing Greenhouse Gas Emissions in California. Climate Change Center at UC Berkley.

Serrano, S., Ürge Vorsatz, D., Barreneche, C., Palacios, A., Cabeza, L. F., 2017. Heating and cooling energy trends and drivers in europe. Energy 119, 425 – 434.

Smith, J., Sweetman, A., 2016. Viewpoint: Estimating the causal effects of policies and programs. Canadian Journal of Economics 49 (3), 871 – 905.

Smith, V., 1976. Experimental economics: Induced value theory. The American Economic Review 66 (2), 274 – 279.

Smith, V. L., 1989. Theory, experiment and economics. Journal of Economic Perspectives 3 (1), 151–169.

Sutton, A. J., Higgins, J. P., 2008. Recent developments in meta-analysis. Statistics in Medicine 27, 625–650.

Toomet, O., Henningsen, A., 2008. Sample selection models in r: Package sampleselection. Journal of Statistical Software 27 (7).

Truelove, H. B., Carrico, A. R., Weber, E. U., Raimi, K. T., Vandenbergh, M. P., 2014. Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. Global Environmental Change 29, 127–138.

Viechtbauer, W., 2010. Conducting meta-analyses in r with the metafor package. Journal of Statistical Software 36 (3).

Wilson, C., Dowlatabadi, H., 2007. Models of decision making and residential energy use. Annu. Rev. Environ. Resour. 32, 169–203.

Wooldridge, J. M., 2010. Introductory Econometrics: A Modern Approach, 4th Edition. Nelson Education.

Chapter 2

Can Household Characteristics Consistently Explain the Heterogeneity in Households' Energy Efficiency Investments? A Meta-Analysis

Géraldine Henningsen and Sebastian Petersen

Abstract

The Energy-Efficiency-Gap (EEG) often rationalises broadly-targeted policies for energy-efficiency measures in residential buildings. However, the idea of a pervasive EEG finds little support in the economic literature. One implication of this finding is that household heterogeneity is more consequential than previously assumed. Hence, knowledge on what kinds of household are likely to engage in energy efficiency investments becomes crucial for formulating effective policies. Although 'policy targeting' and 'household heterogeneity' have been extensively covered in the energy-policy literature, a meta-analysis of how observable characteristics correlate with energy-efficiency investments has not been conducted to date. This article aims at reviewing the empirical evidence in the energy-policy literature in order to assess the existence of behavioural patterns connected with the following four frequently included and easily observable household characteristics: income, age, education, and household size. Additionally, by means of a standard investment model, we compare the theoretical and empirical impact of these four factors on households' investment decisions. We find consistently positive effects for income, but more ambiguous positive effects for age and education, as well as ambiguous negative effects for household size. Our results cannot confirm an influence of the decision domain on the effect of these four household characteristics.

2.1 Introduction

Policies to encourage energy efficiency investments in residential housing are common in many developed countries (e.g., *Weatherization Assistance Program* in the United States, *Green Deal* in the UK, *Energieeffizient Sanieren* in Germany or *Bedre Bolig* in Denmark). Such policies are often broadly targeted, i.e., most households are potentially encouraged through information provision and subsidies. Two common arguments expressed in favour of these policy campaigns are reduced negative externalities from mitigating CO_2

emissions, and untapped financial gains for households, the so-called Energy Efficiency Gap (EEG). Hence, residential energy efficiency policies are often depicted as a win-win for both private households and society.

A review of the EEG by Allcott and Greenstone (2012), however, shows that the magnitude of the EEG is smaller than has been suggested by engineering studies that often motivate these policies; also, these policies often fail to create significant future welfare gains under the assumption of reasonable discount rates. These findings have been confirmed by recent ex-post evaluations of residential energy efficiency programmes (Davis et al., 2014; Fowlie et al., 2015; Zivin and Novan, 2015).

To address these findings, Allcott and Greenstone (2012) suggest designing policies that consider the heterogeneity in energy inefficiencies across households; a point which has also been emphasised in other studies (e.g., Stern, 1992; Allcott et al., 2014; Gillingham and Palmer, 2013). The rationale behind this suggestion is straightforward: if only a subset of the population stands to gain from a policy, targeting this part of the population, for example, through target-oriented information (Fogg, 2003), will be more effective and eventually more efficient than targeting the whole population. However, despite calls for the increased utilisation of household heterogeneity in policy design, it is unclear whether systematic and exploitable patterns in household heterogeneity exist. Secondly, if patterns do exist, are they connected to variables that are accessible enough for energy modellers and policy makers who want to incorporate household heterogeneity into their analyses and policy designs?

Empirical evidence for heterogeneity across households has been documented in various studies (e.g., Hausman, 1979; Newell and Siikamäki, 2013), but a systematic meta-analysis with the purpose of establishing systematic and exploitable patterns connected to accessible household characteristics is still missing to date. We address this gap and contribute to the energy policy discussion on households' energy efficiency investment decisions by analysing and evaluating the empirical evidence.

Allcott and Greenstone (2012) identify unobserved costs and benefits of an energy efficiency measure as a major source of the divergence between economic and engineering studies. Consequently, if these costs and benefits are mostly unobservable, the question arises as to whether they can be proxied by observable variables, e.g., socio-economic variables. Train (1985) investigates the impact of socio-economic variables on (implied) discount rates in the adoption of energy efficiency measures. The study identifies a positive influence of income on the propensity of adoption, but produces mixed results for age and education due to limited evidence. The present study takes up this discussion and is, to the best of our knowledge, the first meta-analysis to investigate systematic patterns in household characteristics that have shown to influence households' decisions to invest in energy efficiency.

In order to screen the results from 19 empirical and experimental studies, we develop a systematic approach to evaluate and compare the empirical results with respect to the following four frequently included and easily observable socio-economic variables: income, age, education, and household size. We investigate the existence of consistent patterns across these four variables with respect to a household's propensity to conduct energy efficiency investments. Furthermore, we apply a classic microeconomic approach to evaluate the theoretical connection between observable socio-economic variables and the latent variables that influence a household's investment decision. This enables us to compare empirical evidence with the assumptions from classic micro-economic theory.

The article is structured as follows: section two describes the methodology of our literature search; section three outlines the theoretical economic model and derives the hypotheses; section four presents the empirical findings; section five discusses the empirical findings and compares them to our hypotheses; and section six concludes.

2.2 Literature Search

To identify the relevant literature, we screened the literature for empirical studies that analysed the determinants of households' energy efficiency investment decisions both under market conditions and as a reaction to policies. For each identified study, we conducted a forward and backward citation search using the Google Scholar, Scopus, EconStor, and EconPapers databases.

We focused our search on three broad categories: real market behaviour, stated preference studies, and policy evaluations. In order to increase comparability, we concentrated on studies conducted in industrialised countries. We screened these studies for observable household characteristics in order to determine the most frequently used variables. Studies use a multitude of different household characteristics, although the most frequently included are income, age, education, and household size. Additional frequently applied characteristics are race and children living in the household, while variables such as home ownership, household debt, employment status, and gender are infrequently used variables. Environmental attitude and political affiliation are often included, especially in studies from the political science and psychological literature. However, as these characteristics are not easily observable for energy modellers and policy makers, but require extensive surveys, we abstract from them in our study. Given these results, we focus on the following four most frequently used variables: income, age, education, and household size.

From the potentially relevant literature, we selected articles fulfilling the following criteria:

- published in a peer-reviewed journal or in an established working paper series,
- presenting empirical results of the determinants of private households' engagement in energy efficiency,
- containing at least one of the four selected household characteristics as a covariate.

Broadly defined, the studies included in our analysis present empirical results that allow inference about the propensity of households to invest in energy-efficiency.

2.3 Model and Hypotheses Formulation

To set a theoretical framework for the analysis of the empirical results, we start by defining a simple investment model such as in Allcott and Greenstone (2015). Households can improve energy efficiency by investing into energy-efficient portable and non-portable assets such as energy-efficient appliances or weatherisation improvements.

Let $\theta_{ij} = (e_{ij}, \xi_{ij}, c_{ij}, \mathscr{T}_{ij})'$ be a vector, where $i = 1, \ldots, \mathscr{I}$ is the household index, and $j \in \mathscr{J}_i$ indicates a specific energy efficiency measure from the set of all feasible measures, $\mathscr{J}_i$, available to household i. e_{ij} is the expected monetary Present Day Value (PDV) of the energy savings of the investment; ξ_{ij} is the expected PDV of the monetised non-monetary benefits of the investment (e.g., better indoor climate); c_{ij} are the monetary costs of the investment and $\mathscr{T}_{ij}$ are its expected monetised non-monetary costs (e.g., disturbances through construction work).

We set up the following expected utility function:

$$E(U(y_i, e_{i0}, \mathscr{B}_{i0}, \Theta_i, \mathbf{I}_i)) = y_i - e_{i0} + \mathscr{B}_{i0} + \sum_{j \in \mathscr{J}_i} I_{ij}(e_{ij} + \xi_{ij} - c_{ij} - \mathscr{T}_{ij}), \tag{2.1}$$

where y_i is household income, a proxy for wealth[1]; e_{i0} is the PDV of the expenditures of the future baseline energy consumption without improvement; $\mathscr{B}_{i0}$ are the monetised non-monetary features of the status quo; $\Theta_i = \{\theta_{ij}; j \in \mathscr{J}_i\}$ is the set of the costs and benefits of all energy efficiency measures available to household i; I_{ij} is a dummy variable indicating whether household i adopts the $j$the energy efficiency measure, and $\mathbf{I}_i = \{I_{ij}; j \in \mathscr{J}_i\}$.[2]

[1] Although income may be relevant by itself, we would expect overall wealth to be more relevant. However, because data on wealth is seldom included in empirical studies, we omit it in our model.

[2] We presume that all energy efficiency measures in set $\mathscr{J}_i$ are independent of each other. In consequence,

As all of these variables, except for y_i and I_{ij}, are usually unobserved latent variables, we suggest expressing them through functions that depend on four observable household characteristics: income, y_i; age, a_i; education, d_i, and household size, z_i:

$$\begin{aligned} E(U(y_i, e_{i0}, \mathscr{B}_{i0}, \theta_{i1}, \ldots, \theta_{iJ})) = & \\ & y_i - e_{i0}(y_i, a_i, d_i, z_i) + \mathscr{B}_{i0}(y_i, a_i, d_i, z_i) \\ & + \sum_{j \in \mathscr{J}_i} I_{ij}(e_{ij}(y_i, a_i, d_i, z_i) + \xi_{ij}(y_i, a_i, d_i, z_i) - c_{ij}(y_i, a_i, d_i, z_i) - \mathscr{T}_{ij}(y_i, a_i, d_i, z_i)) \end{aligned} \tag{2.2}$$

Equation (2.3) shows the effect of an investment j on the expected utility of household i:

$$\lambda_{ij}(\cdot) = e_{ij}(y_i, a_i, d_i, z_i) + \xi_{ij}(y_i, a_i, d_i, z_i) - c_{ij}(y_i, a_i, d_i, z_i) - \mathscr{T}_{ij}(y_i, a_i, d_i, z_i), \tag{2.3}$$

where $\lambda_{ij} = E(U(\cdot)\,|\,I_{ij} = 1) - E(U(\cdot)\,|\,I_{ij} = 0)$. Equation (2.3) shows that λ_{ij} will be positive, negative or neutral depending on the Net Present Value (NPV) of the investment. We argue that the NPV of an investment depends on latent variables that are functions of heterogeneous household characteristics. Hence, in the following, we examine the effects for each of the four household attributes on the NPV of an investment, λ_{ij}, keeping as close as possible to classic micro-economic theory. Hypotheses derived from these effects will serve as a benchmark in the evaluation of the empirical results in section 2.5.

2.3.1 Income

$$\frac{\partial \lambda_{ij}}{\partial y_i} = \overbrace{\frac{\partial e_{ij}}{\partial y_i}}^{(+)} + \overbrace{\frac{\partial \xi_{ij}}{\partial y_i}}^{(-)} - \overbrace{\frac{\partial c_{ij}}{\partial y_i}}^{(-)} - \overbrace{\frac{\partial \mathscr{T}_{ij}}{\partial y_i}}^{(+)} \tag{2.4}$$

some energy efficiency measures are package solutions if their conservation effect is dependent on the combination of several single conservation measures, e.g., a household with two possible investments A and B has three option: 'A', 'B', or 'A and B'.

The direction of the overall effect of income, y_i, on λ_{ij} depends on the magnitude of the single effects in equation (2.4):

1. We presume that income, y_i, has a positive effect on the energy conservation potential of an investment j. An energy efficiency investment will increase the efficiency of the provision of corresponding energy services and, hence, reduce the costs of these services. Assuming that energy services are a normal good, a reduction in their costs will translate into an increase in demand (rebound effect) (Fouquet and Pearson, 2012). However, we argue that wealthier households have a smaller price elasticity of the demand for energy services (see Appendix 2.A) as their initial demand for energy services is already high. In consequence, wealthier households will be less prone to the rebound effect and the net-savings effect, e_{ij}, of an energy efficiency measure j that increases energy efficiency by a certain factor will be larger (in absolute terms) in wealthier households than in poorer households.

2. An increase in the consumption of energy services implies an increase in non-monetary benefits. It then follows from the previous point that poorer household will benefit more from the investment, i.e., have greater non-monetary benefits due to a larger rebound effect.

3. We assume that the monetary costs, c_{ij}, will be lower for wealthier households. Although pure purchasing costs are expected to be the same for all households, the associated capital costs might vary considerably. Wealthier households have better access to capital and lower interest rates because they own more assets as collateral. This effect should be more pronounced the larger the investment sum associated with the energy efficiency measure.

4. Non-monetary costs, $\mathscr{T}_{ij}$, are higher for high income households because their leisure time is more valuable to them (in monetary terms). This is based on the standard assumption that individuals work until their marginal benefit of leisure is

equal to their marginal benefit of work. Thus, the higher the household's income from labour, the higher the costs linked to lost leisure time resulting from energy efficiency measures.[3]

The overall effect of income, to a large extent, will depend on the energy efficiency measure at hand. Irrespective of the income level, households will profit from the investment: either through monetary savings or through an increase in the non-monetary benefits. Hence, the impact of income is most pronounced on the cost side. A higher income will be most advantageous if the investment is capital-intensive, while on the other hand, it will create higher non-monetary costs for time-intensive measures. Based on these considerations, we formulate the following testable hypothesis for income, y_{ij}:

Hypothesis 1 *A household's propensity to invest will increase with income in tact with the capital intensity of the investment.*

2.3.2 Age

$$\frac{\partial \lambda_{ij}}{\partial a_i} = \overbrace{\frac{\partial e_{ij}}{\partial a_i}}^{(-)} + \overbrace{\frac{\partial \xi_{ij}}{\partial a_i}}^{(-)} - \overbrace{\frac{\partial c_{ij}}{\partial a_i}}^{(+/-)} - \overbrace{\frac{\partial \mathscr{T}_{ij}}{\partial a_i}}^{(-)} \tag{2.5}$$

The overall effect of age is ambiguous and complicated by the fact that many effects will be non-linear across a life span:

1. Because of a lower expectation in healthy life years, elder households might have a shorter time horizon to accumulate the benefits of an energy efficiency investment. Hence, the PDV of the energy efficiency measure, e_{ij}, and the PDV of the non-monetary benefits, ξ_{ij}, decrease with age, implying that age has a negative effect on the valuation of the benefits if we assume a common discount rate across all households.

[3] However, these considerations are no longer valid if the main source of income is not or no longer from labour, e.g., in the case of retirees or households with a large income from capital investments.

2. The monetary costs of an investment, c_{ij}, follow a more complicated pattern across age groups. While overall age seems to decrease credit-constraints (Jappelli, 1990; Lyons, 2003), capital costs eventually sharply increase after a certain age (lenders evaluate the risk of giving loans to elderly households as high). This effect might be strongest for energy efficiency measures that require larger investments.

3. Finally, as age increases, the share of labour income from total income will decrease considerably for most households (Aaronson et al., 2014). Hence, elder households will *ceteris paribus* (e.g., for a given total income) have a lower marginal income from labour, i.e. they will have lower opportunity costs of leisure time. Consequently, the valuation of leisure time will be lower and, hence, elder households will consider a loss of leisure time as less costly.

Overall age appears to reduce the benefits, but also decreases the costs of an energy efficiency investment. We expect a more pronounced negative effect of age on the benefits of energy efficiency measures with long amortisation periods. On the other hand, with the exception of the elderly, age has a positive effect on both monetary and non-monetary costs. Based on these considerations, we formulate the following hypothesis for age, a_i:

Hypothesis 2 *The effect of age on the likelihood of a household investing in energy efficiency measures is ambiguous for capital-intensive investments with long amortisation periods.*

2.3.3 Education

$$\frac{\partial \lambda_{ij}}{\partial d_i} = \overbrace{\frac{\partial e_{ij}}{\partial d_i}}^{(+)} + \overbrace{\frac{\partial \xi_{ij}}{\partial d_i}}^{(+)} - \overbrace{\frac{\partial c_{ij}}{\partial d_i}}^{(0)} - \overbrace{\frac{\partial \mathscr{T}_{ij}}{\partial d_i}}^{(0)} \tag{2.6}$$

Although the effect of education on the expected utility of an investment is predominantly positive, the magnitude and total direction of the overall effect of education will depend on the magnitude of the single effects.

1. Harrison et al. (2002) find large and significantly negative effects of education on the discount rate of the future benefits of an investment. Hence, individuals with a longer education are, on average, more patient, i.e., able and willing to wait for the future benefits of their investment. Hence, we expect that an increase in years of education correlates negatively with the discount rate of future monetary and non-monetary benefits of an energy efficiency investment. This effect should be more prominent, the greater the benefits and the longer the amortisation period.

2. We presume that education has no effect on the up-front monetary and non-monetary costs of the investment.

The positive impact of education will mainly manifest itself on the benefit side, and will be more significant for benefits that accrue further in the future. Thus, we posit the third hypothesis:

Hypothesis 3 *Education is expected to increase the propensity to invest in energy efficiency measures, particularly if the amortisation period is long.*

2.3.4 Household size

$$\frac{\partial \lambda_{ij}}{\partial z_i} = \overbrace{\frac{\partial e_{ij}}{\partial z_i}}^{(+)} + \overbrace{\frac{\partial \xi_{ij}}{\partial z_i}}^{(0)} - \overbrace{\frac{\partial c_{ij}}{\partial z_i}}^{(+)} - \overbrace{\frac{\partial \mathscr{T}_{ij}}{\partial z_i}}^{(0)} \tag{2.7}$$

Household size is primarily a control variable and it will, therefore, only impact the propensity to invest, λ_i, through other variables:

1. Larger household size will, *ceteris paribus*, correlate with a higher demand for energy services. If these energy services can be provided more efficiently, large households will profit over-proportionally through greater energy savings, e_{ij}.

2. On the other hand, larger household size means lower per capita income, which eventually translates into higher costs of financing capital intensive investments.

3. The impact of household size on non-monetary costs and benefits is unclear, and might be more influenced by the structure of the household, e.g., age and number of children, number of elderly, etc.[4]

Hence, the overall impact of household size depends on the magnitude of the two opposing effects. Based on these assumptions, we derive the following hypothesis:

Hypothesis 4 *The effect of household size on the likelihood of a household to invest in energy efficiency measures is ambiguous for capital-intensive investments, but positive for less capital-intensive investments.*

2.4 Analysis

2.4.1 Overview of studies

Table 2.1 summarises the empirical evidence for heterogeneity in household energy efficiency investment behaviour. Houde (2014) appears twice as the study conducts two analyses that fall in different categories.

The column "Data Source" documents the kind of data used as the primary input. Stated Preference studies are referred to as "Choice Experiment (SP)" or "Survey Data (SP)". They survey preferences through choice experiments or through survey questions. Revealed Preference (RP) studies, which are based on statements about current or past behavior, are labelled "Survey Data (RP)". "Empirical Data" are data on household behaviour collected from a third source (e.g., tax authorities).

[4]E.g., a household with small children or elderly persons might put a higher value on a better indoor climate. On the other hand, these households might have less leisure time or will be over proportionally affected by the inconvenience of a large building project.

The column "Policy" indicates whether the study evaluates an energy policy. Observing only general market behaviour, however, does not imply that energy policies were not effective, but rather that the study does not focus on measuring the effect of a particular policy. The distinction is important as the sole participation in a policy evaluation study (experiment) might be sufficient to trigger behavioural changes in households that are very different from their market behaviour.

The "Decision domain" column describes the type of conservation decision that a study investigates. Studies investigating more than one decision are referred to as "Multiple". The included studies cover a vast spectrum of different investment decisions, which can have very different non-monetary benefits. Housing investments often lead to improved indoor climate; appliances, on the other hand, have important comfort features in operating them and might also entail status considerations. These differences in non-monetary benefits may strongly correlate with household characteristics and, thereby, further impact the effect of these characteristics on a household's investment decision.

The column labelled 'n' indicates the sample size of the study. For studies that run different model specifications based on various subsets, we note the range of the sample sizes.

Finally, the columns labelled "Inc", "Age", "Educ", and "Size" display the correlation between the likelihood to invest in energy efficiency and the four household characteristics: income, age, education, and household size. This correlation is inferred from the different dependent variables used in the studies (e.g., adoption decision, investment sums, etc.). A positive (negative) correlation, indicating a higher (lower) propensity to invest in energy efficiency, is represented by '+' ('-'). Studies where no statistically significant correlation is found are indicated by "0". Studies that do not include a determinant are indicated by "ϕ". The values in parenthesis show the t-value for the estimates, where bold font indicates statistical significance. In cases where multiple models were estimated (e.g., different subgroups or model specifications), we indicate the range of the t-values for the different estimates. Cases where t-values were unobtainable or inappropriate for the methodology applied in a study are indicated by “NA”.

Table 2.1: Evidence of Heterogeneity in Household Behaviour related to Energy Efficiency

Study	Data Source	Policy?	Decision domain	n	Inc	Age	Educ	Size
Powers et al. (1992)	Survey Data (RP)	Yes	General Retro-fit	690	+ (**2.7**)	0 (-0.68)	+ (**2.75**)	- (1.57)
Ward et al. (2011)	Choice Experiment (SP)	Yes	Refrigerators	355	**0** (-0.18,0.74)	**-** (**-3.64,-5.51**)	- (-0.37)	ϕ
Mills and Schleich (2010)	Survey Data (RP)	No	Multiple (Appliance)	6,993 - 19,014	0/+ (-0.13 /**2.07**)	0/0 (0.26 / 0.34)	0/0 (0.55 / 1.2)	0/+ (1.32 / **0.045**)
Gamtessa (2013)	Empirical Data	Yes	General Retro-fit	181,558	- **NA**	+ **NA**	- **NA**	- **NA**
Davis and Metcalf (2016)	Choice Experiment (SP)	Yes	Air Conditioners	2,440	+ (**4.03,5.58**)	+ (**3.00,3.69**)	0/+ (0.30,1.13)/(**1.95**)	ϕ
Houde (2014)	Empirical Data	No	Refrigerators	49,279 - 76,115	+ (**2.35**)	+ (**4.84,20.85**)	+ (**3.14,15.75**)	+ (**6.5, 5.9**)
Houde (2014)	Empirical Data	Yes	Refrigerators	49,279 - 76,115	+ (**2.24**)	0/+ (-1.49,**2.95**)	+ (**2.78,3.47**)	+ (**8.96, 6.3**)
Newell and Siikamäki (2015)	Choice Experiment (SP)	Yes	Water Heaters	1,217	+ (**2.79**)	0 (1.12,1.6)	0 (0.41)	- (0.05, 1.46
Fowlie et al. (2015)	Randomised Trial	Yes	General Retro-fit	613	**+** (**2.69**)	0 (0.78)	ϕ	0 (1.5)
Frondel and Vance (2013)	Empirical Data	Yes	General Retro-fit	2,530	0 (-1.5,-1.4)	+ (**10, 16**)	ϕ	ϕ
Walsh (1989)	Survey Data (RP)	No	General Retro-fit	787 - 2,124	0/+ (1.33)/(**4.00**)	0/- (-0.67)/(**-10.00**)	ϕ	ϕ
Long (1993)	Empirical data	No	General Retro-fit	5,871	+ (**8.02**)	+ (**2.49**)	ϕ	- (-1.42)

Table 2.1: Evidence of Heterogeneity in Household Behaviour related to Energy Efficiency

Study	Data Source	Policy?	Decision domain	n	Inc	Age	Educ	Size
Brechling and Smith (1994)	Survey Data (RP)	No	Multiple (Insulation)	5,271 - 6,395	+ (**3.75,3.86**)	0 (0.00,0.48)	ϕ	ϕ
Murray and Mills (2011)	Survey Data (RP)	No	Multiple (Appliance)	1116 - 1920	0 (-0.88, 0.07, 1.41)	0 (-0.24, 1.03, 1.04)	ϕ	0 (0.64, 1.24, 1.05)
Metcalf and Hassett (1999)	Survey Data (RP)	No	Attic Insulation	765	0 NA	ϕ	ϕ	+ (**0.006**)
Dubin and Henson (1988)	Empirical Data	Yes	General Retro-fit	688	+ (**9.47**)	ϕ	ϕ	ϕ
Hasset and Metcalf (1995)	Empirical Data	Yes	General Retro-fit	8,496 - 74,792	+ (**4.58**)	ϕ	ϕ	ϕ
Nauleau (2014)	Survey Data (RP)	No	Multiple (Insulation)	23,879	+ NA	ϕ	ϕ	+ NA
Palmer et al. (2015)	Empirical Data	Yes	Multiple (Insulation)	249 - 301	**0/-** (-1.5)/(**-2.57**)	ϕ	ϕ	ϕ
Sardianou (2007)	Survey Data (SP)	No	Conservation Attitude	500	+ (**2.63**)	- (**-2.29**)	0 (0.64)	+ (**1.95**)

Legend: '+' positive correlation, '-' negative correlation, '0' no correlation, 'ϕ' not part of the study. Results separated by a '/' present results for different subsamples.

2.4.2 Quality of the studies

Table 2.2 gives an overview of the different quality aspects of the collected studies, which obviously are very heterogeneous. Many studies do not conduct any form of robustness analysis and often accept the results from only a single – untested – model specification at face value, which unfortunately reduces the contribution of these studies considerably.

Most studies use a wide set of variables on household and housing characteristics of which many are probably highly correlated (e.g., income and home ownership, family size and dwelling area, or income and education). However, we did not find a single study that discusses this issue. Although collinearity does not bias estimators, it leads to inefficient estimators (large standard errors), which might be problematic in studies that are primarily investigating the determinants of households' energy decisions or studies that are interested in the correlation between policy effects and household and housing characteristics.

A further aspect is the external validity of the study results. As Table 2.2 shows, very few studies work with a final sample that is unbiased and facilitates full generalisation of the results. Most studies – unsurprisingly – experience considerable problems with low response-rates, non-random attrition, or problems of missing observations in certain variables for non-random subsets of the sample (e.g., missing taxation data for households in certain income groups). Other studies (e.g., Murray and Mills, 2011), voluntarily reduce a representative random sample to an unrepresentative sub-sample.

This problem is not in any way specific to the selected studies, but is a widespread problem of observational studies in economics and other social sciences. What is problematic, though, is that many of the cited studies use their results as the basis for extensive policy recommendations, often without discussing the external validity of their results or checking the representatives of their sample. In this respect, the growing body of studies using data from carefully designed field experiments (e.g., Wilhite and Ling, 1995; Abrahamse and Steg, 2009; Ito et al., 2014; Fowlie et al., 2015) is a positive trend in the literature that should be further encouraged.

Finally, we encountered a number of studies that either ignore self-selection of households, e.g., self-selection into energy efficiency programmes, or treat self-selection problems in a very nonchalant way in analyses evaluating the effectiveness of policies promoting energy conservation in households. Since the seminal work by James Heckman (e.g., Heckman, 1979; Heckman and Vytlacil, 2001, 2007) it has become apparent that ignoring self-selection of decision-makers into policy treatments can lead to inconsistent estimates of policy effects. Unfortunately, recent studies also often ignore self-selection problems in their policy analyses, which indicates that awareness of this problem is still low. Therefore, studies following approaches like in Murray and Mills (2011) or Mills and Schleich (2010), which fully account for self-selection, should be further encouraged.

Table 2.2: Quality of selected studies

Study	Research question	Sample	Appropriate study design	Risk of bias	Outcome measure	Statistical Issues	Generalisability
Powers et al. (1992)	Factors determining participation in energy efficiency programs.	Survey data,	Yes	Non-random opt-out during survey.	Participation in energy efficiency programme and subsequent investment.	No correction for sample selection problem of the investment decision, both participation and investment decision treated as two separate decisions.	No
Ward et al. (2011)	Influence of the Energy star label on consumers' willingness to pay for an appliance.	Online survey, USA, 2009, representative online research panel.	Study based on stated preferences, danger that respondents give 'correct' answer because there are no real financial consequences of their decision.	None, sample corrected for non-random opt-out.	Choice of refrigerator model	None	Yes
Mills and Schleich (2010)	Determinants of households' knowledge and choice of class A energy appliances.	Mail survey, Germany, 2002.	Yes	Response rate 75%, authors do not discuss or control for potential bias.	Knowledge and ownership of class A energy appliances.	None	Unclear, as potential bias in the sample is not discussed.
Gamtessa (2013)	Determinants of households' participation in the EGH home retrofit program.	EGH audit reports submitted to NRCan, Canada, 1998–2005.	Study design does not include valid exclusion restriction.	Non-random attrition of households due to missing values in some variables.	Number of implemented retro-fit measures.	Do not correct for self-selection problem in investment decision, due to missing IV.	No

Table 2.2: Quality of selected studies

Study	Research question	Sample	Appropriate study design	Risk of bias	Outcome measure	Statistical Issues	Generalisability
Davis and Metcalf (2016)	The effect of tailoring energy label information to the state of residence on households' purchase decisions of air conditioners.	Experimental data, USA.	Study based on stated preferences, risk that measured behaviour does not reflect real world behaviour. However, legal restrictions do not allow a field experiment.	Response rate 62.5%, risk of non-random opt-out of households.	Efficiency of the choice of air conditioner model as a measure of the extent to which households absorb information on energy label.	None	Unlikely
Houde (2014)	The effect of energy labels and energy cost information on households' appliance purchase decisions.	Scanner data from large retailer, USA, 2007–2011	Yes	Households' demographic data only available for 40% of all purchases, no discussion of eventual implications for the final sample.	Purchase of refrigerator models.	None	Unclear
Newell and Siikamäki (2015)	The effect of individual discount rates on energy efficiency decisions.	Experimental data based on representative sample of single-family home-owners, USA.	Study based on stated preferences that can lead to overstated willingness to pay and the omission of unobserved costs that households would face in real life	None	Choice of boiler model.	None	Yes

Table 2.2: Quality of selected studies

Study	Research question	Sample	Appropriate study design	Risk of bias	Outcome measure	Statistical Issues	Generalisability
Fowlie et al. (2015)	Determinants of low-income families' adoption decision of the Federal Weatherization Assistance Program.	Experimental data on low-income households, Michigan (USA), 2011	Yes	It is unclear to what extent the treated households could communicate with untreated households, which could eventually lead to spill-over effects of the treatment to the control group, thereby, diluting the treatment results.	Household's participation in programme.	None	Yes

Table 2.2: Quality of selected studies

Study	Research question	Sample	Appropriate study design	Risk of bias	Outcome measure	Statistical Issues	Generalisability
Frondel and Vance (2013)	The impact of home energy audits on households' retro-fit decisions.	Residential Energy Consumption Survey of detached house owners, Germany, 2007.	Study design does not include valid exclusion restriction.	Unclear	Audit participation and implementation of at least one of four renovation measures (roof insulation, facade insulation, window replacement, and replacement of heating equipment)	Do not correct for self-selection problem in investment decision, due to missing IV.	No
Walsh (1989)	The impact of income and tax credits on households' retro-fit decisions.	Residential Energy Consumption survey, USA, 1982.	Yes	Unclear, not discussed.	Binary variable on household's investment.	Does not take unobserved heterogeneity into account.	Unclear.

Table 2.2: Quality of selected studies

Study	Research question	Sample	Appropriate study design	Risk of bias	Outcome measure	Statistical Issues	Generalisability
Long (1993)	Determinants of retro-fit investments.	IRS tax files containing energy investment related tax returns, USA, 1978–85.	Study misses counter factual to be able to quantify effect of tax returns on retro-fit investments.	Households with $> \$200,000$ excluded, Alaska and Hawaii excluded. extensive discussion of the impact of these selection criteria on the final sample.	Investment sum and retro-fit measures.	None	Unclear
Brechling and Smith (1994)	Determinants of retro-fit investments.	English House Condition, loft insulation, wall insulation, double glazing Survey (EHCS), England, 1986.	Yes	None	Binary variables on decision to conduct loft insulation, wall insulation, and double glazing investments.	Decisions to conduct these three measures most probably not independent, but estimated as three separate models. Results still consistent, but eventually less efficient. No robustness analysis and no tests of significant differences in the parameters across all three investment models.	Yes

Table 2.2: Quality of selected studies

Study	Research question	Sample	Appropriate study design	Risk of bias	Outcome measure	Statistical Issues	Generalisability
Murray and Mills (2011)	Determinants of households knowledge and choice of energy Energy Star white appliances.	Residential Energy Consumption Survey (RECS), USA, 2005.	Yes	Only include households with appliances < 5 years, leaves only 55%–60% of initially representative sample.	Binary variables on knowledge and purchase of energy star labelled white appliances.	Decision to purchase energy star refrigerator, dish washer, and washing machine might be correlated. Modelled as three independent decisions. Results still consistent, but eventually less efficient.	No
Dubin and Henson (1988)	Distributional consequences of the Energy Tax Act.	IRS Taxpayer Compliance Measurement Program, USA, 1979.	Yes	Only get data from long tax form, which is more often filled out by wealthier households. These might be over represented in the sample.	Average dollar amount of credit claimed per return file.	Does not take unobserved heterogeneity into account.	No

Table 2.2: Quality of selected studies

Study	**Research question**	**Sample**	**Appropriate study design**	**Risk of bias**	**Outcome measure**	**Statistical Issues**	**Generalisability**
Hasset and Metcalf (1995)	Impact of tax return programmes on households' willingness to invest in retro-fit measures.	TCMP tax return data, Ernst & Young/University of Michigan Tax Research Database, simple random sample based on social security number, USA, 1978-1985.	Yes	None	Binary variable, tax credit filed.	None	Yes
Nauleau (2014)	Share of free riders of an energy renovation tax deduction scheme.	Energy Management (EM) annual survey, France, 2002-2011 (unbalanced panel).	Analysis based on observational data. True impact analysis only possible in controlled experimental setting or in natural experiment. Use of 'naive' difference measure, which basically compares situation before and after policy introduction while controlling for trends in variables.	Various selection criteria lead to unrepresentative sample.	Binary variable on decision to retrofit.	Assume independence between time-invariant random effects and explanatory variables without testing claim. As data is observational, independence is highly unlikely, therefore, problematic choice of estimator.	No

Table 2.2: Quality of selected studies

Study	Research question	Sample	Appropriate study design	Risk of bias	Outcome measure	Statistical Issues	Generalisability
Palmer et al. (2015)	Factors explaining households' follow-up on home energy audits.	Survey data, USA, 2014.	Yes	Unclear	Follow-up investment	None	Unclear
Sardianou (2007)	The impact of lifestyle variables on households' energy conservation patterns.	Survey data, Athens (Greece), 2003.	Analysis based on stated preferences. Only surveys intentions to implement measure. Treats binary intention variables as count data, by simply adding them up, not taking into account the large qualitative differences in the surveyed measures.	Unclear	Binary variables on household's intention to adopt specific energy conservation behaviours.	Unclear which estimator is used for the model with count data as a dependent variable and the model with a binary dependent variable.	No

2.4.3 Methodology and Results of the meta-analysis

The difficulty of conducting a meta-analysis on correlation studies is the vast difference in the chosen estimates, model specifications, and units of measurement of the endogenous variable as well as the covariates, e.g., a simple shift in the scale of the income variable (income in \$ versus income in €) makes the parameters from two studies incomparable.

To address the first problem, we divide the studies into two categories: (i) those with a binary outcome variable, usually adoption or participation studies, and; (ii) those with a continuous outcome variable measuring investments in energy efficiency. Dividing the studies up into these two categories provides the basis for comparing the magnitude of the effects. However, even within models with binary outcome variables and models with continuous outcome variables, we find a multitude of different econometric approaches applied to the data. Several studies compare various econometric approaches to the same model specification in order to test different econometric assumptions. This creates a trade-off for our choice of estimates between the quality of the estimator and the comparability of the estimates across studies, i.e., a study might contain both the results from a standard econometric approach and those from a very specialised approach, with the latter returning more valid results. In order to increase the comparability of the estimates across all studies, we choose to include, as far as possible, the results from the standard estimator.

The second point is the influence of the inclusion of other covariates on the effects of the four extracted variables. Evidently, the number and type of covariates vary considerably across studies. In cases where some of these additional covariates are strongly correlated with one or several of our four household characteristics, the measured effect of the latter may be influenced by their mere presence or absence in the analysis. We see no possibility of controlling for this influence in our analysis as the variation in the model specifications across all studies is too large to be systematised in any meaningful way.

Another problem that results from different model specifications is that not all studies include all four variables (income, age, education, and household size), which means that the mean effects of these four variables are based on different subsets of the included studies. This is a weakness in our analysis, which we cannot control for as reducing our sample to the subset of studies containing all four variables would leave us with only five studies. However, we thoroughly check the influence of each study on the mean effect and we discuss this point more thoroughly in section 2.5.

A further difficulty are studies that use ordinal covariates such as income or age categories, or educational degrees. Naturally, the number of categories and the interval borders are seldom identical across studies and, hence, we need to find a mean effect across all categories, whilst still preserving any eventual non-linear effect of the variable. We solve this problem in the following way:

- Only one study uses income categories and, as our theoretical model and most studies assume a linear effect of income, we calculate the weighted mean effect across all income categories.[5]

- Many studies report a non-linear influence of age, and either include a quadratic age term or several age categories. As a non-linear influence of age is plausible, also from the point of view of our theoretical model, we resume the estimates (also linear estimates) into estimates for three age categories: young, 18–35 (base line), middle-age, 36–55, and old-age, > 55.

- In the case of education, most studies differentiate between households with or without a college degree. We maintain this distinction and calculate the effect of a college versus non-college degree, with the latter as the base category. For studies that use a different base category, we re-base the categories to fit our narrative.

- Household size is predominately included as a count data variable assuming a linear

[5] $\bar{\beta} = \frac{1}{K}\sum_k \left(\frac{\beta_k - \beta_{k-1}}{A_k - A_{k-1}}\right)$, with K the number of categories $k = 1, \ldots, K$, and A_k the mid-range value of category k.

effect. For the few studies that use categories for household size, we calculate the weighted mean effect across all categories (same approach as in income).

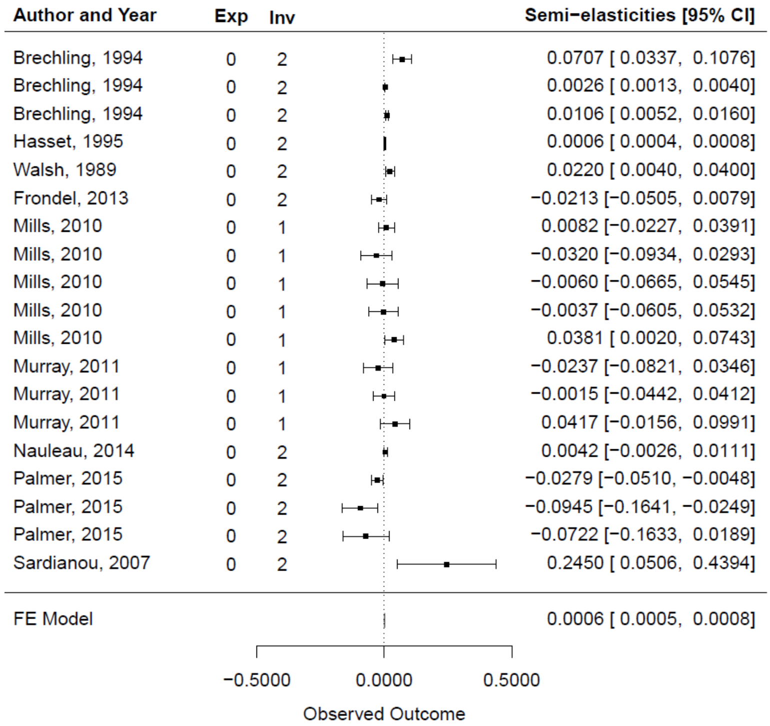
Finally, in order to overcome the problem of different units of measurement for both the endogenous variable and the covariates, we calculate the semi-elasticity of the marginal effect for the continuous variables (income and household size) in studies of category (i) and the elasticity for each of these two variables in studies of type (ii) at the sample mean.[6] For the categorical variables (age and education), we calculate the effect at the sample mean. In order to calculate approximate standard errors of the calculated (semi-)elasticities, we first applied the delta-method (Greene, 2012) at the sample mean. However, as none of the included studies published the covariance matrix of the estimated parameters, we had to set the covariances in the variance-covariance matrix to zero. Tests showed that the calculated standard errors proved to be too far off the mark. Hence, we decided instead to disregard the variance of $\phi(X\beta)$ for probit analyses and the variance of $p(1-p)$ for logit analyses because the inclusion of these terms requires the non-available variance-covariance matrix of the estimates. As such, this approximation of the standard errors of the calculated elasticities then only requires the standard error (or t-value) of the estimated parameter and not the non-available covariance matrix.

A number of studies do not report summary statistics and/or the standard errors of their estimates or only measure pairwise correlations between the endogenous variable and household characteristics, which means that we, unfortunately, had to remove these studies from the meta-analysis, although we include the direction of the measured effects in our discussion.[7] However, a number of the remaining studies conduct analyses on

[6] For studies of category (i): the semi-elasticity of 'income' and 'household size' is calculated as $\varepsilon_{x_i} = \frac{\partial p(y=1)}{\partial x_i} \cdot \bar{x}_i$, with y the binary dependent variable, and x_i the continuous explanatory variable, as well as $\bar{x}_i$ the sample mean of the continuous explanatory variable. The effect of 'age' and 'education' is calculated as $\Delta y = p(y=1|x_i = A_k) - p(y=1|x_i = A_{base})$, with k the respective category. For studies of category (ii): the elasticity of 'income' and 'household size' is calculated as $\varepsilon_{x_i} = \frac{\partial y}{\partial x_i} \cdot \frac{\bar{x}_i}{\bar{y}}$ and the effect of 'age' and 'education' as $\Delta y = \frac{y(x_i = A_k) - y(x_i = A_{base})}{\bar{y}}$.

[7] Studies not included in the meta-analysis are: Powers et al. (1992); Gamtessa (2013); Houde (2014);

multiple cases, which boosts the number of observations in our meta-analysis.[8] Our final



Exp = dummy variable experimental study, Inv = investment category.

Figure 2.1: Semi-elasticities for income, binary outcome

sample contains 13 articles, with a total of 22 observations for income; 21 observations for age; 12 observations for education; and 14 observations for household size.

We use the fixed-effects estimator from the add-on package ‘metafor’ (Viechtbauer, 2010) to the statistical software R (R Core Team, 2017), estimating the mean effect weighted by the standard errors of the estimates across all studies. We also test the influence of two moderators: ‘experiment’ and ‘investment’, measuring the effect of ex-

Fowlie et al. (2015); Davis and Metcalf (2016).

[8] One could argue that by using multiple cases from the same study, estimates are no longer independent observations, which would require a hierarchical model to estimate the mean effect. However, all studies that report results from multiple cases base their estimates on different sub-samples from their target population, which means that independence between observations might not be completely fulfilled, although correlations across observations are expected to be reasonably low.

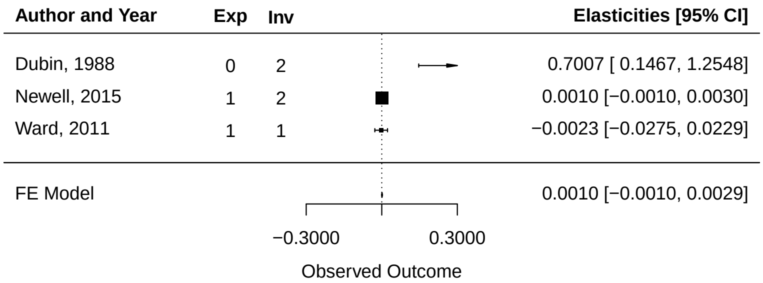


Figure 2.2: Elasticities for income, continuous outcome

perimental data (0 = 'no experiment', 1 = 'experiment') and the investment level (1 = 'medium' – mainly appliances – and 2 = 'high' – mainly retrofit investments) on the magnitude of the (semi-)elasticities, respectively.

Figures 2.1–2.10 report the results for the weighted mean effects of the semi-elasticities and elasticities, as well as the respective semi-elasticities and elasticities and the associated 95% confidence intervals for each included study.

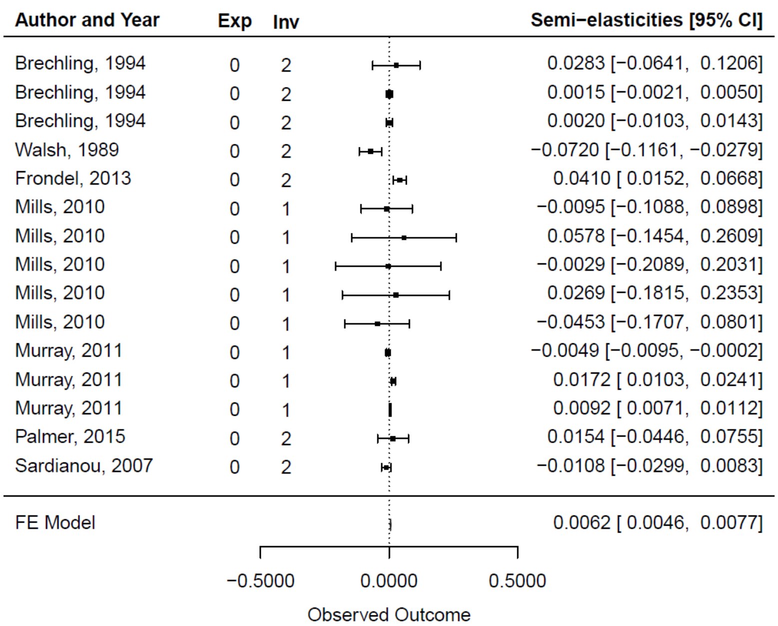


Figure 2.3: Semi-elasticities for medium-age (35–55), binary outcome

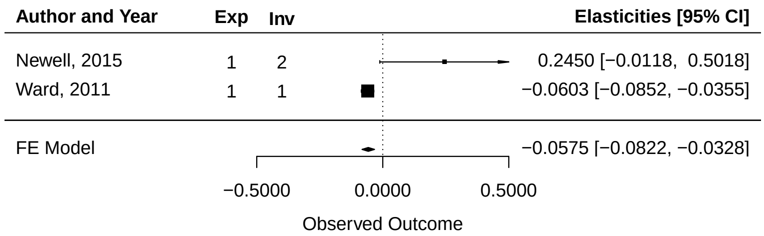


Figure 2.4: Elasticities for medium-age (35–55), continuous outcome

Table 2.3 reports the effects of the two moderators for each of the 5 models. Unfortunately, our data does not contain enough information to allow us to test all possible combinations of moderator effects. Additionally, some combinations contain very few observations. Results from estimations with fewer than 4 observations have been omitted.

As the dependent variables do not always follow a normal distribution, we adopt the approach suggested by Viechtbauer (2010) and also calculate exact permutation tests in order to ensure valid and more conservative p-values (see Table 2.8 in the appendix).

In order to identify influential observations in the estimation of the mean effects, we compare the z-standardised residuals, Cook's distance, and the calculated weights of each observation for each estimation model. Results from Hasset and Metcalf (1995) have a considerable influence on the fixed-effect estimate of the semi-elasticity for income as the estimate from this study has a relatively small standard error. Results from Murray and Mills (2011) and Ward et al. (2011) have a considerable influence on the fixed-effect estimates of the old-age and education effect. None of these influences vanish when moderators are included.

To check the de facto effect of these studies on the magnitude of the fixed-effect estimates, we follow Viechtbauer (2010) and conduct a leave-one-out analysis, where each study is omitted from the estimation in turn (see Table 2.6 for details). Overall, the results are remarkably stable over all subsets of the studies. However, the impact

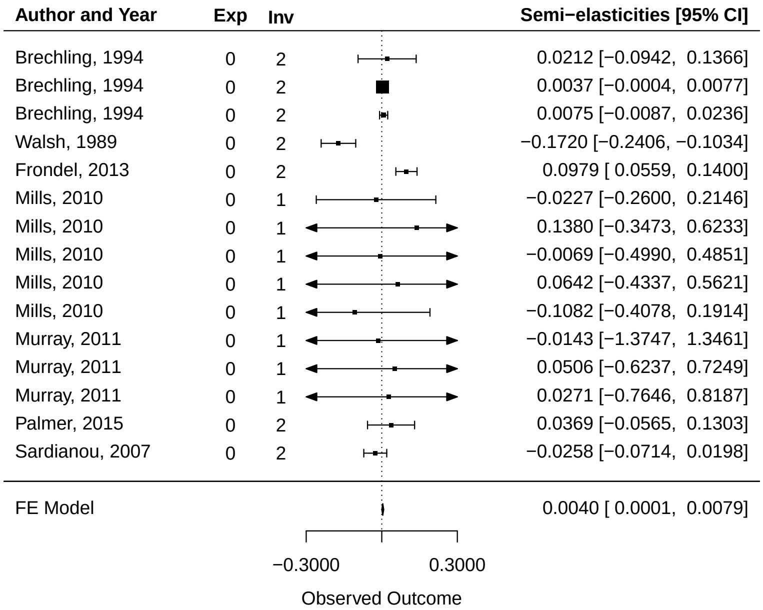


Figure 2.5: Semi-elasticities for old-age (> 55 years), binary outcome

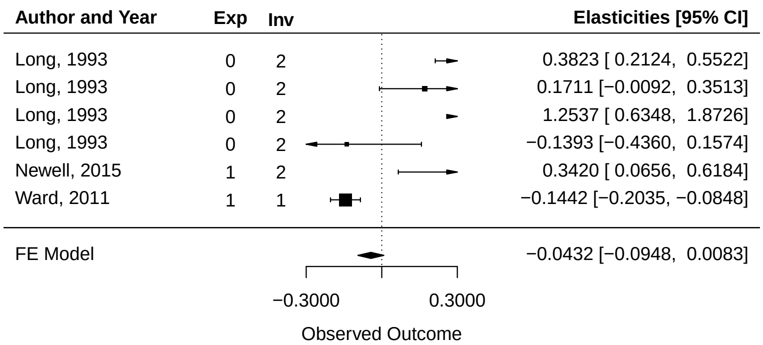


Figure 2.6: Elasticities for old-age (> 55 years), continuous outcome

of Hasset and Metcalf (1995) on income and Murray and Mills (2011) and Ward et al. (2011) on the fixed-effect estimates of age are still obvious. As all three studies report very low standard errors for their estimates, we are reluctant to remove these studies from our analysis.

Finally, in order to check the normality of the residuals from the models that include

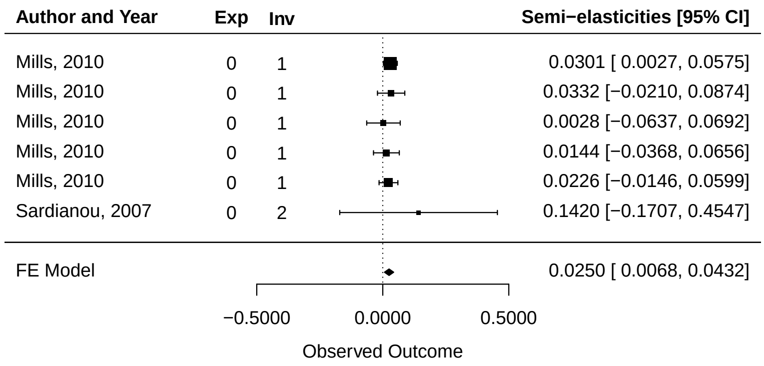


Figure 2.7: Semi-elasticities for education, binary outcome

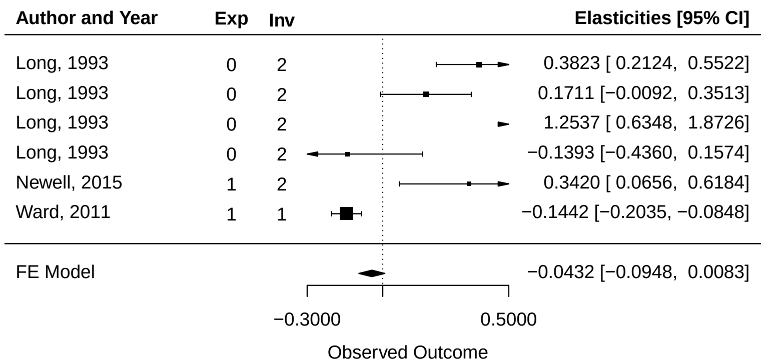


Figure 2.8: Elasticities for education, continuous outcome

moderators, we plot the residuals in a Q-Q plot (see Figure 2.11). The plots show that almost all models have residual distributions with heavy tails, revealing that some observations still markedly deviate from the mean, which indicates that our moderators might be insufficient to capture all the systematic variance in the elasticities (dependent variables) and that other factors might affect the deviation of these observations from the mean effect.

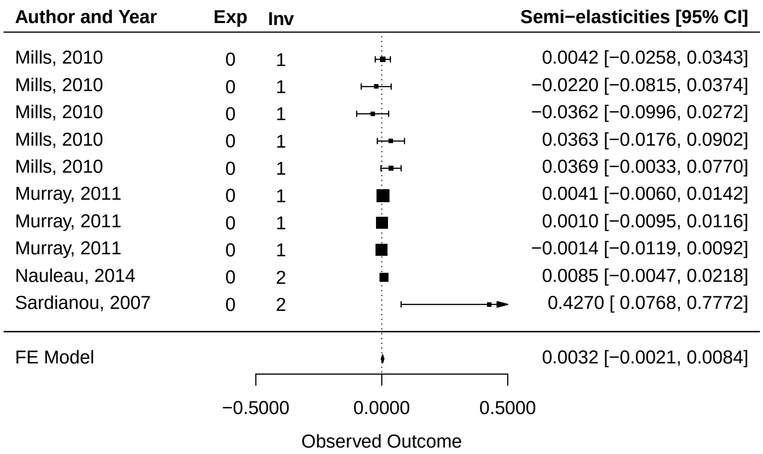


Figure 2.9: Semi-elasticities for household size, binary outcome

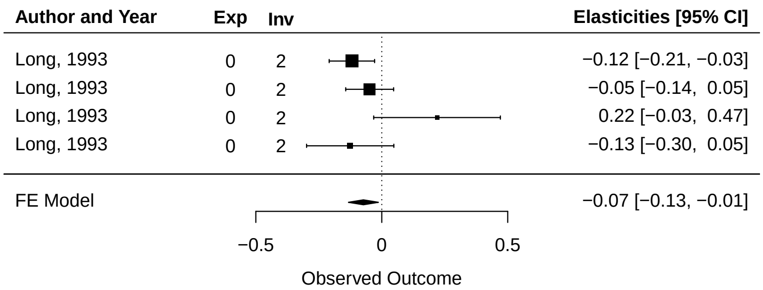


Figure 2.10: Elasticities for household size, continuous outcome

2.5 Discussion

We set out with the goal to identify the existence of clear patterns across four frequently used and easily observable household characteristics (income, age, education, and household size) that can be exploited to design and model heterogeneous energy efficiency policies. Our results confirm to some extent the existence of patterns in the effects of the four variables across the included studies.

However, as discussed in the previous section, the fact that the fixed-effects estimates are based on different subsets of the studies – with the exception of middle-age and old-age – means that a direct comparison of the magnitude of the different effects would

Table 2.3: The influence of moderators (experimental study and investment level)

(a) Income, binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.01	0.01	1.00	0.32	-0.01	0.02
investment: 2	-0.01	0.01	-0.92	0.36	-0.02	0.01

(b) Income, continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.70	0.28	2.46	0.01	0.14	1.25
experiment: 1	-0.70	0.28	-2.48	0.01	-1.25	-0.15
investment: 2	0.00	0.01	0.26	0.80	-0.02	0.03

(c) Mid-age (45), binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.01	0.00	8.21	0.00	0.01	0.01
investment: 2	-0.01	0.00	-3.21	0.00	-0.01	-0.00

(d) Old-age (70), binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-0.01	0.08	-0.19	0.85	-0.16	0.13
investment: 2	0.02	0.08	0.24	0.81	-0.13	0.17

(e) Old-age (70), continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-0.23	0.16	-1.49	0.14	-0.54	0.07
experiment: 1	0.09	0.15	0.58	0.57	-0.21	0.39
investment: 2	0.49	0.14	3.37	0.00	0.20	0.77

(f) Education, binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.02	0.01	2.64	0.01	0.01	0.04
investment: 2	0.12	0.16	0.73	0.46	-0.20	0.43

(g) Education, continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-0.01	0.02	-0.37	0.71	-0.05	0.03
investment: 2	0.23	0.14	1.65	0.10	-0.04	0.50

(h) HH size, binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.00	0.00	0.70	0.48	-0.00	0.01
investment: 2	0.01	0.01	0.96	0.34	-0.01	0.02

experiment: 0 = no experimental data, 1 = experimental data
investment: 1 = medium scale monetary investment (appliances),
2 = large scale monetary investments (retrofit measures)

be fallacious. Influential studies, e.g., Hasset and Metcalf (1995) or Murray and Mills (2011), do not include all four variables and, hence, are not included in the estimation of each of the four effects. Furthermore, we cannot disentangle the weight of their estimates in the calculation of the mean effect from the selected model specification. If, for example, age, education, and income are strongly correlated, estimates from a study such as Hasset and Metcalf (1995), which only includes income, might get a relatively small standard error for the income estimate compared to studies that include all three variables, e.g., Mills and Schleich (2010), and, consequently, will dominate the mean effect of income much more than the latter[9]. Nevertheless, despite these shortcomings, we can examine the direction of the mean effect and compare the consistency of the results for each of the four effects.

Although some results are ambiguous, we can distinguish some tendencies in all four mean effects. Tendencies that are mirrored in the estimates from the studies not included in our meta-analysis (see Table 2.1):

- Although we find largely contradictory results for income across the included studies – sometimes even within the same study – income still displays the least ambiguous effect of all the four variables; the positive effect of income is confirmed by the mean effects of studies with binary and those with continuous outcome variables – although the latter is statistically insignificant (small sample size) (see Figures 2.1 and 2.2). In the case of the binary studies, a 1% increase in income increases the propensity to invest in energy efficiency by 0.06%. This effect remains consistently positive under all combinations of the included studies (see Table 2.6 (a)), which confirms that the positive effect of income is not driven by a single study, but remains stable irrespective of the specific subset of studies we apply. Furthermore, four of the five excluded studies confirm a positive effect of income on the propensity to invest. Hence, we can state that income has a consistent and positive effect on the propensity to invest.

- Age shows more ambiguous results (see Figures 2.3–2.6). For middle-age (45 years)

[9] It could be argued that the exclusion of certain household characteristics in some of the studies leads to omitted variable bias, but we refrain from discussing this further.

we find a significantly positive result for studies with binary outcome variables. In comparison to the young group, the likelihood of middle-aged households investing in energy efficiency increases by 0.62 percentage points. However, for studies with continuous outcome variables, the effect of being middle-aged is significantly negative. Compared to the young group, middle-aged households invest 5.57% less. However, as this effect is only based on two studies, with Ward et al. (2011) negatively dominating the effect, we discount this finding in favour of the positive effect from the first set of studies.

The picture looks quite similar for the effect of old-age (70). The mean effect of the old-age category based on studies with binary outcome variables is significantly positive. Compared to the young group, households in the old-age category are 0.4 percentage points more likely to invest in energy efficiency. A simple Wald test confirms that for studies with a binary outcome variable, the effect of old-age is significantly smaller than the effect of middle-age, which hints at an inverted U-shape effect for age. The effect of old-age based on studies with continuous outcome variables is negative, though statistically insignificant.

For studies with binary outcome variables, both effects remain positive and stable under all combinations of the included studies (see Table 2.6 (c) and (e)). However, the negative effect of old-age for studies with continuous outcome variables persists in all combinations of included studies but one (see Table 2.6 (f)), which strongly indicates that the negative effect of this subset is mainly driven by the results from Ward et al. (2011).

The five excluded studies show ambiguous results with a tendency towards a positive effect, which complements the overall picture of a tendentially positive, albeit non-linear, effect for age.

- Education shows a distinctively positive effect in studies with a binary outcome variable; on average a college educated household will be 2.5 percentage points more likely to invest in energy efficiency compared to a non-college educated household. For studies with continuous outcome variables, however, this effect is reversed; college educated households invest 4.3 % less than non-college educated households. However, this effect is statistically insignificant. Of the five excluded studies, three confirm a positive effect of education. Hence, we carefully conclude that college education tends to have a positive effect on the propensity to invest.

- Finally, household size has the most ambiguous effect of all four variables. The mean effect of household size for studies with binary outcome variables is positive, albeit statistically insignificant. For studies with a continuous outcome variable, the effect is statistically significant and negative. A 1% increase in household size leads to a 7% decrease in investments. However, this result is based on cases from a single study (Long, 1993) and should, therefore, be interpreted with care.

  The results in Tables 2.6 (g) and (h) show that the positive and the negative effects are consistent across all combinations of the included studies. This seems to indicate that household size has a dual effect on the propensity to invest, in the sense that household size might increase the likelihood of adopting energy efficient appliances or retrofitting measures, but might decrease the amount spent on such measures.

  The five excluded studies confirm a negative effect of household size on the propensity to invest. We, therefore, carefully conclude that household size displays a weak tendency to influence the propensity to invest negatively.

We now turn to the hypothesis derived in section 2.3. Table 2.5 compares each hypothesis to the results from the mean effect as well as the effects of the moderators from Table 2.3. As can be seen in Table 2.3, most estimates for 'investment' are statistically insignificant, and the few significant effects vanish when we use the more conservative

Table 2.5: Comparison of hypotheses with empirical results

Hypotheses	Results	Confirmed?
Hyp 1 A household's propensity to invest will increase with income for capital intensive investments.	Positive effect of income on both adoption and investment sum. No impact of investment category on income elasticities.	Partly
Hyp 2 The effect of age on the likelihood of a household to invest in energy efficiency measures is ambiguous for capital-intensive investments with long amortisation periods.	Positive and non-linear effect of age on adoption. Ambiguous effect of investment sum. Higher investment category significantly reduces likelihood of participation for middle-age, but increases semi-elasticity of investment sum for old-age. However, estimates for old-age are lower than for mid-age, which indicates decreasing propensity to invest from a certain age onwards.	Partly
Hyp 3 Education is expected to increase the propensity to invest in energy efficiency measures, particularly if the amortisation period is long.	The impact of education tends to be positive for adoption, but negative for investment sum. However, the effect of income on investment sum is more positive for a high level of investment.	Partly
Hyp 4 The effect of household size on the likelihood of a household to invest in energy efficiency measures is ambiguous for capital-intensive investments, but positive for less capital-intensive investments.	Negative effect of household size on investment sum, but positive effect on adoption. However, there is no significant effect of investment level on likelihood of adoption.	Partly

permutation test (see Table 2.8). Furthermore, the few significant effects seem to contradict our theoretical assumptions (see effects age-middle and age-old). However, given the limited number of observations in our analyses, these results should be interpreted with care. Although we only include up to two moderators in our analyses, many combinations have too few observations or not enough variation in the moderators to return meaningful results (see Figures 2.1-2.10). Although we cannot confidently fully reject or confirm the hypotheses based on our results, we can conclude that our results support the derived hypotheses to some extent. However, our results cannot confirm a qualitative distinction of the effects of the four household characteristics with regard to the magnitude of the energy efficiency investment, i.e., we find no confirmation that household characteristics have a different impact for larger energy efficiency investments compared to smaller energy efficiency investments.

2.6 Conclusion

Our results indicate that some patterns exist across the four investigated household characteristics. However, the results show that the limiting factor in our analysis is still the number of observations. Although our largest data set comprises 22 observations, in total we had to base our study on only 13 articles. A number of studies had to be removed due to simple reasons such as missing standard errors or t-values, or missing summary statistics. Naturally, we strongly recommend that these measures are included in all empirical studies. The limited number of studies puts considerable strain on our ability to further investigate and explain the partly considerable variance in the estimated elasticities and semi-elasticities. We, therefore, strongly encourage repetition studies and further investigations in the field of households' energy efficiency investment behaviour in order to augment the volume of available data for future analyses. Currently, we deem the empirical evidence to be insufficient to support valid policy recommendations for targeted energy policies.

Although we found a considerable number of studies that investigate adoption or participation decisions, we found few studies that investigate the determinants of the actual investment sums. Adoption of energy efficiency measures or participation in energy efficiency programmes only reveals one side of household heterogeneity with regard to the energy efficiency gap. If we want to gain a thorough understanding of the lack of households' investments in energy efficiency in order to design targeted and tailored energy policies, we need more information on the determinants of actual investment decisions. As our results demonstrate, using the decision domain as a proxy for the investment sum is too rough a measure to return valid estimates of the correlation between investment sum and household characteristics. Future work should, therefore, concentrate on the heterogeneity in the households' monetary (and non-monetary) costs of energy efficiency investments.

Acknowledgements

We are highly indebted to Arne Henningsen, Frits Møller Andersen, Lars Gårn Hansen, and Elizabeth Wilson for helpful comments and discussions on this and previous versions of this paper. This work was supported by the Danish Innovation Fund as part of the SAVE-E project.

Bibliography

Aaronson, S., Cajner, T., Fallick, B., Galbis-Reig, F., Smith, C. L., Wascher, W., 2014. Labor force participation: Recent developments and future prospects. Tech. Rep. 2014-64, Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

Abrahamse, W., Steg, L., 2009. How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings? Journal of Economic Psychology 30, 711–720.

Allcott, H., Greenstone, M., 2012. Is there an energy efficiency gap? Journal of Economic Perspectives 26 (1), 3–28.

Allcott, H., Greenstone, M., 2015. Measuring the welfare effects of energy efficiency programs. Tech. rep., NYU.

Allcott, H., Mullainathan, S., Taubinsky, D., 2014. Energy policy with externalities and internalities. Journal of Public Economics 112, 72–88.

Brechling, V., Smith, S., 1994. Household energy efficiency in the uk. Fiscal Studies 15 (2), 44–56.

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

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

Dubin, J. A., Henson, S. E., 1988. The distributional effects of the federal energy tax act. Resources and Energy 10, 191–212.

Fogg, B., 2003. Persuative Technology. Morgan Kaufmann Publishers.

Fouquet, R., Pearson, P. J., 2012. The long run demand for lighting: Elasticities and rebound effects in different phases of economic development. Economics of Energy and Environmental Policy 1 (1), 83–100.

Fowlie, M., Greenstone, M., Wolfram, C., 2015. Are the non-monetary costs of energy efficiency investments large? understanding low take-up of a free energy efficiency program. American Economic Review: Papers & Proceedings 105(5), 201–204.

Frondel, M., Vance, C., 2013. Heterogeneity in the effect of home energy audits: Theory and evidence. Environmental and Resource Economics 55 (3), 407–418.

Gamtessa, S. F., 2013. An explanation of residential energy-efficiency retrofit behavior in canada. Energy and Buildings 57, 155–164.

Gillingham, K., Palmer, K., 2013. Bridging the energy efficiency gap: insights for policy from economic theory and empirical analysis. Tech. rep., Resources for the Future.

Greene, W. H., 2012. Econometric Analysis. Pearson.

Harrison, G. W., Lau, M. I., Williams, M. B., 2002. Estimating individual discount rates in denmark: A field experiment. American Economic Review 92 (5), 1606 – 1617.

Hasset, K. A., Metcalf, G. E., 1995. Energy tax credits and residential conservation investment: Evidence from panel data. Journal of Public Economics 57, 201–217.

Hausman, J. A., 1979. Individual discount rates and the purchase and utilization of energy-using durables. The Bell Journal of Economics, 33–54.

Heckman, J. J., 1979. Sample selection bias as a specification error. Econometrica: Journal of the econometric society 47 (1), 153–161.

Heckman, J. J., Vytlacil, E., 2001. Policy-relevant treatment effects. The American Economic Review 91 (2), 107–111.

Heckman, J. J., Vytlacil, E., 2007. Econometric evaluation of social programs, part i: Causal models, structural models and econometric policy evaluation. In: Heckman, J. J., Leamer, E. E. (Eds.), Handbook of Econometrics. Vol. Volume 6, Part B. North Holland, Ch. 70, p. 4779–4874.

Houde, S., 2014. How consumers respond to environmental certification and the value of energy information. Tech. rep., National Bureau of Economic Research.

Ito, K., Takanori, I., Tanaka, M., 2014. The persistence of moral suasion and monetary incentives: Experimental evidence from energy demand. Working Paper 20910, NBER Working Paper Series.

Jappelli, T., 1990. Who is credit constrained in the u. s. economy? The Quarterly Journal of Economics 105 (1), 219–234.

Long, J. E., 1993. An econometric analysis of residential expenditures on energy conservation and renewable energy sources. Energy Economics 15 (4), 232–238.

Lyons, A. C., 2003. How credit access has changed over time for u.s. households. The Journal of Consumer Affairs 37 (2), 231–255.

Metcalf, G. E., Hassett, K. A., 1999. Measuring the energy savings from home improvement investments: evidence from monthly billing data. Review of economics and statistics 81 (3), 516–528.

Mills, B., Schleich, J., 2010. What's driving energy efficient appliance label awareness and purchase propensity? Energy Policy 38 (2), 814–825.

Murray, A. G., Mills, B. F., 2011. Read the label! energy star appliance label awareness and uptake among us consumers. Energy Economics 33 (6), 1103–1110.

Nauleau, M.-L., 2014. Free-riding on tax credits for home insulation in france: An econometric assessment using panel data. Energy Economics 46, 78–92.

Newell, R., Siikamäki, J., 2013. Nudging energy efficiency behavior: The role of information labels. Journal of the Association of Environmental and Resource Economists 1(4), 555 – 598.

Newell, R., Siikamäki, J., 2015. Individual time preferences and energy efficiency. American Economic Review: Papers & Proceedings 105(5), 196–200.

Palmer, K., Walls, M., O'Keeffe, L., 2015. Putting information into action: What explains follow-up on home energy audits? Tech. Rep. DP 15-34, RFF.

Powers, T. L., Swan, J. E., Lee, S.-D., 1992. Identifying and understanding the energy conservation consumer: A macromarketing systems approach. Journal of Macromarketing 12.2, 5–15.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL `http://www.R-project.org/`

Sardianou, E., 2007. Estimating energy conservation patterns of greek households. Energy Policy 35, 3778 – 3791.

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

Train, K., 1985. Discount rates in consumers' energy related decisions: A review of the literature. Energy 10(12), 1243 – 1253.

Viechtbauer, W., 2010. Conducting meta-analyses in r with the metafor package. Journal of Statistical Software 36 (3).

Walsh, M. J., 1989. Energy tax credits and housing improvement. Energy Economics 11 (4), 275–284.

Ward, D. O., Clark, C. D., Jensen, K. L., Yen, S. T., Russell, C. S., 2011. Factors influencing willingness-to-pay for the energy star® label. Energy Policy 39 (3), 1450–1458.

Wilhite, H., Ling, R., 1995. Measured energy savings from a more informative energy bill. Energy and buildings 22 (2), 145–155.

Zivin, J. G., Novan, K., 2015. Upgrading efficiency and behavior: Electricity savings from residential weatherization programs. Tech. rep., UC Davis.

Appendix

2.A Effect of income on energy savings

This point becomes more obvious when considering the Slutzky equation in elasticities: $\varepsilon_{i,x,p_x} = \varepsilon^c_{i,x,p_x} - s_{i,x}\varepsilon_{i,x,m_i}$, where ε_{i,x,p_x} is the uncompensated (Marshallian) price elasticity for energy services of household i; ε^c_{i,x,p_x} is the compensated (Hicksian) price elasticity for energy services of household i; $s_{i,x}$ is household i's budget share spent on energy services; and ε_{i,x,m_i} is the income elasticity for energy services of household i. As we assume that energy services are a normal good, the second term on the right will be positive, i.e., both terms move in the same direction. Engel's law states that the budget share spent on a normal commodity is lower the wealthier the household, hence, the higher y the lower $s_{i,x}$. Consequently, under the assumption of constant Hicksian price elasticities and constant income elasticities across all households, the Marshallian price elasticity will be smaller in absolute terms for wealthier households due to the lower $s_{i,x}$.

However, relaxing the assumption of constant compensated price elasticities and constant income elasticities across all income classes may either reinforce or weaken this effect. In the case of the income elasticity

$$\varepsilon_{y,i} = \frac{\partial x_{e,i}}{\partial y_i}\frac{y_i}{p_e x_i} \cdot p_e = \frac{\partial x_{e,i}}{\partial y_i}\frac{1}{s_{e,i}} \cdot p_e \tag{2.8}$$

the overall change depends on whether $\frac{\partial x_{e,i}}{\partial y_i}$ decreases faster than $\frac{1}{s_{e,i}}$ or vice versa, i.e., the overall change depends on the shape of the Engel curve. On the other hand, there is no theoretical foundation whatsoever for the effect of income on the Hicksian price elasticity of demand. Hence, the outcome of relaxing the assumption of a constant compensated demand elasticity remains an empirical question.

2.B Additional Tables and Figures

Table 2.6: Leave-one-out analysis

(a) Income, binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
1	0.00063	0.00008	8.01676	0.00000	0.00048	0.00079
2	0.00061	0.00008	7.64943	0.00000	0.00045	0.00077
3	0.00063	0.00008	7.92462	0.00000	0.00047	0.00078
4	0.00315	0.00065	4.83720	0.00000	0.00187	0.00442
5	0.00063	0.00008	8.01205	0.00000	0.00048	0.00079
6	0.00064	0.00008	8.04016	0.00000	0.00048	0.00079
7	0.00064	0.00008	8.02993	0.00000	0.00048	0.00079
8	0.00064	0.00008	8.03505	0.00000	0.00048	0.00079
9	0.00064	0.00008	8.03296	0.00000	0.00048	0.00079
10	0.00064	0.00008	8.03281	0.00000	0.00048	0.00079
11	0.00063	0.00008	8.02365	0.00000	0.00048	0.00079
12	0.00064	0.00008	8.03458	0.00000	0.00048	0.00079
13	0.00064	0.00008	8.03274	0.00000	0.00048	0.00079
14	0.00064	0.00008	8.02861	0.00000	0.00048	0.00079
15	0.00063	0.00008	8.00698	0.00000	0.00048	0.00079
16	0.00064	0.00008	8.04847	0.00000	0.00048	0.00079
17	0.00064	0.00008	8.03839	0.00000	0.00048	0.00079
18	0.00064	0.00008	8.03509	0.00000	0.00048	0.00079
19	0.00064	0.00008	8.03047	0.00000	0.00048	0.00079

(b) Income, continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
1	0.00098	0.00100	0.98302	0.32560	-0.00097	0.00293
6	-0.00086	0.01284	-0.06720	0.94643	-0.02602	0.02430
7	0.00101	0.00100	1.00876	0.31309	-0.00095	0.00297

(c) Mid-age (45), binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
1	0.00615	0.00081	7.59275	0.00000	0.00456	0.00774
2	0.00735	0.00091	8.10567	0.00000	0.00558	0.00913
3	0.00623	0.00082	7.62539	0.00000	0.00463	0.00783
5	0.00626	0.00081	7.72218	0.00000	0.00467	0.00785
6	0.00603	0.00081	7.42493	0.00000	0.00444	0.00762
7	0.00616	0.00081	7.60591	0.00000	0.00458	0.00775
8	0.00616	0.00081	7.59782	0.00000	0.00457	0.00774
9	0.00616	0.00081	7.60238	0.00000	0.00457	0.00775
10	0.00616	0.00081	7.60024	0.00000	0.00457	0.00775
11	0.00617	0.00081	7.61152	0.00000	0.00458	0.00776
12	0.00761	0.00086	8.83302	0.00000	0.00592	0.00930
13	0.00555	0.00083	6.66115	0.00000	0.00391	0.00718
14	0.00166	0.00128	1.29973	0.19369	-0.00084	0.00416
16	0.00615	0.00081	7.59127	0.00000	0.00456	0.00774
19	0.00628	0.00081	7.72119	0.00000	0.00468	0.00787

(d) Mid-age (45), continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
6	-0.06035	0.01267	-4.76252	0.00000	-0.08518	-0.03551
7	0.24500	0.13100	1.87023	0.06145	-0.01176	0.50176

(e) Old-age (70), binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
1	0.00397	0.00198	1.99944	0.04556	0.00008	0.00786
2	0.00753	0.00699	1.07746	0.28128	-0.00617	0.02123
3	0.00377	0.00204	1.84696	0.06475	-0.00023	0.00778
5	0.00455	0.00199	2.29254	0.02187	0.00066	0.00845
6	0.00318	0.00199	1.59557	0.11059	-0.00073	0.00708
7	0.00399	0.00198	2.01377	0.04403	0.00011	0.00788
8	0.00398	0.00198	2.00603	0.04485	0.00009	0.00786
9	0.00399	0.00198	2.01071	0.04436	0.00010	0.00787
10	0.00398	0.00198	2.00852	0.04459	0.00010	0.00787
11	0.00401	0.00198	2.01978	0.04341	0.00012	0.00789
12	0.00399	0.00198	2.01050	0.04438	0.00010	0.00787
13	0.00399	0.00198	2.00961	0.04447	0.00010	0.00787
14	0.00399	0.00198	2.01012	0.04442	0.00010	0.00787
16	0.00393	0.00198	1.97995	0.04771	0.00004	0.00782
19	0.00421	0.00199	2.11283	0.03462	0.00030	0.00811

(f) Old-age (70), continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
2	-0.08634	0.02759	-3.12920	0.00175	-0.14042	-0.03226
3	-0.06231	0.02744	-2.27085	0.02316	-0.11608	-0.00853
4	-0.05229	0.02638	-1.98173	0.04751	-0.10400	-0.00057
5	-0.04025	0.02670	-1.50746	0.13169	-0.09257	0.01208
6	-0.05711	0.02676	-2.13403	0.03284	-0.10956	-0.00466
7	0.26683	0.05306	5.02929	0.00000	0.16285	0.37082

(g) HH size, binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
7	0.00314	0.00273	1.14891	0.25059	-0.00221	0.00849
8	0.00337	0.00270	1.24858	0.21182	-0.00192	0.00866
9	0.00344	0.00270	1.27683	0.20166	-0.00184	0.00873
10	0.00285	0.00270	1.05561	0.29115	-0.00244	0.00814
11	0.00258	0.00271	0.95135	0.34143	-0.00273	0.00789
12	0.00282	0.00315	0.89344	0.37162	-0.00336	0.00899
13	0.00387	0.00310	1.25004	0.21129	-0.00220	0.00995
14	0.00467	0.00310	1.50657	0.13192	-0.00141	0.01075
15	0.00217	0.00293	0.73954	0.45958	-0.00357	0.00790
19	0.00307	0.00269	1.14357	0.25280	-0.00219	0.00834

(h) HH size, continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
2	-0.03769	0.04037	-0.93364	0.35049	-0.11682	0.04143
3	-0.08889	0.03886	-2.28715	0.02219	-0.16506	-0.01272
4	-0.09023	0.03122	-2.89025	0.00385	-0.15141	-0.02904
5	-0.06592	0.03229	-2.04114	0.04124	-0.12921	-0.00262

(i) Education, binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
7	0.02096	0.01245	1.68335	0.09231	-0.00344	0.04537
8	0.02396	0.00988	2.42587	0.01527	0.00460	0.04331
9	0.02681	0.00967	2.77253	0.00556	0.00786	0.04576
10	0.02655	0.00995	2.66739	0.00764	0.00704	0.04605
11	0.02574	0.01066	2.41493	0.01574	0.00485	0.04664
19	0.02460	0.00931	2.64137	0.00826	0.00635	0.04286

(j) Education, continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
6	-0.00796	0.02154	-0.36972	0.71159	-0.05017	0.03425
7	0.22250	0.13800	1.61232	0.10689	-0.04798	0.49298

Table 2.8: The influence of moderators using permutation tests (experimental study and investment level)

(a) Income, binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.01	0.01	1.00	0.32	-0.01	0.02
investment: 2	-0.01	0.01	-0.92	0.36	-0.02	0.01

(b) Income, continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.70	0.28	2.46	1.00	0.14	1.25
experiment: 1	-0.70	0.28	-2.48	0.67	-1.25	-0.15
investment: 2	0.00	0.01	0.26	1.00	-0.02	0.03

(c) Mid-age (45), binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.01	0.00	8.21	0.23	0.01	0.01
investment: 2	-0.01	0.00	-3.21	0.45	-0.01	-0.00

(d) Old-age (70), binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-0.01	0.08	-0.19	0.98	-0.16	0.13
investment: 2	0.02	0.08	0.24	0.88	-0.13	0.17

(e) Old-age (70), continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-0.23	0.16	-1.49	0.77	-0.54	0.07
experiment: 1	0.09	0.15	0.58	0.90	-0.21	0.39
investment: 2	0.49	0.14	3.37	0.27	0.20	0.77

(f) Education, binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.02	0.01	2.64	0.50	0.01	0.04
investment: 2	0.12	0.16	0.73	0.17	-0.20	0.43

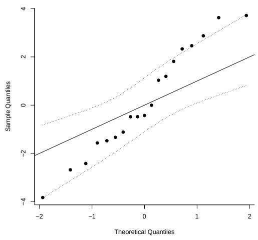
(g) Education, continuous outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	-0.01	0.02	-0.37	1.00	-0.05	0.03
investment: 2	0.23	0.14	1.65	1.00	-0.04	0.50

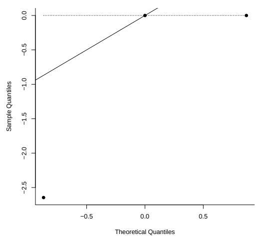
(h) HH size, binary outcome

	estimate	se	zval	pval	ci.lb	ci.ub
intrcpt	0.00	0.00	0.70	0.89	-0.00	0.01
investment: 2	0.01	0.01	0.96	0.24	-0.01	0.02

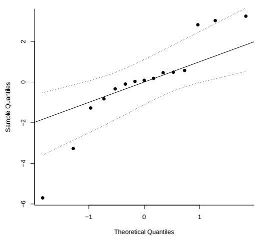
experiment: 0 = no experimental data, 1 = experimental data
investment: 1 = medium scale monetary investment (appliances), 2 = large scale monetary investments (retrofit measures)



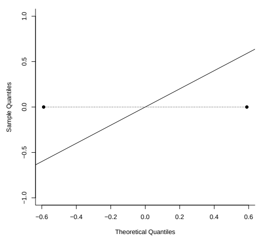
(a) Income, binary outcome



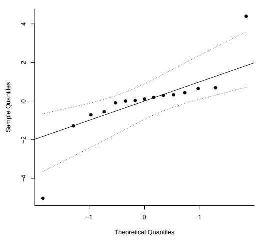
(b) Income, continuous outcome



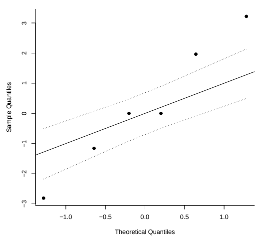
(c) Mid-age (45), binary outcome



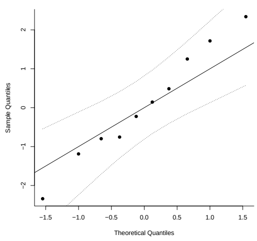
(d) Mid-age (45), continuous outcome



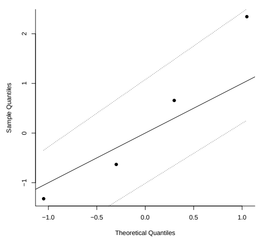
(e) Old-age (70), binary outcome



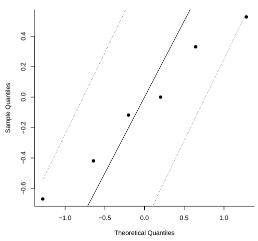
(f) Old-age (70), continuous outcome



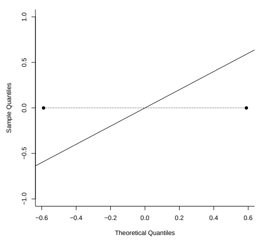
(g) HH size, binary outcome



(h) HH size, continuous outcome



(i) Education, binary outcome



(j) Education, continuous outcome

Figure 2.11: Residual Q-Q plots - regressions with moderators

Chapter 3

Interval Regression with Sample Selection

Arne Henningsen, Sebastian Petersen, and Géraldine Henningsen

This vignette derives the econometric specification used to program the Interval Regression estimator with sample selection as part of the sampleSelection package (Toomet and Henningsen, 2008) in the statistical software R (R Core Team, 2017). The specification finds a practical application in Petersen et al. (2017) (see Chapter 4).

3.1 Model Specification

We formulate a general specification of an interval regression model with sample selection as follows:

$$y_i^{S*} = \boldsymbol{\beta}^{S'} \boldsymbol{x}_i^S + \varepsilon_i^S \tag{3.1}$$

$$y_i^S = \begin{cases} 0 & \text{if } y_i^{S*} \leq 0 \\ 1 & \text{otherwise} \end{cases} \tag{3.2}$$

$$y_i^{O*} = \boldsymbol{\beta}^{O'} \boldsymbol{x}_i^O + \varepsilon_i^O \tag{3.3}$$

$$y_i^O = \begin{cases} \text{unknown} & \text{if } y_i^S = 0 \\ 1 & \text{if } \alpha_1 < y_i^{O*} \leq \alpha_2 \text{ and } y_i^S = 1 \\ 2 & \text{if } \alpha_2 < y_i^{O*} \leq \alpha_3 \text{ and } y_i^S = 1 \\ \vdots & \\ M & \text{if } \alpha_M < y_i^{O*} \leq \alpha_{M+1} \text{ and } y_i^S = 1 \end{cases} \tag{3.4}$$

$$\begin{pmatrix} \varepsilon_i^S \\ \varepsilon_i^O \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix} \right), \tag{3.5}$$

where subscript i indicates the observation, y_i^{O*} is a latent outcome variable, y_i^O is a partially observed categorical variable that indicates in which interval y_i^{O*} lies, M is the number of intervals, $\alpha_1, \ldots, \alpha_{M+1}$ are the boundaries of the intervals (whereas frequently but not necessarily $\alpha_1 = -\infty$ and $\alpha_{M+1} = \infty$), y_i^S is a binary variable that indicates whether

y_i^O is observed, y_i^{S*} is a latent variable that indicates the "tendency" that y_i^S is one, $\boldsymbol{x}_i^S$ and $\boldsymbol{x}_i^O$ are (column) vectors of explanatory variables for the selection equation and outcome equation, respectively, $\boldsymbol{\varepsilon}_i^S$ and $\boldsymbol{\varepsilon}_i^O$ are random disturbance terms that have a joint bivariate normal distribution, and $\boldsymbol{\beta}^S$ and $\boldsymbol{\beta}^O$ are (column) vectors and ρ and σ are scalars of unknown model parameters.

3.2 Log-Likelihood Function

Based on the general specification above we derive probabilities for the different possible outcomes, and the log-likelihood function for the maximum likelihood estimator.

The probability that y_i^O is unobserved is:

$$P\left(y_i^S = 0\right) = P\left(y_i^{S*} \leq 0\right) \tag{3.6}$$

$$= P\left(\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \varepsilon_i^S \leq 0\right) \tag{3.7}$$

$$= P\left(\varepsilon_i^S \leq -\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right) \tag{3.8}$$

The probability that y_i^O is observed and indicates that y_i^{O*} lies in the mth interval is:

$$P\left(y_i^S = 1 \wedge y_i^O = m\right) = P\left(y_i^{S*} > 0 \wedge \alpha_m < y_i^{O*} \leq \alpha_{m+1}\right) \tag{3.9}$$

$$= P\left(\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \varepsilon_i^S > 0 \wedge \alpha_m < \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O + \varepsilon_i^O \leq \alpha_{m+1}\right) \tag{3.10}$$

$$= P\left(\varepsilon_i^S > -\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S \wedge \alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O < \varepsilon_i^O \leq \alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O\right) \tag{3.11}$$

The log-likelihood contribution of the ith observation is:

$$
\begin{aligned}
\ell_i =& (1-y_i^S)\ln\left[\Phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\right] \\
&+ \sum_{m=1}^{M} y_i^S(y_i^O = m)\ln\left[\Phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)\right. \\
&\left. -\Phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)\right],
\end{aligned}
\tag{3.12}
$$

where $\Phi(.)$ indicates the cumulative distribution function of the univariate standard normal distribution and $\Phi_2(.)$ indicates the cumulative distribution function of the bivariate standard normal distribution.

3.3 Restricting coefficients ρ and σ

The parameter ρ needs to be in the interval $(-1,1)$. In order to restrict ρ to be in this interval, we estimate $\operatorname{arctanh}(\rho)$ instead of ρ so that the derived parameter $\rho = \tanh(\operatorname{arctanh}(\rho))$ is always in the interval $(-1,1)$. We use the delta method to calculate approximate standard errors of the derived parameter ρ, whereas the corresponding element of the Jacobian matrix is:

$$
\frac{\partial \tanh(\operatorname{arctanh}(\rho))}{\partial \operatorname{arctanh}(\rho)} = \frac{\partial \rho}{\partial \operatorname{arctanh}(\rho)} = (1+\rho^2) \tag{3.13}
$$

The parameter σ needs to be strictly positive, i.e. $\sigma > 0$. In order to restrict σ to be strictly positive, we estimate $\log(\sigma)$ instead of σ or σ^2 so that the derived parameters $\sigma = \exp(\log(\sigma))$ and $\sigma^2 = \exp(2\ \log(\sigma))$ are always strictly positive. We use the delta method to calculate approximate standard errors of the derived parameters σ and σ^2, whereas the corresponding elements of the Jacobian matrix are:

$$\frac{\partial \exp(\log(\sigma))}{\partial \log(\sigma)} = \exp(\log(\sigma)) = \sigma \tag{3.14}$$

$$\frac{\partial \exp(2 \log(\sigma))}{\partial \log(\sigma)} = 2 \exp(2 \log(\sigma)) = 2 \sigma^2 \tag{3.15}$$

3.4 Gradients of the CDF of the bivariate standard normal distribution

In order to facilitate the calculation of the gradients of the log-likelihood function, we calculate the partial derivatives of the cumulative distribution function (CDF) of the bivariate standard normal distribution:

$$\Phi_2(x_1, x_2, \rho) = \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} \phi_2(a_1, a_2, \rho)\, da_1\, da_2, \tag{3.16}$$

where $\phi_2(.)$ is the probability density function (PDF) of the bivariate standard normal distribution:

$$\phi_2(x_1, x_2, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp\left(-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)}\right) \tag{3.17}$$

In the following, we check equation (3.17) by a simple numerical example:

```
> library( "mvtnorm" )
> library( "maxLik" )
> x1 <- 0.4
> x2 <- -0.3
> rho <- -0.6
> sigma <-  matrix( c( 1, rho, rho, 1 ), nrow = 2 )
> dens <- dmvnorm( c( x1, x2 ), sigma = sigma )
```

```
> print( dens )
[1] 0.1831324
> all.equal( dens, ( 2 * pi * sqrt( 1 - rho^2 ) )^(-1) *
+    exp( - ( x1^2 - 2 * rho * x1 * x2 + x2^2 ) / ( 2 * ( 1 - rho^2 ) ) ) )
[1] TRUE
```

3.4.1 Gradients with respect to the limits (x_1 and x_2)

$$\begin{aligned}
\frac{\partial \Phi_2(x_1, x_2, \rho)}{\partial x_2} &= \int_{-\infty}^{x_1} \phi_2(a_1, x_2, \rho)\, da_1 && (3.18)\\
&= \int_{-\infty}^{x_1} \phi(a_1|x_2, \rho)\phi(x_2)\, da_1 && (3.19)\\
&= \int_{-\infty}^{x_1} \tilde{\phi}\left(a_1, \rho x_2, 1-\rho^2\right) \phi(x_2)\, da_1 && (3.20)\\
&= \int_{-\infty}^{x_1} \phi\left(\frac{a_1 - \rho x_2}{\sqrt{1-\rho^2}}\right) \left(\sqrt{1-\rho^2}\right)^{-1} \phi(x_2)\, da_1 && (3.21)\\
&= \int_{-\infty}^{x_1} \phi\left(\frac{a_1 - \rho x_2}{\sqrt{1-\rho^2}}\right) \left(\sqrt{1-\rho^2}\right)^{-1} da_1\, \phi(x_2) && (3.22)\\
&= \int_{-\infty}^{\frac{x_1 - \rho x_2}{\sqrt{1-\rho^2}}} \phi(a_1)\, da_1\, \phi(x_2) && (3.23)\\
&= \Phi\left(\frac{x_1 - \rho x_2}{\sqrt{1-\rho^2}}\right) \phi(x_2), && (3.24)
\end{aligned}$$

where $\tilde{\phi}(\ ,\mu,\sigma^2)$ indicates the density function of a normal distribution with mean μ and variance σ^2.

In the following, we use the same simple numerical example as in the beginning of section 3.4 to check the above derivations. First, we check whether the PDF of the bivariate standard normal distribution, i.e. $\phi_2(x_1, x_2, \rho)$ (part of equation 3.18), is equal to $\tilde{\phi}\left(x_1, \rho x_2, 1-\rho^2\right)\phi(x_2)$ (part of equation 3.20) and equal to $\phi\left((x_1 - \rho x_2)/(\sqrt{1-\rho^2})\right)$ $\left(\sqrt{1-\rho^2}\right)^{-1} \phi(x_2)$ (part of equations 3.21 and 3.22):

```
> all.equal( dens, dnorm( x1, rho * x2, sqrt( 1 - rho^2 ) ) * dnorm(x2) )

[1] TRUE

> all.equal( dens, ( dnorm( ( x1 - rho * x2 ) / sqrt( 1 - rho^2 ) ) /
+    sqrt( 1 - rho^2 ) ) * dnorm(x2) )

[1] TRUE
```

In the following, we will numerically calculate the derivative of the cumulative distribution function of the bivariate normal distribution (equation 3.16) with respect to x_2 and check whether this partial derivative is equal to the right-hand sides of equations 3.18, 3.21, 3.22, and 3.24:

```
> funX2 <- function( a2 ) {
+    prob <- pmvnorm( upper = c( x1, a2 ), sigma = sigma )
+    return( prob )
+ }
> grad <- c( numericGradient( funX2, x2 ) )
> print( grad )

[1] 0.2320142

> funX1 <- function( a1 ) {
+    dens <- rep( NA, length( a1 ) )
+    for( i in 1:length( a1 ) ) {
+       dens[i] <- dmvnorm( c( a1[i], x2 ), sigma = sigma )
+    }
+    return( dens )
+ }
> all.equal( grad, integrate( funX1, lower = -Inf, upper = x1 )$value )
```

```
[1] TRUE

> funX1a <- function( a1 ) {
+    dens <- rep( NA, length( a1 ) )
+    for( i in 1:length( a1 ) ) {
+       dens[i] <- ( dnorm( ( a1[i] - rho * x2 ) / sqrt( 1 - rho^2 ) ) /
+             sqrt(1-rho^2) ) * dnorm(x2)
+    }
+    return( dens )
+ }
> all.equal( grad, integrate( funX1a, lower = -Inf, upper = x1 )$value )

[1] TRUE

> funX1b <- function( a1 ) {
+    dens <- rep( NA, length( a1 ) )
+    for( i in 1:length( a1 ) ) {
+       dens[i] <- dnorm( ( a1[i] - rho * x2 ) / sqrt( 1 - rho^2 ) ) /
+          sqrt(1-rho^2)
+    }
+    return( dens )
+ }
> all.equal( grad,
+    integrate( funX1b, lower = -Inf, upper = x1 )$value * dnorm(x2) )

[1] TRUE

> all.equal( grad,
+    pnorm( ( x1 - rho * x2 ) / sqrt( 1 - rho^2 ) ) * dnorm( x2 ) )

[1] TRUE
```

3.4.2 Gradients with respect to the coefficient of correlation (ρ)

$$\frac{\partial \Phi_2(x_1, x_2, \rho)}{\partial \rho} \tag{3.25}$$

$$= \frac{\partial \left[\int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \phi_2(a_1, a_2, \rho)\, da_2\, da_1\right]}{\partial \rho} \tag{3.26}$$

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{\partial \phi_2(a_1, a_2, \rho)}{\partial \rho}\, da_2\, da_1 \tag{3.27}$$

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{\partial}{\partial \rho} \left(\frac{\exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right)}{2\pi\sqrt{1-\rho^2}} \right) da_2\, da_1 \tag{3.28}$$

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{1}{2\pi} \frac{\partial}{\partial \rho} \left(\frac{\exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right)}{\sqrt{1-\rho^2}} \right) da_2\, da_1 \tag{3.29}$$

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \frac{1}{2\pi} \left(\frac{\frac{\partial}{\partial \rho}\left(\exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right)\right) \cdot \sqrt{1-\rho^2}}{1-\rho^2} - \frac{\frac{\partial}{\partial \rho}(\sqrt{1-\rho^2}) \cdot \exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right)}{1-\rho^2} \right) da_2\, da_1 \tag{3.30}$$

$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \frac{1}{2\pi} \left(\frac{\frac{\partial}{\partial \rho}\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right) \cdot \exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right) \cdot \sqrt{1-\rho^2}}{1-\rho^2} \right. \tag{3.31}$$

$$\left. - \frac{\left(-\frac{\rho}{\sqrt{1-\rho^2}}\right) \cdot \exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right)}{1-\rho^2} \right) da_2\, da_1$$

$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \frac{1}{2\pi} \left(\frac{\frac{\left(-4\rho(a_1^2 - 2\rho a_1 a_2 + a_2^2) - 2(1-\rho^2)(-2a_1 a_2)\right)}{4(1-\rho^2)^2} \cdot \exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right) \cdot \sqrt{1-\rho^2}}{1-\rho^2} \right. \tag{3.32}$$

$$\left. - \frac{\left(-\frac{\rho}{\sqrt{1-\rho^2}}\right) \cdot \exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right)}{1-\rho^2} \right) da_2\, da_1$$

$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \frac{1}{2\pi} \left(\frac{\frac{\left(-4\rho(a_1^2 - 2\rho a_1 a_2 + a_2^2) - 2(1-\rho^2)(-2a_1 a_2)\right)}{4(1-\rho^2)^2} \cdot \sqrt{1-\rho^2}}{1-\rho^2} - \frac{\left(-\frac{\rho}{\sqrt{1-\rho^2}}\right)}{1-\rho^2} \right) \tag{3.33}$$

$$\cdot \exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right) da_2\, da_1$$

$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \frac{1}{2\pi} \left(\frac{\left(-4\rho(a_1^2 - 2\rho a_1 a_2 + a_2^2) - 2(1-\rho^2)(-2a_1 a_2)\right)}{4(1-\rho^2)^{\frac{5}{3}}} + \frac{\rho}{(1-\rho^2)^{\frac{3}{2}}} \right) \tag{3.34}$$

$$\cdot \exp\left(-\frac{a_1^2 - 2\rho a_1 a_2 + a_2^2}{2(1-\rho^2)}\right) da_2\, da_1$$

$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \frac{1}{2\pi}\left(\frac{\left(-4\rho(a_1^2-2\rho a_1a_2+a_2^2)-2(1-\rho^2)(-2a_1a_2)\right)}{4(1-\rho^2)^{\frac{5}{2}}}+\frac{\rho}{(1-\rho^2)^{\frac{3}{2}}}\right) \cdot \exp\left(-\frac{a_1^2-2\rho a_1a_2+a_2^2}{2(1-\rho^2)}\right) da_2\, da_1 \tag{3.35}$$

$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \frac{1}{2\pi}\left(\frac{\rho}{(1-\rho^2)^{\frac{3}{2}}}-\frac{\rho(a_1^2-\rho a_1a_2+a_2^2)-a_1a_2}{(1-\rho^2)^{\frac{5}{2}}}\right) \cdot \exp\left(-\frac{a_1^2-2\rho a_1a_2+a_2^2}{2(1-\rho^2)}\right) da_2\, da_1 \tag{3.36}$$

$$= \frac{1}{2\pi\sqrt{(1-\rho^2)}}\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\left(\frac{\rho}{1-\rho^2}-\frac{\rho(a_1^2-\rho a_1a_2+a_2^2)-a_1a_2}{(1-\rho^2)^2}\right) \cdot \exp\left(-\frac{a_1^2-2\rho a_1a_2+a_2^2}{2(1-\rho^2)}\right) da_2\, da_1 \tag{3.37}$$

$$= \frac{1}{2\pi\sqrt{(1-\rho^2)}}\int_{-\infty}^{x_1}\left|\left(-\frac{2a_1-2\rho a_2}{2(1-\rho^2)}\right)\cdot\exp\left(-\frac{a_1^2-2\rho a_1a_2+a_2^2}{2(1-\rho^2)}\right)\right|_{-\infty}^{x_2} da_1 \tag{3.38}$$

$$= \frac{1}{2\pi\sqrt{(1-\rho^2)}}\int_{-\infty}^{x_1}\left(\left(-\frac{2a_1-2\rho x_2}{2(1-\rho^2)}\right)\cdot\exp\left(-\frac{a_1^2-2\rho a_1x_2+x_2^2}{2(1-\rho^2)}\right) - \lim_{a_2\to-\infty}\frac{1}{2(1-\rho^2)}\frac{-2a_1+2\rho a_2}{\exp\left(\frac{a_1^2-2\rho a_1a_2+a_2^2}{2(1-\rho^2)}\right)}\right) da_1 \tag{3.39}$$

Applying L'Hospital on the last term leads to

$$=\frac{1}{2\pi\sqrt{(1-\rho^2)}}\int_{-\infty}^{x_1}\left(\left(-\frac{2a_1-2\rho x_2}{2(1-\rho^2)}\right)\cdot\exp\left(-\frac{a_1^2-2\rho a_1x_2+x_2^2}{2(1-\rho^2)}\right)-0\right)\,da_1 \tag{3.40}$$

$$=\frac{1}{2\pi\sqrt{(1-\rho^2)}}\int_{-\infty}^{x_1}\left(-\frac{2a_1-2\rho x_2}{2(1-\rho^2)}\right)\cdot\exp\left(-\frac{a_1^2-2\rho a_1x_2+x_2^2}{2(1-\rho^2)}\right)\,da_1 \tag{3.41}$$

$$=\frac{1}{2\pi\sqrt{(1-\rho^2)}}\left|\exp\left(-\frac{a_1^2-2\rho a_1x_2+x_2^2}{2(1-\rho^2)}\right)\right|_{-\infty}^{x_1} \tag{3.42}$$

$$=\frac{1}{2\pi\sqrt{(1-\rho^2)}}\cdot\exp\left(-\frac{x_1^2-2\rho x_1x_2+x_2^2}{2(1-\rho^2)}\right) \tag{3.43}$$

$$=\phi_2(x_1,x_2,\rho) \tag{3.44}$$

This result is in line with Sibuya (1960) and Sungur (1990).

In the following, we will numerically calculate the derivative of the cumulative distribution function of the bivariate normal distribution (equation 3.26) with respect to ρ and check whether this partial derivative is equal to the right-hand sides of equation 3.44:

```
> # Numerical gradient of the PDF w.r.t. rho
> funrho <- function( p ) {
+    prob <- dmvnorm( x = c( x1, x2 ),
+        sigma = matrix( c( 1, p, p, 1 ), nrow = 2 ) )
+    return( prob )
+ }
> grad <- c( numericGradient( funrho, rho ) )
> print( grad )

[1] -0.1775883

> # Comparison with analytical gradient for rho
```

```
> efun <- exp(-(x1^2 - 2 * rho * x1 * x2 + x2^2)/(2*(1 - rho^2)))
> all.equal( grad,
+ ( (-((2*rho*(-2*rho*x1*x2+x1^2+x2^2) - 2*x1*x2*(1-rho^2)) * efun)/
+     (2*(1-rho^2)^(3/2) )) +
+     ((rho*efun)/(sqrt(1-rho^2))) ) /
+     (2*pi*(1-rho^2)) )
[1] TRUE

> #Eq29
> all.equal(grad,
+     (1/(2*pi)) * (
+     ((((-4*rho*(x1^2-2*rho*x1*x2+x2^2)-2*(1-rho^2)*(-2*x1*x2))/(4*(1-rho^2)^2)) *
+             efun * sqrt(1-rho^2))/(1-rho^2)) -
+             ((-(rho/(sqrt(1-rho^2)))*efun)/(1-rho^2))
+     ))
[1] TRUE

> #Eq33
> all.equal(grad,
+     (1/(2*pi)) *
+     ((rho/((1-rho^2)^(3/2))) - ((rho*(x1^2-rho*x1*x2+x2^2)-x1*x2)/
+             ((1-rho^2)^(5/2)))) * efun
+     )
[1] TRUE

> #Eq34
> all.equal(grad,
```

```
+      (1/(2*pi*sqrt(1-rho^2))) *
+      (((rho/(1-rho^2)) - ((rho*(x1^2-rho*x1*x2+x2^2)-x1*x2)/
+              ((1-rho^2)^2))) * efun)
+      )
[1] TRUE
>
>
> # Numerical gradient of the CDF w.r.t. rho
> cdfRho <- function( p, xa = x1, xb = x2 ) {
+    prob <- pmvnorm( upper = c( xa, xb ),
+       sigma = matrix( c( 1, p, p, 1 ), nrow = 2 ) )
+    return( prob )
+ }
> grad <- c( numericGradient( cdfRho, rho ) )
> print( grad )
[1] 0.1831324
> # comparison with analytical gradient
> all.equal( grad, dmvnorm( x = c( x1, x2 ),
+       sigma = matrix( c( 1, rho, rho, 1 ), nrow = 2 ) ) )
[1] TRUE
> # comparisons with other values
> compDerivRho <- function( xa, xb, p ) {
+    dn <- c( numericGradient( cdfRho, p, xa = xa, xb = xb ) )
+    da <- dmvnorm( x = c( xa, xb ),
```

```
+       sigma = matrix( c( 1, p, p, 1 ), nrow = 2 ) )
+    return( all.equal( dn, da ) )
+ }
> compDerivRho( x1, x2, rho )

[1] TRUE

> compDerivRho( 0.5, x2, rho )

[1] TRUE

> compDerivRho( 2.5, x2, rho )

[1] TRUE

> compDerivRho( x1, -2, rho )

[1] TRUE

> compDerivRho( x1, x2, 0.2 )

[1] TRUE

> compDerivRho( x1, x2, 0.98 )

[1] TRUE
```

3.5 Gradients of the Log-Likelihood Function

3.5.1 Gradients with respect to the parameters in the self-selection decision ($\boldsymbol{\beta}^S$)

First, we use equation 3.24, to determine the derivative of the bivariate standard normal distribution with respect to the parameter $\boldsymbol{\beta}^S$ as part of the loglikelihood function:

$$\frac{\partial \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)}{\partial \beta^S} = \Phi\left(\frac{\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma} - \rho\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S}{\sqrt{1-\rho^2}}\right)\phi(\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S)\cdot\frac{\partial \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S}{\partial \boldsymbol{\beta}^S} \tag{3.45}$$

$$= \Phi\left(\frac{\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma} + \rho\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S}{\sqrt{1-\rho^2}}\right)\phi(\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S)\cdot\boldsymbol{x}_i^S \tag{3.46}$$

Using this result we can now derive the gradient for $\boldsymbol{\beta}^S$ in the log-likelihood function:

$$\begin{aligned}
\frac{\partial \ell_i}{\partial \beta^S} = \frac{\partial}{\partial \beta^S}\Bigg((1-y_i^S)\ln\left[\Phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\right] \qquad\qquad & (3.47)\\
+\sum_{m=1}^{M} y_i^S(y_i^O = m)\ln\Bigg[\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right) &\\
-\Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)\Bigg]\Bigg) &\\
= (1-y_i^S)\frac{\partial}{\partial \beta^S}\left(\ln\left[\Phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\right]\right) \qquad\qquad & (3.48)\\
+\sum_{m=1}^{M} y_i^S(y_i^O = m)\frac{\partial}{\partial \beta^S}\Bigg(\ln\Bigg[\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right) &\\
-\Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)\Bigg]\Bigg) &
\end{aligned}$$

$$= (1-y_i^S)\frac{\phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\cdot(-\boldsymbol{x}_i^S)}{\Phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)} \tag{3.49}$$

$$+\sum_{m=1}^{M} y_i^S(y_i^O=m)\frac{\frac{\partial\Phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S,-\rho\right)}{\partial\beta^S}-\frac{\partial\Phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S,-\rho\right)}{\partial\beta^S}}{\Phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S,-\rho\right)-\Phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S,-\rho\right)}$$

$$= (1-y_i^S)\frac{\phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\cdot(-\boldsymbol{x}_i^S)}{\Phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)} \tag{3.50}$$

$$+\sum_{m=1}^{M} y_i^S(y_i^O=m)\frac{1}{\Phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S,-\rho\right)-\Phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S,-\rho\right)}$$

$$\left(\Phi\left(\frac{\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}+\rho\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S}{\sqrt{1-\rho^2}}\right)\phi\left(\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\cdot\boldsymbol{x}_i^S\right.$$

$$\left.-\Phi\left(\frac{\frac{\alpha_m-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}+\rho\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S)}{\sqrt{1-\rho^2}}\right)\phi\left(\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\cdot\boldsymbol{x}_i^S\right)$$

$$= (1-y_i^S)\frac{\phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\cdot(-\boldsymbol{x}_i^S)}{\Phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)} \tag{3.51}$$

$$+\sum_{m=1}^{M} y_i^S(y_i^O=m)\frac{\left(\Phi\left(\frac{\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}+\rho\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S}{\sqrt{1-\rho^2}}\right)-\Phi\left(\frac{\frac{\alpha_m-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}+\rho\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S}{\sqrt{1-\rho^2}}\right)\right)\phi\left(\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\cdot\boldsymbol{x}_i^S}{\Phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S,-\rho\right)-\Phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S,-\rho\right)}$$

3.5.2 Gradients with respect to the parameters in the outcome decision ($\boldsymbol{\beta}^O$)

Analogous to $\boldsymbol{\beta}^S$ and by using equation 3.24 we derive the gradient of $\boldsymbol{\beta}^O$:

$$\frac{\partial \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S, -\rho\right)}{\partial \beta^O} = \Phi\left(\frac{\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma}\right)\cdot\left(-\frac{\boldsymbol{x}_i^O}{\sigma}\right) \tag{3.52}$$

Using this result we derive the gradient for the outcome parameter $\boldsymbol{\beta}^O$ for the log-likelihood function:

$$\begin{aligned}
\frac{\partial \ell_i}{\partial \beta^O} = \frac{\partial}{\partial \beta^O}\Bigg(& (1-y_i^S)\ln\left[\Phi\left(-\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S\right)\right] \qquad (3.53)\\
& + \sum_{m=1}^{M} y_i^S(y_i^O = m)\ln\left[\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S, -\rho\right)\right.\\
& \left.\left. -\Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S, -\rho\right)\right]\right)\\
= \sum_{m=1}^{M} y_i^S(y_i^O = m)\frac{\partial}{\partial \beta^O}\Bigg(& \ln\left[\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S, -\rho\right)\right. \qquad (3.54)\\
& \left.\left. -\Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S, -\rho\right)\right]\right)
\end{aligned}$$

$$= \sum_{m=1}^{M} y_i^S(y_i^O = m) \frac{\frac{\partial \Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)}{\partial \beta^O} - \frac{\partial \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)}{\partial \beta^O}}{\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right) - \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)} \tag{3.55}$$

$$\begin{aligned}
&= \sum_{m=1}^{M} y_i^S(y_i^O = m) \cdot \frac{1}{\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right) - \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)} \qquad (3.56)\\
&\quad \left(\Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\phi\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\cdot\left(-\frac{\boldsymbol{x}_i^O}{\sigma}\right)\right.\\
&\qquad \left. -\Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\cdot\left(-\frac{\boldsymbol{x}_i^O}{\sigma}\right)\right)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{m=1}^{M} y_i^S(y_i^O = m) \cdot \frac{1}{\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right) - \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)} \qquad (3.57)\\
&\quad \left(\Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\phi\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\right.\\
&\qquad \left. -\Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\right)\cdot\left(-\frac{\boldsymbol{x}_i^O}{\sigma}\right)
\end{aligned}$$

3.5.3 Gradients with respect to the coefficient of correlation (ρ)

Given the result that the derivative of the CDF with respect to ρ is equal to the PDF (see equation 3.44), we can also derive the gradient of the correlation parameter (ρ):

$$
\begin{aligned}
\frac{\partial \ell_i}{\partial \rho} = \frac{\partial}{\partial \rho}\Bigg(&(1-y_i^S)\ln\left[\Phi\left(-\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S\right)\right] \\
&+ \sum_{m=1}^{M} y_i^S(y_i^O=m)\ln\left[\Phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)\right. \\
&\left.\left. -\Phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)\right]\right)
\end{aligned}
\tag{3.58}
$$

$$
\begin{aligned}
= \sum_{m=1}^{M} y_i^S(y_i^O=m)\frac{\partial}{\partial \rho}\Bigg(&\ln\left[\Phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)\right. \\
&\left.\left. -\Phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)\right]\right)
\end{aligned}
\tag{3.59}
$$

$$
= \sum_{m=1}^{M} y_i^S(y_i^O=m)\frac{\phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)-\phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)}{\Phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)-\Phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)}
\tag{3.60}
$$

As we estimate $\operatorname{arctan}(\rho)$ to ensure that ρ is always in the interval (-1,1), we have to adjust the gradient accordingly:

$$
\frac{\partial \ell_i}{\partial\, \operatorname{arctanh}(\rho)} = \frac{\partial \ell_i}{\partial \rho}\frac{\partial \rho}{\partial\, \operatorname{arctanh}(\rho)} = \frac{\partial \ell_i}{\partial \rho}\,(1-\rho^2)
\tag{3.61}
$$

3.5.4 Gradients with respect to the standard deviation used for normalisation (σ)

Finally, we derive the gradient for $\boldsymbol{\sigma}$ in the same way as we did for β^S and β^O, by first calculating the derivative of the bivariate cdf with respect to $\boldsymbol{\sigma}$:

$$\frac{\partial \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)}{\partial \sigma}$$

$$= \Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\cdot\frac{\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O - \alpha_m}{\sigma^2} \tag{3.62}$$

$$\lim_{\alpha_m \to \infty}\frac{\partial \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)}{\partial \sigma}$$

$$= \lim_{\alpha_m \to \infty}\Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\cdot\frac{\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O - \alpha_m}{\sigma^2} \tag{3.63}$$

$$= \Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\lim_{\alpha_m \to \infty}\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\cdot\frac{\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O - \alpha_m}{\sigma^2} \tag{3.64}$$

$$= \Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\lim_{\alpha_m \to \infty}\frac{\frac{\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O - \alpha_m}{\sigma^2}}{\left(\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\right)^{-1}} \tag{3.65}$$

$$= \Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right) \tag{3.66}$$

$$\lim_{\alpha_m \to \infty}\frac{-\frac{1}{\sigma^2}}{-\left(\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\right)^{-2}\left(-\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\frac{1}{\sigma}}$$

$$= \Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\lim_{\alpha_m\to\infty}\frac{-\frac{1}{\sigma}}{\left(\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)\right)^{-1}\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}} \tag{3.67}$$

$$= \Phi\left(\frac{\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S + \rho\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}}\right)\lim_{\alpha_m\to\infty}\frac{-\frac{1}{\sigma}\phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}\right)}{\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}} \tag{3.68}$$

$$= 0 \tag{3.69}$$

Similarly, for the negative limes of a_m:

$$\lim_{\alpha_m\to-\infty}\frac{\partial\Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)}{\partial\sigma} = 0 \tag{3.70}$$

We use this result to calculate the derivative of the log-likelihood function with respect to $\boldsymbol{\sigma}$:

$$\begin{aligned}
\frac{\partial\ell_i}{\partial\sigma} = \frac{\partial}{\partial\sigma}\Bigg(& (1-y_i^S)\ln\left[\Phi\left(-\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S\right)\right] \qquad (3.71)\\
& + \sum_{m=1}^{M} y_i^S(y_i^O = m)\ln\Bigg[\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)\\
& -\Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)\Bigg]\Bigg)\\
= \sum_{m=1}^{M} y_i^S(y_i^O = m)\frac{\partial}{\partial\sigma}\Bigg(& \ln\Bigg[\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right) \qquad (3.72)\\
& -\Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S, -\rho\right)\Bigg]\Bigg)
\end{aligned}$$

$$= \sum_{m=1}^{M} y_i^S (y_i^O = m) \frac{\frac{\partial \Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S, -\rho\right)}{\partial \sigma} - \frac{\partial \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S, -\rho\right)}{\partial \sigma}}{\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S, -\rho\right) - \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S, -\rho\right)} \tag{3.73}$$

$$= \sum_{m=1}^{M} \frac{y_i^S (y_i^O = m)}{\Phi_2\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S, -\rho\right) - \Phi_2\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma}, \boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S, -\rho\right)} \tag{3.74}$$

$$\left(\Phi\left(\frac{\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S + \rho \frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}} \right) \phi\left(\frac{\alpha_{m+1} - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma} \right) \cdot \frac{\boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O - \alpha_{m+1}}{\sigma^2} \right.$$

$$\left. - \Phi\left(\frac{\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S + \rho \frac{\alpha_m - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma}}{\sqrt{1-\rho^2}} \right) \phi\left(\frac{\alpha_m - \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sigma} \right) \cdot \frac{\boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O - \alpha_m}{\sigma^2} \right)$$

As we estimate $log(\boldsymbol{\sigma})$ to ensure that $\boldsymbol{\sigma}$ is strictly positive, we also have to adjust this gradient:

$$\frac{\partial \ell_i}{\partial \log(\sigma)} = \frac{\partial \ell_i}{\partial \sigma} \frac{\partial \sigma}{\partial \log(\sigma)} = \frac{\partial \ell_i}{\partial \sigma} \, \boldsymbol{\sigma} \tag{3.75}$$

Bibliography

Petersen, S., Henningsen, G., Henningsen, A., 2017. Can changes in households' life situations predict participation in energy audits and investments in energy savings?, unpublished Manuscript. Department of Management Engineering, Technical University of Denmark.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL `http://www.R-project.org/`

Sibuya, M., 1960. Bivariate extreme statistics, i. Annals of the Institute of Statistical Mathematics 11 (2), 195 – 210.

Sungur, E., 1990. Dependence information in parameterized copulas. Communications in Statistics - Simulation and Computation 19 (4), 1339 – 1360.

Toomet, O., Henningsen, A., 2008. Sample selection models in r: Package sampleselection. Journal of Statistical Software 27 (7).

Chapter 4

Can Changes in Households' Life Situations Predict Participation in Energy Audits and Investments in Energy Savings?

Sebastian Petersen, Géraldine Henningsen, and Arne Henningsen

Abstract

We suggest changes in the life situation of private home owners as important drivers of the decision to have an energy audit and energy retrofit their home. Earlier findings indicate that households seldom conduct renovations solely for energy efficiency purposes, and that households may be inattentive to the energy properties of their home. Identifying situations in which households are more likely to pay attention to energy matters and take action could therefore be crucial to improve household response to energy efficiency programmes. Using a comprehensive data set on participants of a free energy audit programme and a representative control group of home owners from Sønderborg municipality in Southern Denmark, we investigate the relation between the initial life situation and changes in the life situation during the policy period on a household's decision to (a) participate in the programme and (b) invest into an energy renovation. We estimate the initial decision to have an energy audit as a weighted probit to account for uneven representation of audit and non-audit households in our sample. We then analyse the final investment into energy renovations by estimating an interval regression model with sample selection to account for an interval coded outcome variable and endogenous selection into the audit programme. The results suggest that relocation and entering retirement are significant predictors of joining the audit programme, while including the current life situation variables does not significantly improve the fit of our estimation model. Furthermore, relocation and getting married in the policy period are associated with a higher investment. The results support the notion that timing is important for reaching private home owners with energy efficiency programmes.

4.1 Introduction

When people perceive a change in their life situation, they are more likely to make aspirational changes in their life, sometimes referred to as a Fresh Start Effect (Peetz and

Wilson, 2012; Dai et al., 2014). Based on this concept we hypothesize that entering certain life stages could provide an opportunity for policy programmes to reach out to private households during a time when they are susceptible to new information. We investigate the influence of four life situational variables of home owners – getting married, entering retirement, loosing a job, and moving into a new home – on their propensity to join a free energy audit programme, and the decision to invest in energy renovations. By testing both the effect of the current life situation and changes in the life situation we will be able to test our hypothesis of a Fresh Start Effect.

Policy instruments to promote energy efficiency in private households have been popular since the late 1970's (Hirst et al., 1981; Hirst and Grady, 1983; Hirst and Goeltz, 1985; Wirtshafter, 1985; Fuller et al., 2010). First set in place as a reaction to the oil crises, they experienced revived interest as concerns about climate change started to grow. A large fraction of these programmes comprise some form of a home energy audit, sometimes coupled with a financial investment incentive (subsidies, loans, and investment bonuses (Lutzenhiser, 1993; Wilson et al., 2015), or third-party pre-payment schemes (Miller and Ford, 1985)). Although, the pursuit of post-audit energy investments is generally high (between 50 to 80% of all participating households conduct at least one suggested investment) (e.g. Miller and Ford 1985; Stern 1999; Abrahamse et al. 2007; Frondel and Vance 2013; Palmer et al. 2015), the initial uptake of energy efficiency programmes amongst eligible households is depressingly low (between 0.9 - 8%) (e.g. Fuller et al. 2010; LaRiviere et al. 2014; Fowlie et al. 2015a,b). This seems remarkable as Granade et al. (2009) estimate that by 2020, 29 % of the predicted baseline energy use in buildings could be saved through investments with favourable amortisation periods.

However, it has become increasingly clear over the past three decades that households face many hurdles and unobserved costs in connection with the realisation of these energy saving potentials (e.g., Allcott and Greenstone, 2012; Davis and Metcalf, 2014; Zivin and Novan, 2015; Fowlie et al., 2015b; Wilson et al., 2015). In fact, when taking a closer look at the many factors that need to be in place for a household to finally invest, the low uptake numbers are less surprising. In order to finally conduct an energy efficiency investment, the household needs: to have an energy savings potential that surmounts the investment costs (Stern, 1992; Hoicka et al., 2014; Palmer et al., 2015; Pettifor et al., 2015); to be eligible to the energy efficiency programme in place (Hoicka et al., 2014; Fowlie et al., 2015b); to be informed about the energy efficiency programme (Archer et al., 1984; Wilson and Dowlatabadi, 2007; Pettifor et al., 2015; Palmer and Walls, 2015; Fowlie et al., 2015b) and to trust the source of information and the agency conducting the programme (Stern, 1992; Osterhus, 1997; Hoffman and High-Pippert, 2010); and finally, to be able to conduct the investment financially (Stern et al., 1985; Wilson et al., 2015; Palmer et al., 2015; Pettifor et al., 2015) and with regard to other limiting factors, such as time (Palmer et al., 2015; Fowlie et al., 2015b), non-monetary transaction costs (Stern et al., 1985; Palmer et al., 2015), and personal or social norms (Wilson and Dowlatabadi, 2007).

Recent work by Palmer and Walls (2015) highlights the role of household inattention to the energy properties of their dwelling in the decision to have an energy audit. Home owners that score high on an inattentiveness indicator are much less likely to have an energy audit, which suggests that those most affected by an information gap are in fact less likely to seek out additional information. Inattentiveness to the energy properties of ones home may also explain why energy renovations are rarely motivated by energy efficiency concerns alone, as a UK based study by Wilson et al. (2015) finds that only one in ten energy renovators considered a renovation solely for energy efficiency purposes.

We propose that changes in certain life situations may indicate periods in which households are more likely to be attentive to new information and make changes in their home, which presents a window of opportunity to reach inattentive households. If changes in a household's life situation do play an influential role in the decision to join an energy efficiency programme, marketing strategies for future programmes should not only consider whom they target (Allcott and Greenstone, 2012; Henningsen and Petersen, 2017), but also when they target households (Wilson et al., 2015).

Using the case of the ProjectZero home energy audit programme conducted between September 2010 and June 2013 in Sønderborg municipality in Southern Denmark, we investigate the influence of a home owner's initial life situation and changes in the life situation during the policy period on the decision to participate in the programme. We use data from 978 participating home owners[1] as well as a representative sample of 2,219 non-participating home owners from Sønderborg municipality. In a first analysis we investigate the decision to participate in the audit programme with a weighted probit, accounting for an uneven representation of audit and non-audit households in our sample, due to endogenous sampling. In a separate analysis we model the household's decision making process of first conducting an audit and subsequently investing in energy renovation. Contrary to earlier analyses modelling the same decision making process (e.g., Hartman, 1988; LaRiviere et al., 2014; Frondel and Vance, 2013), we estimate our model as a joint maximum likelihood estimation with a probit selection equation and an interval regression outcome equation in a regression model that accounts for non-random sample selection.

Results indicate that two changes in the life situation – entering retirement and relocation – positively influence the probability of joining the audit programme, while the initial life situational factors do not seem to be associated with the propensity to join the audit programme. For the final investment in energy renovations we find that getting married and relocation are associated with a higher investment sum. The estimation results also suggest that energy use in the form of heat consumption shows a remarkably small association with both having an audit and investing, supporting the notion that households are either inattentive to the energy properties of their home (Palmer and Walls, 2015), or that improvements in energy efficiency are only considered as a

[1] A total number of 1107 audits have been registered, but some entries could not be matched with register data or were removed, as the property was not registered as single-family residential home.

side-benefit in home improvements (Wilson et al., 2015).

The next section develops a theoretical model that describes the audit and investment decision in formal terms and gives a foundation for the need of joint estimation to account for self-selection. This will be followed by a description of the data from the policy intervention, and the econometric approach. Results are then presented for both the audit and investment decision, and the conclusion sums up the findings.

4.2 Theory

We follow Frondel and Vance (2013) and consider a utility-maximising household that faces the decision to invest in retrofitting measures (e.g., improve wall insulation, change windows, etc.) by either using an energy audit, or not. Hence, the household is faced with two decisions: the decision to request an energy audit and the decision to make an investment, which also comprises the choice to invest nothing, i.e., keep the status quo.

At the outset, we model the dichotomous decision of household i ($i = 1,\ldots,n$) to have an energy audit. Household i has the following baseline characteristics $\mathscr{B}_i = \{w_i, y_i, l_i, \rho_i\}$, where w_i is the wealth of the household, y_i is its income, l_i is the household's leisure time, and ρ_i are the household's preferences over time, the environment, and all consumption goods, including energy services.

Household i has the following feasible retrofitting options $\mathbf{V}_i = \{v_{i1}, v_{i2}, \ldots, v_{im}\}$, which are determined by the technical condition of the household's dwelling. As conservation effects of different combinations of these single options might be non-additive (e.g., the effect of a new heating system may depend on whether or not the building is re-insulated), we define $\mathscr{J}_i$ as the power set of $\mathbf{V}_i$, i.e., $\mathscr{J}_i = \{J_{ij} : j = 1,\ldots,2^m\} = P(\mathbf{V}_i)$, which contains all 2^m possible combinations of the measures in set $\mathbf{V}_i$, including the empty set $\emptyset$, i.e., the status quo (no investment). Each element, $J_{ij} \in \mathscr{J}_i$, has an *objective* net present value (NPV), based on the following function of its characteristics $\theta_{ij}(\mathscr{B}_i) = e_{ij}(\mathscr{B}_i) - c^I_{ij}(\mathscr{B}_i) +$

$\xi^I_{ij}(\mathscr{B}_i)$, where e_{ij} are the NPV of the future financial gains from a reduced energy consumption, c^I_{ij} are the monetary cost of the investment, and ξ^I_{ij} is the NPV of the monetised non-monetary future costs and benefits of the investment, e.g., better indoor climate or the inconveniences that follow from construction work. All three components of θ_{ij} are functions of the household characteristics in $\mathscr{B}_i$, which allows us to incorporate the household's budget and time restrictions directly into the NPVs of all alternatives $J_{ij} \in \mathscr{J}_i$.

We define the set of *objective* NPVs over all options $J_{ij} \in \mathscr{J}_i$ as $\Theta_i = \{\theta_{i1}(\mathscr{B}_i), \ldots, \theta_{i2^m}(\mathscr{B}_i)\}$. However, we assume that prior to the audit, the household has imperfect information about Θ_i, i.e., it has a *subjective* set of NPVs over all options in $\mathscr{J}_i$, which may both be biased as well as uncertain. To model the effect of imperfect information on Θ_i, we express the household's uncertainty through a prior probability distribution for each of the NPVs $\theta_{ij} \in \Theta_i$, $f_{ij}(\theta_{ij}(\mathscr{B}_i) + b_{ij}, \sigma^2_{ij})$, with $\theta_{ij} + b_{ij} = E[f_{ij}]$ and $\sigma^2_{ij} = VAR[f_{ij}]$, and which follows an unknown distribution and is specific to each investment option and each household, which in the latter case takes the household's prior knowledge into account.[2] In the case of a large $|b_{ij}|$ the expected value of each distribution $\theta_{ij} + b_{ij}$ can be far off θ_{ij}. We define the household's *subjective* set of probability distributions over the NPVs as $\Theta^{Prior}_i = \{f_{i1}, \ldots, f_{i2^m}\}$, with $f_{i\emptyset}(\theta_{i\emptyset} + 0, 0)$, i.e. we assume that the household has perfect information about the status quo situation. However, while the household knows its *subjective* probability distributions over the NPVs, it has no information about the *objective* NPVs, θ_{ij}, nor the size or direction of its biases, b_{ij}.

The household now faces two options: either to choose the option that maximises expected utility from the set Θ^{Prior}_i, which includes choosing the status quo, $\emptyset$, or to conduct an energy audit and then to chose to maximise utilty over the *objective* set Θ_i, which is yet unknown to the household.

For the first option, we define for each of the 2^m investment options, $u(\theta_{ij})$ as the

[2] We do not follow the *common prior* assumption (Harsany, 1955) in standard economic modeling, but assume subjective uncertainty, irrespective of the access to the same information, as in Savage (1954).

Bernoulli utility function that describes the utility the household derives over the different outcomes of θ_{ij}.[3] The household is then able to determine its von Neumann-Morgenstern (vNM) (expected) utility for each option $J_{ij} \in \mathscr{J}_i$ as

$$E[U_{ij}] = \int_{-\infty}^{\infty} u(\theta_{ij}) f(\theta_{ij}) d\theta_{ij}. \tag{4.1}$$

Assuming that all preference conditions over the 2^m investment options are fulfilled[4], the household can now determine which investment option J_{ij} derives the highest vNM utility

$$J_i^* = \underset{j}{\operatorname{argmax}}\{E[U_{ij}] : j = 1, \dots, 2^m\}. \tag{4.2}$$

In the second option, to maximize utility with an audit, the household compares the expected utility from the energy audit,

$$\begin{aligned} E[U_i|A_i = 1] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u\left(\max_j(\theta_{i1}, \dots, \theta_{i2^m}) - c_i^A + \xi_i^A\right) \\ \cdot f(\theta_{i1}, \dots, \theta_{i2^m})\, d\theta_{i2^m} \dots d\theta_{i1}, \end{aligned} \tag{4.3}$$

where $A_i = \{0, 1\}$ indicates the decision to conduct an audit or not, and c_i^A and ξ_i^A are the monetary costs[5] and net-benefits of the audit, respectively, with the alternative to maximise utility over the set Θ_i^{Prior}. When uncertainty is present the value of additional and accurate information consists in shifting towards better choices given a fixed set of options (Hirshleifer and Riley, 1979), hence, the value of the contrived information is the difference in utility between choosing an investment option under uncertainty and choosing an investment option with certainty. Using Jensens's inequality it follows then

[3]The shape of the Bernoulli utility function will determine whether the household acts risk averse, risk neutral, or risk seeking.
[4]The household has a rational relation over Θ_i^{Prior} which implies completeness and transitivity.
[5]Audits can be subsidised to various extents. In our case study, audits are 100 % subsidised.

that

$$E[U_i|A_i = 1] \geq E[U_{iJ_i^*}]. \tag{4.4}$$

which means that as long as the net-gain in utility from conducting the audit is positive, the household will conduct the audit.

We assume that the energy audit is an expert opinion that provides the household with perfect information, which means that the bias and the uncertainty disappear after the audit. Hence, the household is now provided with the *objective* NPVs, Θ_i, and will use these as a base for the investment decision. The household derives the following set of utilities, which are a monotone function of the NPVs, and maximises over the set of utilities in order to chose the option $J_{ij} \in \mathscr{J}_i$ that generates the maximal utility

$$J_i^{A*} = \underset{j}{\operatorname{argmax}}\{U(\theta_{ij}(\mathscr{B}_{ij})) : j = 1, \ldots, 2^m\}. \tag{4.5}$$

Please note, that $\emptyset \in \mathscr{J}_{ij}$ is still an option for the household, i.e., the household can still chose the status quo if it is the utiliy maximising option.

4.3 Data

The empirical analysis is based on data from a policy intervention in Sønderborg municipality in Southern Denmark. In the intervention, single-family home owners were offered free energy audits with the goal to increase investments in energy saving measures[6] . The intervention was designed and carried out by the private-public partnership ProjectZero. The organisation carries out various initiatives with the goal of making the municipality CO_2-neutral by 2029. Energy audits were offered and carried out between September 2010 and June 2013. The intervention was marketed broadly through a wide range of media channels (newspapers, radio, TV, trade shows, etc.), and households self-selected into the energy audit programme.

The data collected in the course of the policy intervention notes when a free energy audit was requested and carried out, as well as details on investment activity prior to the audit. Investment activity after an audit was registered through follow-up phone surveys carried out between Autumn 2012 and Fall 2013. During this period audit participants were called up multiple times, capturing investment decisions at a later time. Overall, 44.6 percent of the audit population have answered at least one survey. When investments were registered (either during the audit or in the follow-up), it was noted what kinds of measures where installed, and how much was invested in Danish Kroner. Although the investment costs were registered based on actual billings, there is some uncertainty about their accuracy, due to black labour and assigning costs when larger renovations were carried out. Nevertheless, we consider the sum invested as an approximation for the extent of the energy renovation activity of a household.

As households self-selected into policy participation, it is likely that the sample that received energy audits is not representative of the population of single family home-owners in the municipality. As a way to address this, a random sample of 2,219 single

[6]These measures also include installation of renewable energy supply, e.g., photovoltaic, which strictly speaking do not qualify as energy saving. We still include them in the analysis, as they constitute energy savings for the commercial energy grid.

Table 4.1: Comparison of means between different samples in 2010 (Note: right hand panel is only preliminary (see Section 4.3.1))

	Audit Decision			Investment Decision		
	No Audit	Audit	Diff.	No Investment	Investment	Diff.
Number of Observations	2,219	978		377	601	
Socio-economic variables						
Gross Family Income (DKK/Year)	621,183.28	671,073.22	49,889.94*	639,937.5	690,463.8	50,526.3*
Property Value (DKK)	1,167,245.35	1,129,142.19	-38,103.16	1,140,730.95	1,121,925	-18,805.95
Family Debt (DKK)	1,093,474.42	1,107,695.95	14,221.53	1,180,010.47	1,062,660.22	-117,350.26
Family Wealth (DKK)	1,690,405.7	2,164,850.56	474,444.87*	2,316,611.55	2,070,337.5	-246,274.05
Share with Higher Education	0.85	0.93	0.08*	0.94	0.92	-0.01
Age of Household Head	51.48	51.33	-0.16	50.37	51.93	1.56
Number of Adults	1.9	1.9	0	1.9	1.9	0
Number of Children	0.89	0.77	-0.12*	0.74	0.79	0.05
House characteristics						
Building Age	60.11	63.88	3.76*	61.1	65.61	4.51
Number of Rooms	4.94	5.17	0.23*	5.2	5.15	-0.05
Size of Dwelling (m^2)	149.69	160.53	10.83*	160.7	160.42	-0.28
Heat Consumption (kWh/Year)	19,341.82	21,429.19	2,087.37*	20,940.86	21,752.1	811.25
Heat Consumption ([kWh/m^2]/Year)	134.3	142.39	8.1*	138.71	144.85	6.14
Share with Oil Heating	0.06	0.08	0.02*	0.07	0.08	0.02
Share with Gas Heating	0.33	0.31	-0.02	0.34	0.29	-0.04
Share with District Heating	0.37	0.33	-0.04*	0.35	0.32	-0.03
City Center Distance (in m)	8,476.19	8,039	-437.19*	7,494.05	8,377.78	883.73*
Life Situation Variables						
Share Married	0.73	0.73	0	0.71	0.75	0.04
Share Retired	0.23	0.21	-0.02	0.23	0.2	-0.03
Share Unemployed	0.02	0.01	0	0.02	0.01	0

Note: "*" indicates a mean difference with p-value < 0.05 based on a two-tailed t-test.

family home owners living in their homes from the same municipality was drawn from register data available through Statistics Denmark (DST). The data from DST contains household and house characteristics for the randomly sampled households, and households that participated in the audit programme. Finally, we obtained data on annual heat consumption from the Danish Ministry of Taxation. The original dataset from the policy intervention contains 1,107 observations. In the process of matching and merging with register data, only 978 could be clearly identified and were registered as single family home owners.[7]

Table 4.1 shows mean values for socio-economic variables, house characteristics and

[7] Since not all home owners were clearly identified, unmatched households may have been sampled from the remaining population. Assuming that all unmatched households are in the remaining population, this could affect around 0.6% of the observations in the random draw.

initial life situation variables for the sample populations in the audit and investment decision in 2010. The left panel of the table reveals that mean values in the audit sample are different from the non-audit sample for most of the variables. Intuitively, the observed differences appear to make sense, with households in the audit sample being richer on average (availability of capital), having less children (lower living costs/more time available), living in older homes that are larger and have more rooms (more possibilities for improvements), and having a higher annual heat consumption (higher savings potential) both in total and per square meter.

Mean differences in the investment sample (right hand panel of Table 4.1) are almost all insignificant, with the exception of Gross Family Income and City Center Distance. Households that make an investment have a higher mean income, indicating that liquidity might play a role in the investment decision, and household distance from the city center correlates with higher investment. Generally, the fact that the non-audit and audit sample are so different, while the non-investment and investment sample are not, is a strong indication that the decision to have an audit is a strong selection mechanism for the decision to make an energy saving investment. The histogram in Figure 4.1 shows the distribution of observations in the different investment intervals in the sample that received an energy audit. Each investment interval represents an investment range of DKK 30,000, with only the final interval unbounded in the positive range.

4.3.1 Note of Caution: Issues with Investment Data

The right hand panel of Table 4.1 only shows preliminary mean statistics, due to data issues that were discovered at a very late stage of the analysis. Unfortunately, these issues could not be fixed before the submission of this thesis. More specifically, it was discovered that database entries that were registered as non-investors in the main database had in fact never replied to any of the follow-up surveys. Hence, we have two different types of zero observations (see Figure 4.1): households that did not invest, and households that

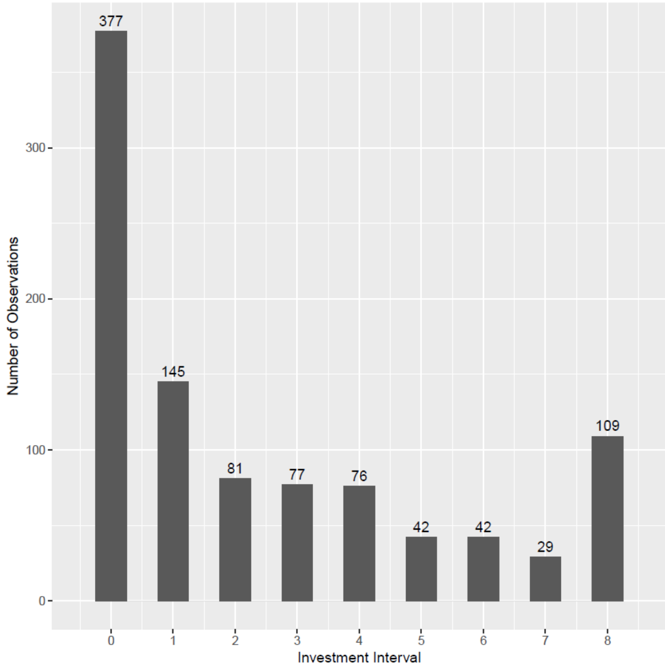


Figure 4.1: Number of observations in the different investment intervals

did not answer the survey. This problem is not easy to address, as the final data analysis was carried out on a server by Statistics Denmark, where it had to be anonymised. This means that we cannot easily match the individual household survey response to the household observation on the server. Furthermore, the new data structure may require a different conceptual approach, which takes into account another selection effect in the second stage, where only a subset of households answered the follow-up survey. While the data issues do not affect the analysis of the decision to join the free energy audit programme, the final paper will only be submitted to an academic journal once the data issues in the investment decision have been resolved and the results have been revised.

4.3.2 Multiple Imputation

There is some degree of missingness in certain registry variables, but one of the main concerns for this analysis is missingness in the values for annual heat consumption at the household level, which is missing for 23.6 % of all observations. As we believe this to be an important variable to explain audit and investment behaviour, we do not consider it an option to leave it out of the analysis. At the same time we are also reluctant to lose a large number of observations through list-wise deletion, as it may create bias in the dataset. We address this through Multiple Imputation (King et al., 2001) of missing values with the R package Amelia II (Honaker et al., 2011).

Based on the assumption that individual variable observations are missing at random and that variables follow a joint multivariate normal distribution, missing values are imputed linearly:

$$\tilde{D_{ij}} = D_{i,-j}\tilde{\beta} + \tilde{\varepsilon}_i, \tag{4.6}$$

where " ˜ " indicates a random draw from a posterior distribution derived from observed data, $\tilde{D_{ij}}$ the variable observation, j, that is imputed for observation, i, $D_{i,-j}$ other variables from the same observation if they are observed, $\tilde{\beta}$ the variable estimate; and $\tilde{\varepsilon}_i$ an error term (King et al., 2001).

In the imputation step we use all the available variable observations used in the analysis, as well as unused register variables available to us (e.g. alternative measures of wealth and income). The random draws introduce a degree of uncertainty in the data, so that every time this process is applied to the data, the imputed values are different. Therefore, the imputation is repeated multiple times, to create a set of m imputed datasets. A rule of thumb is suggested by White et al. (2011), to generate a number of imputed data sets at least as high as the percentage of missing data in the analysis. Using annual heat consumption with a missingness of 23.6 % as guidance, we have generated a total of 25 imputed data sets, and regressions were run on each of these individually.

The estimation results of the regressions were then averaged to generate a single estimation output, and standard errors were calculated as:

$$SE(q) = \sqrt{\frac{1}{m}\sum_{j=1}^{m} SE(q_j)^2 + \frac{\sum_{j=1}^{m}(q_j - \bar{q})^2}{m-1}(1+\frac{1}{m})}, \tag{4.7}$$

with q the parameter estimate, and subscript $j(j = 1,...,m)$ indicating the imputed dataset (Honaker et al., 2011).

As a test of the quality of the imputation we apply overimputation (Honaker et al., 2011) to our data, which treats the existing observations as missing and imputes them based on the estimation model applied to the missing variable observations. Overimputation results for variables with more than 50 missing observations are presented graphically in Figure 4.2. The imputed values are shown as a dot indicating the mean value of 100 repeated imputations, and a line indicating the 90% confidence interval. The diagonal line in each graph indicates perfect imputation where observed and imputed values are identical. Hence, the closer observations are to this line, the better the imputation.

The imputation seems to work quite well for Gross Family Income, Property Value, Family Debt, and Building Sqm, with the general exception of outlier observations (see Figure 4.2). Heat Consumption and Family Wealth are imputed less well, with the overimputed values close to a horizontal trend line. However, the share of overimputed Heat Consumption observations with confidence intervals intersecting the diagonal line is still around 90%, the threshold value given by Honaker et al. (2011) for a good imputation.

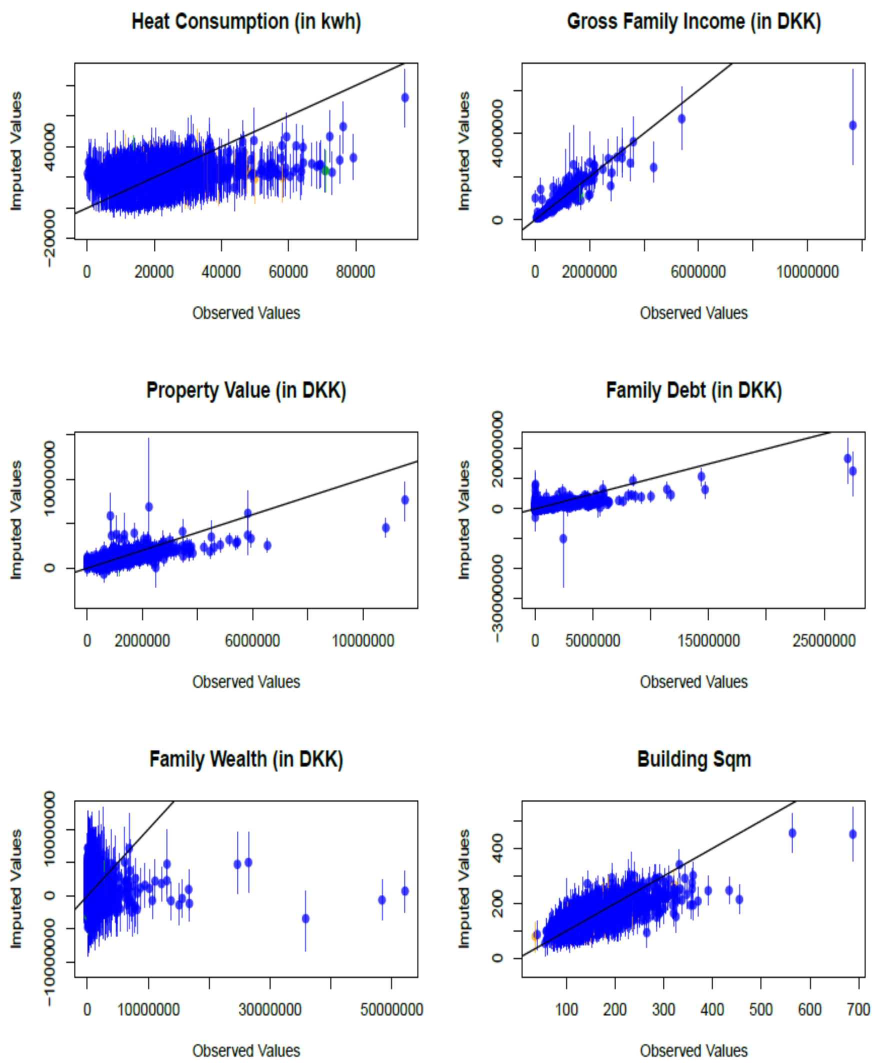


Figure 4.2: Overimputation results for variables with more than 50 missing observations. Note: Results plotted are from 100 repeated imputations for each variable observations; dots indicate the mean, with lines indicating the 90%-confidence interval; the diagonal line in the graphs indicates perfect imputation, where observed and imputed values are equal.

4.4 Econometric Model of Energy Audit and Investment

The policy intervention took place over a time span of three years. Although, data is available for each year individually, this analysis will treat the whole policy period as a cross-sectional data set. Therefore, we estimate participation in the audit programme and investment at any time in the policy period, without considering timing or changes in the explanatory variables over time. To avoid reversed causality issues (e.g. for annual heat consumption or family debt), we use the socio-economic and house characteristics from 2010.

In the data, we observe a household's audit decision through participation in the free audit programme, which corresponds to the audit decision, A_i, in Equation (4.4). In our theoretical model the decision to participate in the policy programme is driven by the benefits and costs of the audit itself, and the benefits and costs of the different retrofitting options, $\boldsymbol{\theta}_{ij}$, which in turn are determined by the household's baseline characteristics, $\mathscr{B}_i$. We include a wide range of socio-economic characteristics to control for the baseline characteristics. Physical characteristics of the dwelling are also included, as this influences the set of available retrofitting measures, and may also be related to the potential benefit of the measures. Our main variables of interest are a group of variables we call life situation variables: marriage, retirement, unemployment, and relocation to a new home. We include these either as a set of dummy variables that captures the household's initial life situation in 2010, or as a set of dummy variables that captures the change in life situation during the policy period, i.e. if people got married, started retirement, got unemployed (at least once), or relocated (at least once) in the policy period (2010 - 2013).

In the theoretical model, Equation (4.4) shows that households already consider potential investments when making the audit decision, which gives rise to selection issues in which the error terms of the two decisions are correlated. The common approach to address sample selection suggested by Heckman (1979) is not suitable for our data, as the final outcome variable (sum invested in energy saving measures) is coded in intervals, and requires non-linear estimation of the second stage (Bhattacharya et al., 2006; Greene and Hensher, 2009; Freedman and Sekhon, 2010).

To address this issue we base the estimation of the investment decision on a probit selection equation in the first stage, and an interval regression in the second stage. Selection issues are accounted for through joint maximum likelihood estimation of the two decisions under the assumption of a bivariate normal distribution of the error terms.

4.4.1 Weighted Probit

While selection bias is an issue for the analysis of the investment decision, this problem does not arise for the analysis of the decision to join the audit programme. However, since only part of our data is based on a random sample of the population, while it contains close to the whole population of the audit participants, the final sample we have is subject to endogenous sampling. This leaves us with the issue of estimating a dichotomous choice, in which the sample cannot be considered representative for the whole population. This can be addressed by estimating a weighted probit model, which is in large parts identical to the general probit model.

The probit equation estimates the decision to request an energy audit, corresponding to A_i in Equation (4.4). This is based on a latent variable that describes the tendency of a household to request an audit, which we can theoretically interpret as the net-benefit of getting an audit:

$$y_i^{S*} = \boldsymbol{\beta}^{S'} \boldsymbol{x}_i^S + \varepsilon_i^S \tag{4.8}$$

$$y_i^S = \begin{cases} 0 & \text{if } y_i^{S*} \leq 0 \\ 1 & \text{otherwise} \end{cases}, \tag{4.9}$$

where subscript i indicates the household observation, y_i^S is a binary variable that indicates whether the household has received an energy audit, y_i^{S*} the latent variable indicating the "tendency" that y_i^S is one, $\boldsymbol{x}_i^S$ is the vector of explanatory variables for the selection equation, $\boldsymbol{\beta}^S$ a vector of parameters, and ε_i^S a random error term.

In the general probit model the individual likelihood contributions to the maximum likelihood estimator are all weighted equally, while in the weighted probit weights are calculated based on the ratio of the population share to the sample share (Solon et al., 2015):

$$w_i(s) = \frac{\bar{N}_s/\bar{N}}{N_s/N}, \tag{4.10}$$

where $\bar{N}_s$ is the total number of observations in the population belonging to sub-group s, $\bar{N}$ the total number of observations in the population, N_s the total number of observations in the sample belonging to sub-group s, and N the total number of observations in the sample.

4.4.2 Interval Regression with Sample Selection

The joint-estimation of the investment decision includes a nested probit estimation, which is then followed by an interval regression, where the latent variable is the financial capital invested in energy saving measures in Danish Kroner. This part of the regression corresponds to the investment decision for households that have received an audit as described in Equation (4.5). An important difference between the theoretical model and the econometric estimation is, however, that our theoretical model describes the decision to invest over a set of discrete choices, while in the estimation the outcome is in terms of the investment sum (though implemented in a discrete choice specification, due to the interval coding). While the data does contain information about how many and what type of measures were implemented, we consider the investment sum a better proxy for the scale of the retro-fitting activity carried out by the household. We define the dependent variable in the selection equation the same as in Equations (4.8) and (4.9), while the outcome dependent variable is specified as follows:

$$y_i^{O*} = \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O + \varepsilon_i^O \tag{4.11}$$

$$y_i^O = \begin{cases} \text{unknown} & \text{if } y_i^S = 0 \\ 1 & \text{if } \alpha_1 < y_i^{O*} \leq \alpha_2 \text{ and } y_i^S = 1 \\ 2 & \text{if } \alpha_2 < y_i^{O*} \leq \alpha_3 \text{ and } y_i^S = 1 \\ \vdots & \\ M & \text{if } \alpha_M < y_i^{O*} \leq \alpha_{M+1} \text{ and } y_i^S = 1 \end{cases}, \tag{4.12}$$

where y_i^{O*} is a latent outcome variable, y_i^O is the partially observed investment sum that indicates in which interval y_i^{O*} lies, M is the number of intervals defined by the boundary values $\alpha_1, \ldots, \alpha_{M+1} = (-\infty, 0,\ 30000,\ 60000,\ 90000,\ 120000,\ 150000,\ 180000,\ 210000, +\infty)$ measured in Danish Kroner, $\boldsymbol{x}_i^O$ is the vector of explanatory variables for the outcome equation, $\boldsymbol{\beta}^O$ is a vector of parameters, and ε_i^O is a random error term following a joint bivariate normal distribution as follows:

$$\begin{pmatrix} \varepsilon_i^S \\ \varepsilon_i^O \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma_2 \\ \rho\sigma_2 & \sigma_2^2 \end{pmatrix} \right), \tag{4.13}$$

with ρ and σ_2 scalars of unknown model parameters.

Based on the estimation of the decision to request an audit in Equation (4.8) we receive the following probability of not receiving an audit (and not observing the investment decision, y_i^O):

$$P\left(y_i^S=0\right)=P\left(y_i^{S*}\leq 0\right) \tag{4.14}$$

$$=P\left(\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S+\varepsilon_i^S\leq 0\right) \tag{4.15}$$

$$=P\left(\varepsilon_i^S\leq -\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S\right). \tag{4.16}$$

Due to the joint distributions of the error terms $\boldsymbol{\varepsilon}_i^S$ and $\boldsymbol{\varepsilon}_i^O$ we receive a positive selection outcome as joint probability of $y_i^S=1$ and $y_i^O=m$:

$$P\left(y_i^S=1\wedge y_i^O=m\right)=P\left(y_i^{S*}>0\wedge \alpha_m<y_i^{O*}\leq \alpha_{m+1}\right) \tag{4.17}$$

$$=P\left(\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S+\varepsilon_i^S>0\wedge \alpha_m<\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O+\varepsilon_i^O\leq \alpha_{m+1}\right) \tag{4.18}$$

$$=P\left(\varepsilon_i^S>-\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S\wedge \alpha_m-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O<\varepsilon_i^O\leq \alpha_{m+1}-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O\right) \tag{4.19}$$

Based on these probabilities we derive the log-likelihood function for each individual household i as:

$$\ell_i=(1-y_i^S)\ln\left[\Phi\left(-\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S\right)\right]+\sum_{m=1}^{M}y_i^S(y_i^O=m)\ln\left[\Phi_2\left(\frac{\alpha_{m+1}-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma_2},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)\right. \tag{4.20}$$
$$\left.-\Phi_2\left(\frac{\alpha_m-\boldsymbol{\beta}^{O'}\boldsymbol{x}_i^O}{\sigma_2},\boldsymbol{\beta}^{S'}\boldsymbol{x}_i^S,-\rho\right)\right],$$

where $\Phi(.)$ indicates the cumulative distribution function of the univariate standard normal distribution and $\Phi_2(.)$ indicates the cumulative distribution function of the bivariate standard normal distribution.

The log-likelihood is similar to an ordered probit with sample selection (theoretically described in Greene and Hensher (2009), an applied example is found in Jimenez and Kugler (1987)) , with the only difference that the boundaries are known.

Exclusion Restriction

In order for the outcome estimation of a selection model to be theoretically identified an exclusion restriction is needed, if one does not want to purely rely on the distributional assumptions of the estimator (Smith and Sweetman, 2016). This exclusion restriction may only affect the outcome of interest through its correlation with the selection decision, i.e. it has to be a sufficiently strong predictor of selection, but must not be correlated with the final outcome. While a variable's predictive qualities on selection can be tested, the latter property cannot be shown in statistical terms. Instead, the case has to be made on theoretical grounds.

In our case we are looking for a variable that predicts selection into the audit programme, but does not play a role in the household investment decision. We believe that exposure to marketing could be such a variable, as it raises attention for the campaign, and chances of participation, but does not change the fundamental properties of the investment decision. Unfortunately, we cannot control for this directly, because marketing was conducted pretty broadly and on multiple channels. Instead, we propose to use distance from the center of the city of Sønderborg, which is the municipal capital. The idea behind this is that households living further away from the city center are more likely to consume local media from neighbouring municipalities that did not advertise the programme, which means they are less likely to be aware of and select into it. The office of ProjectZero is also located in the city center, so people could also be more familiar with the organisation, which may increase participation as well.

Clearly, there are a lot of factors that determine household location in space. Differences in house prices come to mind, which leads to higher income and higher wealth households to live closer to the city, and other socio-economic grouping mechanisms, like education and children. If these socio-economic variables also correlate with the investment decision our exclusion restriction would not work. Thus, we have to assume that we control for the socio-economic mechanisms behind household distance to the city center with our selection of socio-economic variables.

Partial Effects

In the investment data we register households that do not make an investment, which are coded as a zero investment. As the estimation of the interval regression model with sample selection is based on GLM maximum likelihood estimation, we can correct for left-censoring in the same way as in a tobit estimation (Wooldridge, 2010). By doing this we assume that the tendency to make an investment and the final investment sum are influenced by our independent variables in the same way. To overcome this assumption would require a specification with an additional stage that models the selection into the investment category (Sigelman and Zeng, 1999). While theoretically desirable, this would require an additional exclusion restriction, for which we cannot identify a suitable candidate in the data.

Thus, partial effects of the continuous independent variables are calculated as:

$$\begin{aligned}\frac{\partial E(y_i^{O*}|\boldsymbol{x}_i^O, y_i^S=1)}{\partial x_j} =& \frac{\partial P(y_i^{O*}>0|\boldsymbol{x}_i^O, y_i^S=1)}{\partial x_j} \cdot E(y_i^{O*}|y_i^{O*}>0, \boldsymbol{x}_i^O, y_i^S=1) \\ &+ P(y_i^{O*}>0|\boldsymbol{x}_i^O, y_i^S=1) \cdot \frac{\partial E(y_i^{O*}|y_i^{O*}>0, \boldsymbol{x}_i^O, y_i^S=1)}{\partial x_j}.\end{aligned} \tag{4.21}$$

Partial effects for discrete independent variables are calculated based on the discrete changes in the tobit corrected fitted values evaluated for each individual observation.

4.5 Results

In this section we first present and discuss the estimation of the decision to join the policy and have an energy audit. Following this we will direct our attention to the investment decision to investigate determinants of making an energy investment.[8] All estimations were programmed and carried out in the statistical software R (R Core Team, 2017).

4.5.1 Weighted Probit Results

The weighted probit regression was carried out using the 'survey' package (Lumley, 2017), which uses the Horvitz-Thompson estimator instead of Maximum Likelihood estimation. While the general interpretations of the regression output is the same as for a general probit estimated by maximum likelihood, we cannot apply the same statistical tests. Therefore, to compare our different model specifications we use an adjusted Rao-Scott Likelihood-Ratio test (Lumley and Scott, 2014), instead of the standard Likelihood-Ratio test.

Table 4.2 shows the estimated coefficients and standard errors for three different specifications. Panel (1) in Table 4.2 includes socio-economic variables and house characteristics to predict the decision to have an audit. Most of the estimated coefficients in this specification are significant. The second panel adds life situation variables in 2010, to test if being in a certain life situation is associated with audit participation. None of the added initial life situation variables are significant, and an adjusted Rao-Scott Likelihood-Ratio test shows that the addition of these variables does not significantly improve the fit of the model (P-value = 0.73).

Finally, panel (3) shows an alternative specification that instead adds dummy variables for changes in the life situation of households during the policy period. In this estimation, starting retirement is significant at the 10%-level, and moving in the first half of the

[8]Please note that the investment results are only preliminary, until data issues are addressed (see Section 4.3.1).

Table 4.2: Weighted Probit Regression Results for Decision to Join the Audit Programme

	(1)		(2)		(3)	
Variable	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
(Intercept)	-2.9559***	0.3195	-3.0462***	0.3452	-3.0137***	0.3367
Socio-economic Variables						
Gross Family Income	0.0004	0.0005	0.0004	0.0005	0.0004	0.0005
Property Value (in DKK 10.000)	-0.0023***	0.0006	-0.0022***	0.0006	-0.0021***	0.0005
Family Wealth (in DKK 10.000)	0.0008***	0.0002	0.0008***	0.0002	0.0008***	0.0002
Family Debt (in DKK 10.000)	-0.0007**	0.0003	-0.0007**	0.0003	-0.0007**	0.0003
Higher Education (Dummy)	0.3572***	0.0729	0.3628***	0.0732	0.3584***	0.0736
Household Head Age	0.0248***	0.0093	0.0282***	0.0100	0.0260***	0.0100
Household Head Age Squared	-0.0003***	0.0001	-0.0003***	0.0001	-0.0003***	0.0001
Number of Adults	-0.0371	0.0495	-0.0227	0.0609	-0.0399	0.0501
Number of Children	-0.0855***	0.0254	-0.0840***	0.0259	-0.0797***	0.0262
House Characteristics						
Building Age	-0.0011**	0.0005	-0.0011**	0.0005	-0.0010*	0.0005
Building Sqm	0.0057***	0.0011	0.0058***	0.0011	0.0057***	0.0011
Heat Consumption (in 1.000 kWh)	-0.0111	0.0076	-0.0111	0.0076	-0.0113	0.0075
Heat Consumption (kWh) per Building Sqm	0.0032***	0.0010	0.0032***	0.0010	0.0032***	0.0010
City Center Distance (in km)	-0.0232***	0.0049	-0.0231***	0.0049	-0.0227***	0.0049
Oil Heating (Dummy)	0.1345*	0.0801	0.1356*	0.0805	0.1396*	0.0804
Life Situation Variables						
Married			-0.0405	0.0576		
Retired			0.0367	0.0772		
Unemployed			-0.0820	0.1475		
Got Married 2010-2013					0.0857	0.0940
Started Retirement 2010-2013					0.1277*	0.0712
Moved 2010-2011					0.2284***	0.0840
Moved 2012-2013					-0.1143	0.0829

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

policy period is significant at the 1%-level. An adjusted Rao-Scott Likelihood-Ratio test suggests that this specification does provide a better fit for the estimation of participation in the audit programme (P-value = 0.007). Hence, we prefer this specification, and discuss the results below. For a better interpretation of the effects of the different variables we provide Average Partial Effects in Table 4.3.

Our results suggest that both family wealth and family debt are associated with the decision to join the audit programme, while income is not statistically significant. We also find that higher property value has a negative association with the propensity to have an energy audit. This may be due to higher value properties being in better condition both in terms of maintenance, and in terms of thermal properties, reducing the general

Table 4.3: Average Partial Effects for specification (3) in Table 4.2

	Average Partial Effect	Std. error
Socio-economic Variables		
Gross Family Income (in DKK 10.000)	0.00004	0.00005
Property Value (in DKK 10.000)	-0.00022***	0.00006
Family Wealth (in DKK 10.000)	0.00008***	0.00002
Family Debt (in DKK 10.000)	-0.00008**	0.00003
Higher Education (Dummy)	0.02982***	0.00494
Household Head Age	0.00268***	0.00104
Household Head Age Squared	-0.00003***	0.00001
Number of Adults	-0.00412	0.00518
Number of Children	-0.00823***	0.00273
House Characteristics		
Building Age	-0.00010*	0.00006
Building Sqm	0.00059***	0.00011
Heat Consumption (in 1.000 kWh)	-0.00117	0.00078
Heat Consumption (kWh) per Building Sqm	0.00033***	0.00011
Oil Heating (Dummy)	0.01582	0.00995
City Center Distance (in km)	-0.00234***	0.00051
Life Situation Variables		
Got Married 2010-2013	0.00940	0.01091
Started Retirement 2010-2013	0.01428*	0.00857
Lost Job 2010-2013	-0.01555	0.01148
Moved 2010-2011	0.02745**	0.01164
Moved 2012-2013	-0.01090	0.00729

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

need for energy improvements.

The partial effect for higher education is relatively strong, which may partly be explained by higher educated households being better informed through newspapers and other channels, thus having more exposure to the marketing of the policy. An alternative explanations could be that higher education is a correlate of social classes that care about their environmental image, or that households with higher education are generally more patient and therefore apply a lower discount rate to future savings. Household Head Age is estimated both as a linear and quadratic term. Both are significant, but with opposing signs, indicating that age increases the probability of having an audit up to a point, but that at a certain point higher age decreases the likelihood of having an energy audit. The estimates indicate a turning point of the effect of age at around 44.67 years of age. The household composition plays a role, as we see that the number of children in the household is a significant factor. More children reduce the probability of having an audit, while the number of adults in the household is not associated with the audit decision.

The age of a building is negatively associated with the probability to join the audit programme, which is a bit surprising, though the average partial effect is very small. Building square meters and annual heat consumption enter the model as individual parameters and as heat consumption per square meter. Therefore, their average partial effects cannot be interpreted directly, but also need to account for the change in the interaction term. The combined average partial effect of building square meters is 0,00023, which means that an additional square meter increases the probability to have an audit by 0.023 percentage points. Annual heat consumption has a combined average partial effect of 0.0029, i.e. 0.29 percentage points, for a 1,000 kWh increase in consumption, which appears somewhat low, given that the main motivation for policy makers is to encourage energy savings. For oil heating the audit intervention seems to be more on target with regard to the policy intention, as households with oil heating are almost

1.6 percentage points more likely to join the audit programme. Distance from the city centre aims to capture the intensity of marketing exposure to the policy campaign. As such we would expect it to have a negative association with the probability to have an energy audit, which is also the case. The partial effect implies that an additional kilometre of distance decreases the probability of having an audit by 0.2 percentage points on average.

Regarding the dummy variables indicating a change in the life situation, we find that household heads that started retirement or moved in the first half of the policy period are more likely to have an audit, with average partial effects of 1.4 and 2.7 percentage points, respectively. This result supports the notion of a Fresh Start Effect, though the fact that not all the change in life situation variables are significant also highlights that this effect appears to be dependent on the context. An alternative explanation for the influence of relocation could be that relocating households are more likely to engage in other renovation activities, making them more attentive to the potential for energy efficiency improvements from renovation.

Given that relocation in the first half of the policy period shows a relatively strong association with having an audit, it is surprising that moving in the second half has the opposite sign and is not significant. One reason for this could be that households that relocated earlier had a longer exposure to the programme, giving them a longer window of opportunity to join. Another reason could be that the reference group (i.e. households that did not move) is more likely to join in the second half of the policy period compared to the first half. This could be driven by the continuing marketing campaign the non-relocating households are exposed to, which also makes less attentive households aware of the audit programme over time, and relatively more likely to join.

4.5.2 (Preliminary) Interval Regression Results

The maximum likelihood equations for the estimation of the final investment sum were maximised using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm. We also tested the Newton-Raphson (NR) and Berndt–Hall–Hall–Hausman (BHHH) algorithms, but they did not handle the estimations as well as the BFGS algorithm. The NR algorithm performed markedly worse and failed to provide a solution for our final specifications of the interval regression model with sample selection presented below. Results from the BHHH algorithm were generally quite close to the results from BFGS, but the latter was slightly more reliable in converging when applied repeatedly to the full set of imputed data sets (see Section 4.3 regarding the description of the multiple imputation).

In trying different specifications for the estimation of the investments sums we found that the estimations performed much better when applying an inverse hyperbolic sine transformation (asinh) to the interval values in the dependent variable. While somewhat uncommon in economics, the inverse hyperbolic sine transformation is an alternative to the more common logarithmic transformation, but with the advantage of being continuous through zero. (Burbidge et al., 1988) Average partial effects are calculated on the inverse hyperbolic sine transformed latent variable, i.e. estimated investment in energy saving measures. This allows us to interpret the partial effects as percentage changes in the same way as a log-level model specification. To get an idea about numerical values for the partial effects of the significant variables we have created box plots of the partial effects shown in Figure 4.3, with outliers excluded for better readability.

As we described in the previous section, a valid exclusion restriction is crucial for the joint estimation with the interval regression estimator with sample selection. Based on the estimation results from the selection stage we see that City Center Distance, a proxy for marketing intensity, does have the expected sign and is highly significant, which is a necessary condition for its suitability as exclusion restriction. However, we need to ensure that it has strong enough predictive properties to function as instrument. In order to test this we derive F-statistics from a Wald-test of the unrestricted selection stage against a restricted model without the City Center Distance variable. The mean F-statistic for all the estimations run on the imputed data sets is 19.95. According to Smith and Sweetman (2016) this should qualify as strong enough instrument to serve as exclusion restriction, which defines a value of 10 as rule-of-thumb threshold value.

Unfortunately, and despite the positive evaluation of the City Center Distance variable as exclusion restriction, it was not possible to estimate the interval regression with sample selection specification, as none of the algorithms we tried managed to numerically solve the task. Therefore, it was necessary to exclude two additional factors (Higher Education and Building Square-meters), to estimate the model. However, these exclusion restrictions do not have a theoretical foundation, and were selected based on the fact that they were not one of our main variables of interest and the estimation converged properly when excluding them. Therefore, we also estimated a subsample interval regression that does not account for self-selection, but contains the full set of variables, to provide an indication on the robustness of the results.

The estimation results for the selection stage of the interval regression model with sample selection are presented in Table 4.4. In terms of signs and relative magnitude of the coefficients the results are mostly in line with the weighted probit results in Table 4.2. As the weighted probit model is the more appropriate model for investigating the decision to have an audit, we do not further discuss the results of the selection stage here.

The estimation results for the outcome stage are presented in Table 4.5 alongside p-values and average partial effects. The coefficient estimates cannot be directly compared between the two estimations, because they are scaled differently and also depend on the fitted values of each individual observation. Thus, we mainly direct our attention to the significance of the explanatory variables and the average partial effects. Between the two models, results are fairly consistent with most estimated parameters at a similar

Table 4.4: Selection Results for Interval Regression with Sample Selection

Model:	Interval Regression with Sample Selection		
Dependent Variable:	*Audit Decision*		
Variable	Estimate	Std. Error	p-value
(Intercept)	-2.3967***	0.3892	0.0000
Socio-economic Variables			
Gross Family Income (in DKK 10.000)	0.0007	0.0007	0.3193
Property Value (in DKK 10.000)	-0.0023***	0.0005	0.0000
Family Debt (in DKK 10.000)	-0.0010***	0.0003	0.0009
Family Wealth (in DKK 10.000)	0.0011***	0.0002	0.0000
Higher Education (Dummy)	0.4582***	0.0906	0.0000
Household Head Age	0.0364***	0.0124	0.0033
Household Head Age Squared	-0.0004***	0.0001	0.0013
Number of Adults	-0.0562	0.0628	0.3713
Number of Children	-0.0935***	0.0251	0.0002
House Characteristics			
Building Age	-0.0008	0.0006	0.1661
Building Sqm	0.0071***	0.0012	0.0000
Heat Consumption (in 1.000 kwh)	-0.0147*	0.0088	0.0972
Heat Consumption (kwh) per Building Sqm	0.0041***	0.0012	0.0007
City Center Distance (in km)	-0.0273***	0.0054	0.0000
Oil Heating (Dummy)	0.1825*	0.1005	0.0695
Life Situation Variables			
Got Married 2010-2013	0.1086	0.1197	0.3642
Lost Job 2010-2013	-0.2191	0.1866	0.2406
Started Retirement 2010-2013	0.1621*	0.0876	0.0645
Moved 2010-2011	0.2810***	0.1067	0.0085
Moved 2012-2013	-0.1403	0.1092	0.1988
Log Likelihood	-4140.8044		
nObs	3197		

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Table 4.5: Results for estimation of investment sum from Subsample Interval Regression and Interval Regression with sample selection estimation on Investment Intervals

Model:	Subsample Interval Regression			Interval Regression with Sample Selection		
Dependent Variable:	*asinh(Investment Interval)*			*asinh(Investment Interval)*		
Variable	Estimate	p-value	APE	Estimate	p-value	APE
(Intercept)	-12.2489**	0.0287		-17.8598***	0.0028	
Socio-economic Variables						
Gross Family Income (in DKK 10.000)	0.0316**	0.0232	0.0316	0.0329***	0.0077	0.0329
Property Value (in DKK 10.000)	-0.0040	0.5783	-0.0040	-0.0063	0.3944	-0.0063
Family Debt (in DKK 10.000)	-0.0015	0.6659	-0.0015	-0.0031	0.4135	-0.0031
Family Wealth (in DKK 10.000)	-0.0023	0.1954	-0.0023	-0.0011	0.5809	-0.0011
Higher Education (Dummy)	-1.1851	0.4463	-0.9865			
Household Head Age	0.5971***	0.0028	0.6337	0.6513***	0.0017	0.6988
Household Head Age Squared	-0.0052***	0.0080	-0.0052	-0.0058***	0.0044	-0.0058
Number of Adults	-0.6698	0.4606	-0.7318	-0.7442	0.4274	-0.8290
Number of Children	0.1806	0.6202	0.1819	0.0105	0.9782	0.0106
House Characteristics						
Building Age	0.0111	0.1831	0.0111	0.0099	0.2460	0.0099
Building Sqm	-0.0054	0.7013	-0.0054			
Heat Consumption (in 1.000 kwh)	0.1237	0.2570	0.1242	0.1304*	0.0776	0.1308
Heat Consumption (kwh) per Building Sqm	-0.0092	0.5531	-0.0092	-0.0074	0.4153	-0.0074
Oil Heating (Dummy)	2.5040*	0.0702	2.8797	2.7356*	0.0548	2.6721
Life Situation Variables						
Got Married 2010-2013	3.5393**	0.0369	6.8783	3.7362**	0.0348	5.7434
Lost Job 2010-2013	-2.2979	0.4159	-2.5007	-2.6533	0.3777	-2.2967
Started Retirement 2010-2013	0.7518	0.5390	0.5436	0.9502	0.4338	0.5975
Moved 2010-2011	2.3580*	0.0831	2.5649	2.8926*	0.0562	2.9941
Moved 2012-2013	1.0314	0.5159	0.7684	0.8648	0.6272	0.5332
Model Parameters						
sigma	10.2413***	0.0000		10.4377***	0.0000	
rho				0.2273	0.1610	
Log Likelihood	-2304.2303			-4140.8044		
nObs	978			3197		

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

level of p-values and average partial effects in a similar range. The numerical values for the average partial effects plotted in Figure 4.3, however, are consistently higher for the subsample interval regression. This is mainly due to the differences in the intercept, diminishing the overall level of the fitted values in the interval regression with sample selection.

For the interval regression with sample selection we also estimate the rho parameter, which captures the selection effect. The parameter is not significant, indicating that we cannot confirm the presence of a selection effect. However, as we are not completely confident in the set of instruments used, we do not think that this is conclusive evidence of the absence of selection effects. In light of the insignificant rho parameter it makes sense, though, that the interval regression delivers results that are fairly similar to the interval regression with sample selection.

The outcome estimation has generally proven difficult. Only a few factors appear to be associated with the amount that is invested in energy saving measures. This is possibly due to the strong selection effect, which leads to a low degree of variance in most variables of the audit population, which we have already seen in the comparison of means in the different subgroups in Table 4.1. The different models we estimated deliver quantitatively similar results, which gives us some confidence in them with regards to the investment decision of this specific sample. However, given the challenges encountered during estimation the generalizability of the results is put in doubt.

For the financial variables we see a reverse picture of the selection equation, in that Gross Family Income is significant, while the wealth based variables are all insignificant. The association with income is quite small with about 3 percent increase in the investment sum per DKK 10,000 of income. Numerically, the majority of the partial effects in the first boxplot in Figure 4.3 are below DKK 75 for the subsample interval regression and below DKK 25 for the interval regression with sample selection. Age is the only other significant socio-economic variable, with quite a substantial partial effect

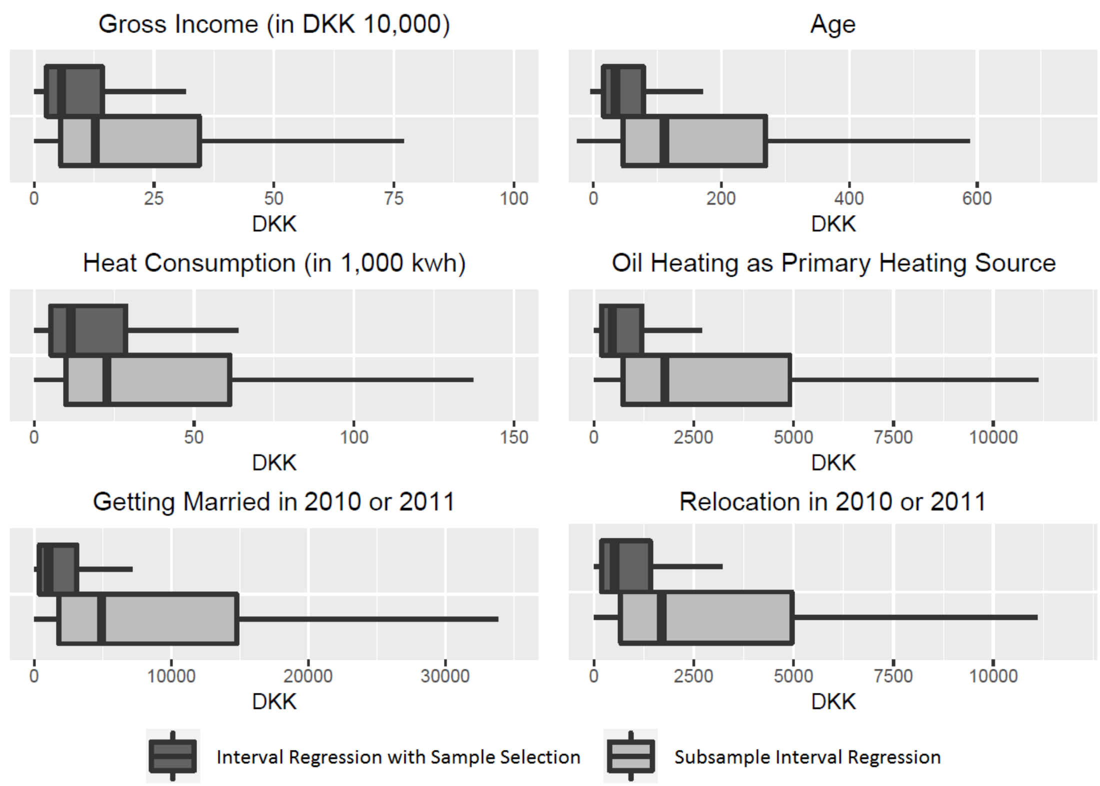


Figure 4.3: Box plots of individual partial effects for significant variables (without outliers)

predicted by the interval regression results (the majority of observations falling between DKK 50 and DKK 300), but in the interval regression with sample selection results the partial effect is much more modest (below DKK 100 for most observations). For some observations the partial effect of age is negative, as the negative quadratic term leads to a turning point of the partial effect of age at 57.41 years for the interval regression, and at 56.14 years for the interval regression with sample selection.

Total heat consumption is only significant at the 10 percent level in the interval regression with sample selection estimation, while in the subsample interval regression it fails to pass at this significance level. The combined average partial effect of heat consumption and heat consumption divided by building square meters is 0.013 in the interval regression with sample selection specification, i.e. an additional 1,000 kWh of annual heat consumption increases the investment sum by approximately 1.3 percent. This translate into a fairly small numerical partial effect with less than DKK 50 additional investment per 1,000 kWh of annual heat consumption for the majority of the observations. This supports the notion that energy efficiency concerns are rarely the main driver of energy renovations (Wilson et al., 2015). As for the audit decision, however, we see that oil heating as primary heating source does show a significant association with the investment sum (significant at the 10 percent level). The average partial effects are quite high, with a 288 percent increase in the investment sum predicted by the interval regression model, and a 276 percent increase predicted by the interval regression with sample selection. The correspoding numerical partial effects predict additional investments between DKK 1,000 and DKK 5,000 for the subsample interval regression, and between DKK 200 and DKK 2,500 for the interval regression with sample selection.

Out of the variables indicating a change in the life situation, getting married in the policy period and relocation in the first half of the policy period are significant in both models. Getting married is strongly associated with the investment sum, with the investment sum predicted to increase by 688 percent in the subsample interval regression, and by 574 percent in the interval regression with sample selection model. Numerically, partial effects for the majority of the observations fall between DKK 2,500 and DKK 20,000 in the subsample interval regression, and between DKK 1,000 and DKK 5,000 in the interval regression with sample selection. Relocation in the first half of the policy period also has a very strong association with the investment decision, though somewhat lower than getting married. The average partial effects indicate an increase of 256 percent for the subsample interval regression, and a 299 percent increase for the interval regression with sample selection. The majority of the partial effects in the interval regression are between DKK 1,000 and DKK 5,000, and between DKK 200 and DKK 2,500 for the interval regression with sample selection estimation.

Across the two decision – having a free energy audit and investing in energy saving measures – we find that households that moved in the first half of the policy period were more likely to have an audit and invested more in energy saving measures. This may indicate this target group as 'low hanging fruit', though this evaluation cannot tell us if the audits have generated any additional investments that might have also happened in absence of policy.

4.6 Conclusion

In this article we have investigated the determinants of households joining a free energy audit programme, and investing in energy saving measures in their home. We investigated the decision to have an energy audit by applying a weighted probit regression framework with a comprehensive set of socio-economic variables and housing characteristics. In two alternative specifications we included dummy variables capturing the initial household life situation, and changes in the life situation of households during the policy period. Findings indicate that while the initial life situation does not play a role for the decision to have an audit, changes in two life situations – starting retirement and relocation in the first half of the policy period – have a positive influence on the propensity to have an energy audit.

In the estimation of the final investment in energy saving measures we applied two regression frameworks: a subsample interval regression and an interval regression accounting for sample selection. The application of the interval regression model accounting for sample selection has proven difficult, as the conceptualized exclusion restriction did not allow for empirical identification of the second stage. This issue may be due to the limited variance in the self-selected sample that had an energy audit. This made it necessary to include two additional (theoretically unfounded) exclusion restrictions. When applying the additional exclusion restrictions, the estimation does deliver results that are qualitatively similar to the subsample interval regression.

In the estimation of the final investment decision we find that getting married and relocation in the first half of the policy period have a strong association with the amount that is invested into energy saving measures. Together with the association found for starting retirement and relocation with the decision to have an audit, this provides support for the notion of a Fresh Start Effect (Peetz and Wilson, 2012), in which private households are more likely to seek out new information or make changes in their life as they find themselves in a new life situation. This insight is potentially useful for future policy design, e.g. by marketing energy audits specifically to households that just moved into a municipality as a window of opportunity to reach households that are susceptible to new information. It could also be interesting to tailor and market energy efficiency programmes to home owners starting retirement, as there is some evidence that elderly households are harder to reach than the general population (Berry and Brown, 1988).

An additional, and somewhat surprising finding indicates that annual heat consumption (both total and per square meter) is only weakly associated with both audit and investment decision. Given that the main motivation for the policy maker is to encourage energy savings, energy consumption appears to play a remarkably small role for home owners. This may be in line with evidence from the literature that households are either inattentive to the energy properties of their home (Palmer and Walls, 2015), or that energy efficiency renovations are rarely motivated by energy concerns alone (Wilson et al., 2015).

Acknowledgements

We are highly indebted to Peter Rathje and ProjectZero for sharing their data and helpful discussions that led to the writing of this paper. Furthermore, we are thankful to Ulrik Nielsen from the Danish Ministry of Taxation for providing the annual heat consumption data. Finally, we thank participants of numerous seminars at the Department of Management Engineering at the Technical University of Denmark, participants of the Danish Environmental Economics Conference 2014 and 2015, and the SAVE-E workshop 2017. This work was supported by the Danish Innovation Fund as part of the SAVE-E project.

Bibliography

Abrahamse, W., Steg, L., Vlek, C., Rothengatter, T., 2007. The effect of tailored information, goal setting, and tailored feedback on household energy use, energy-related behaviors, and behavioral antecedents. Journal of Environmental Psychology 27 (265–276).

Allcott, H., Greenstone, M., 2012. Is there an energy efficiency gap? Journal of Economic Perspectives 26 (1), 3–28.

Archer, D., Costanzo, M., Iritani, B., Pettigrew, T., Walker, I., L.T. White, L. . T., 1984. Energy Efficiency: Perspectives on Individual Behavior. ACEEE Press, Ch. Energy conservation and public policy: The mediation of individual behavior, pp. 69–92.

Berry, L. G., Brown, M. A., 1988. Participation of the elderly in residential conservation programmes. Energy Policy 16 (2), 152 – 163.

Bhattacharya, J., Goldman, D., McCaffrey, D., 2006. Estimating probit models with self-selected treatments. Statistics in Medicine 25, 389 – 413.

Burbidge, J. B., Magee, L., Robb, A. L., 1988. Alternative transformations to handle extreme values of the dependent variable. Journal of the American Statistical Association 83 (401), 123–127.

Dai, H., Milkman, K. L., Riis, J., 2014. The fresh start effect: Temporal landmarks motivate aspirational behavior. Management Science 60 (10), 2563–2582.

Davis, L., Metcalf, G., 2014. Does better information lead to better choices? evidence from energy-efficiency labels. Tech. rep., EI Working Paper.

Fowlie, M., Greenstone, M., Wolfram, C., 2015a. Are the non-monetary costs of energy efficiency investments large? understanding low take-up of a free energy efficiency program. American Economic Review: Papers & Proceedings 105(5), 201–204.

Fowlie, M., Greenstone, M., Wolfram, C., 2015b. Do energy efficiency investments deliver? evidence from the weatherization assistance program. Working Paper 21331, NBER Working Paper Series.

Freedman, D. A., Sekhon, J. S., 2010. Endogeneity in probit response models. Political Analysis 18 (2), 138 – 150.

Frondel, M., Vance, C., 2013. Heterogeneity in the effect of home energy audits: Theory and evidence. Environmental and Resource Economics 55 (3), 407–418.

Fuller, M., Kunkel, C., Zimring, M., Hoffman, I., Soroye, K., Goldman, C., 2010. Driving demand for home energy improvements. Tech. Rep. 3960, Lawrence Berkeley National Labatory.

Granade, H., Creyts, J., Derkach, A., Farese, P., Nyquist, S., 2009. Unlocking energy efficiency in the us economy. Tech. rep., McKinsey /& Company.

Greene, W. H., Hensher, D. A., 2009. Modelling ordered choices.

Harsany, J. C., 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. Journal of Political Economy 63 (4), 309–321.

Hartman, R. S., 1988. Self-selection bias in the evolution of voluntary energy conservation programs. The Review of Economics and Statistics 70 (3), 448–458.

Heckman, J. J., 1979. Sample selection bias as a specification error. Econometrica: Journal of the econometric society 47 (1), 153–161.

Henningsen, G., Petersen, S., 2017. Can household characteristics consistently explain the heterogeneity in households' energy effciency investments? a meta-analysis. Working Paper.

Hirshleifer, J., Riley, J. G., 1979. The analytics of uncertainty and information: An expository survey. Journal of Economic Literature 17 (4), 1375–1421.

Hirst, E., Berry, L., Soderstrom, J., 1981. Review of utility home energy audit programs. Energy 6 (7), 621 – 630.

Hirst, E., Goeltz, R., 1985. Estimating energy savings due to conservation programmes: The bpa residential weatherization pilot programme. Energy Economics 7 (1), 20–28.

Hirst, E., Grady, S., 1983. Evaluatiion of a winsconsin utility home energy audit program. Journal of Environmental Systems 12 (4), 303–320.

Hoffman, S., High-Pippert, A., 2010. From private lives to collective action: recruitment and participation incentives for a community energy program. Energy Policy 38 (12), 7567–7574.

Hoicka, C. E., Parker, P., Andrey, J., 2014. Residential energy efficiency retrofits: How program design affects participation and outcomes. Energy Policy 65, 594–607.

Honaker, J., King, G., Blackwell, M., 2011. Amelia ii: A program for missing data. Journal of Statistical Software 45 (7), 1–47.

Jimenez, E., Kugler, B., 1987. The earnings impact of training duration in a developing country: an ordered probit selection model of colombia's servicio nacional de aprendizaje (sena). Journal of Human Resources 22 (2), 228–247.

King, G., Honaker, J., Joseph, A., Scheve, K., 2001. Analyzing incomplete political science data: An alternative algorithm for multiple imputation. American Polical Science Review 95 (1).

LaRiviere, J., Holladay, S., Novgorodsky, D., Price, M., 2014. Prices vs. nudges: A large field experiment on energy efficiency fixed cost investments.

Lumley, T., 2017. survey: Analysis of Complex Survey Samples. Version 3.32-1.

Lumley, T., Scott, A., 2014. Tests for regression models fitted to survey data. Australian and New Zealand Journal of Statistics 56 (1), 1 – 14.

Lutzenhiser, L., 1993. Social and behavioral aspects of energy use. Annual Review of Energy and the Environment 18, 247–289.

Miller, R., Ford, J., 1985. Shared savings in the residenlial market: A public-private partnership for energy conservation. Tech. rep., MD: Urban Consortium for Technology Initiatives, Energy Task Force.

Osterhus, T. L., 1997. Pro-social consumer influence strategies: When and how do they work? Journal of Marketing 61 (4), 16–29.

Palmer, K., Walls, M., 2015. Limited attention and the residential energy efficiency gap. The American Economic Review 105 (5), 192–195.

Palmer, K., Walls, M., O'Keeffe, L., 2015. Putting information into action: What explains follow-up on home energy audits? Tech. Rep. DP 15-34, RFF.

Peetz, J., Wilson, A. E., 2012. The post-birthday world: Consequences of temporal landmarks for temporal self-appraisal and motivation. Journal of Personality and Social Psychology 104 (2), 249–266.

Pettifor, H., Wilson, C., Chryssochoidis, G., 2015. The appeal of the green deal: Empirical evidence for the influence of energy efficiency policy on renovating homeowners. Energy Policy 79., 161–176.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL http://www.R-project.org/

Savage, L., 1954. The Foundations of Statistics. Wiley.

Sigelman, L., Zeng, L., 1999. Analyzing censored and sample-selected data with tobit and heckit models. Political Analysis 8 (2), 167 – 182.

Smith, J., Sweetman, A., 2016. Viewpoint: Estimating the causal effects of policies and programs. Canadian Journal of Economics 49 (3), 871 – 905.

Solon, G., Haider, S. J., Wooldridge, J. M., 2015. What are we weighting for? Journal of Human resources 50 (2), 301 – 316.

Stern, P., Aronson, E., Darley, J., Hill, D., Hirst, E., Kempton, W., Wilbanks, T., 1985. The effectiveness of incentives for residential energy conservation. Evaluation Review 10 (2), 147–176.

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

Stern, P. C., 1999. Information, incentives, and proenvironmental consumer behavior. Journal of Consumer Policy 22, 461–478.

White, I. R., Royston, P., Wood, A. M., 2011. Multiple imputation using chained equations: Issues and guidance for practice. Statistics in Medicine 30 (4), 377 – 399.

Wilson, C., Crane, L., Chryssochoidis, G., 2015. Why do homeowners renovate energy efficiently? contrasting perspectives and implications for policy. Energy Research & Social Science 7, 12–22.

Wilson, C., Dowlatabadi, H., 2007. Models of decision making and residential energy use. Annu. Rev. Environ. Resour. 32, 169–203.

Wirtshafter, R. M., 1985. Non-participants in utility energy conservation programmes. Energy Policy 13 (2), 143 – 155.

Wooldridge, J. M., 2010. Introductory Econometrics: A Modern Approach, 4th Edition. Nelson Education.

Zivin, J. G., Novan, K., 2015. Upgrading efficiency and behavior: Electricity savings from residential weatherization programs. Tech. rep., UC Davis.

Chapter 5

Detecting Behavioral Spillovers in a Real Effort Public Good Experiment

Sebastian Petersen and Helene Willadsen

Abstract

Social Norm interventions are increasingly used as a policy tool to influence people's everyday consumption behaviour. They promise a significant demand response at low financial costs. Often overlooked in the application of such behavioural interventions, however, is that they can have unintended spillover effects. This paper proposes a novel experimental framework to investigate the conditions determining occurrence and direction of spillovers. Subjects complete real effort tasks in a public good environment, in what we call a Real Effort Public Good game. Earnings are generated through a private task which benefits the individual only, and two public tasks, which benefit a group. A social reference intervention is applied in the treatment group in which participants receive information about session-level contributions to only one public good. We find that contribution patterns in the control group resemble regular public good games in the literature, which shows that our underlying public good framework delivers comparable outcomes. We also find that our social reference intervention does have a positive effect on contributions in the target public good. Finally, we find that the social reference treatment leads to a diminished increase in contributions to the non-targeted public good, indicating a positive spillover.

5.1 Introduction

Making society more sustainable requires individuals to act pro-environmental in their daily life. While this general notion is easy to agree with, when it comes to actual change resistance is often encountered. Due to the nature of democratic and liberal societies this resistance cannot be overcome by mandates that dictate behaviour, which is one of the rationales for libertarian paternalism (Sunstein and Thaler, 2003) Within this context behavioural science based interventions have become increasingly popular as a soft policy alternative. This trend has even sparked the creation of government entities specialized in this type of policies, such as the Behavioural Insights Team in the UK and the Social and Behavioral Sciences Team in the US.

Social referencing is a policy tool proven to be effective at influencing people's behaviour by communicating peer group behaviour in a salient way. Thus, it has the potential to be used by policy makers to encourage pro-social and pro-environmental behaviour. Providing a social reference works in a variety of settings. It has been used to get hotel customers to reuse towels more often (Goldstein et al., 2008), get people to donate more to charity (Frey and Meier, 2004), or help households reduce consumption of energy or water (Schultz et al. 2007; Allcott 2011; Tiefenbeck et al. 2013).

Residential energy consumption is a frequent focus for environmental policy, as it accounts for more than 20 percent of CO_2 emissions in most developed countries. (IEA, 2008) Based on insights from behavioural science, the company Opower implemented large-scale interventions in cooperation with utilities to send consumers home energy reports providing a social reference about other households' energy consumption. Short run evaluations of the Opower intervention show that the average household reduces electricity consumption by 2 percent (Allcott, 2011). In the long run households reduce consumption immediately upon receiving their energy report, but steadily adjust consumption toward previous levels, until receiving a new report.(Allcott and Rogers, 2014).

The policy evaluation of the Opower intervention only measures the household reaction in isolation of other behavioural domains, but Allcott (2011) notices that the treatment could have unanticipated consequences because: '*Consumers could, for example, also become motivated to drive their cars less, or could perhaps even drive more* [...]' (Allcott (2011), p.1089). Since the ultimate goal of encouraging energy savings is to reduce CO_2 emissions, such behavioural spillovers could undermine the efficiency of the policy. (Dolan and Galizzi, 2015) Mixed evidence is provided in the existing literature on spillovers, with studies finding negative spillovers (Mazar and Zhong, 2010; Catlin and Wang, 2013; Tiefenbeck et al., 2013), positive spillovers (Lanzini and Thøgersen, 2014), or no spillovers (Poortinga et al., 2013; Littleford et al., 2014).

In this study we present an experimental framework to systematically explore spillovers in the lab. Inspired by the Opower intervention in the field we apply a social reference treatment in a Real Effort Public Good game. In the experiment, subjects choose to work on a private task that only benefits themselves, or one of two group tasks that generate less profit for themselves, but also provides profit to their group. The task we use is a variety of the slider task (Gill and Prowse, 2011), and the underlying incentive structure corresponds to a regular public good game (Andreoni, 1988; Croson, 1996). We find that our social reference intervention does produce higher contributions to the treated public task, and that this effect also increases contributions in the untreated public task, implying a positive spillover. We hypothesize that this result may be related to the similarity of the two tasks and encourage future research to explore how task similarity affects the occurrence and direction of spillovers.

The following section describes our experimental design and the implementation; Section 5.3 formulates hypotheses based on previous results from the literature; Section 5.4 presents the results; and Section 5.5 concludes.

5.2 Experimental Design and Implementation

Our experimental design is based on a public good incentive structure to capture the dilemma at the heart of pro-environmental behaviour that is beneficial for society, but costly to the individual. The aim of the experiment is to create a simplified version of the real world (Gneezy and Rey-Biel, 2015) that captures the decision environment in which frequent, everyday consumption decisions are made. In line with this idea subjects face a constant stream of decisions with limited time for deliberation in what we call a Real-Effort Public Good (REPG) game. In the game, contributions to a public good encompass opportunity costs and a reduced pay-off for the individual, with additional benefits for the group an individual is part of. In contrast to a regular public good game, subjects have to complete slider tasks to earn points. The slider task we use is based on Gill and Prowse (2012), of which we use the scrollbars version (see Figure 5.1). In the task, subjects have to use their computer mouse to move the slider to the middle of the bar, to a value of 50. The scrollbars are a bit easier than the regular sliders, reducing the potential for subject frustration. At the same time it retains the general advantages of the slider task, in that it is easy to communicate, contains no random elements, and requires very little marginal effort to complete.

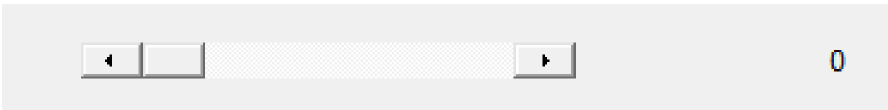


Figure 5.1: Scrollbars version of the slider tasks by Gill and Prowse (2012)

A second difference to a regular public good game is that subjects have the option to contribute to two public goods instead of one. To differentiate the two public tasks we have applied a framing to the experiment, with the two public tasks framed as energy and water task. Contributions to the public goods are made by completing the corresponding sliders, which means that subjects generate earnings and make contribution decisions simultaneously by deciding which sliders they complete.

Figure 5.2 shows the decision environment. The top window is filled with 42 sliders that are always available to subjects (they are replaced by another screen of 42 sliders, if the whole screen is completed), these are framed as private tasks. The bottom two windows are framed as water and energy task (together referred to as public tasks), represented by a water faucet and a light bulb. The public tasks are not always available, but instead appear in a randomized order throughout each period. Each task appears five times per period, and tasks do not overlap. This means that subjects never face the direct choice between the two public tasks, but only through redirecting effort from the private task. The public tasks are available for a limited time of ten seconds, after which they disappear. Availability of the public tasks is communicated by either the light bulb changing colour from transparent to yellow, or a water droplet appearing below the water faucet. Additionally, a slider appears below one of the two pictures. Visual feedback is given to subjects whenever a task is completed, with a coin appearing next to a completed private task, the light bulb going transparent for the treated task, and the water faucet droplet disappearing for the untreated task.

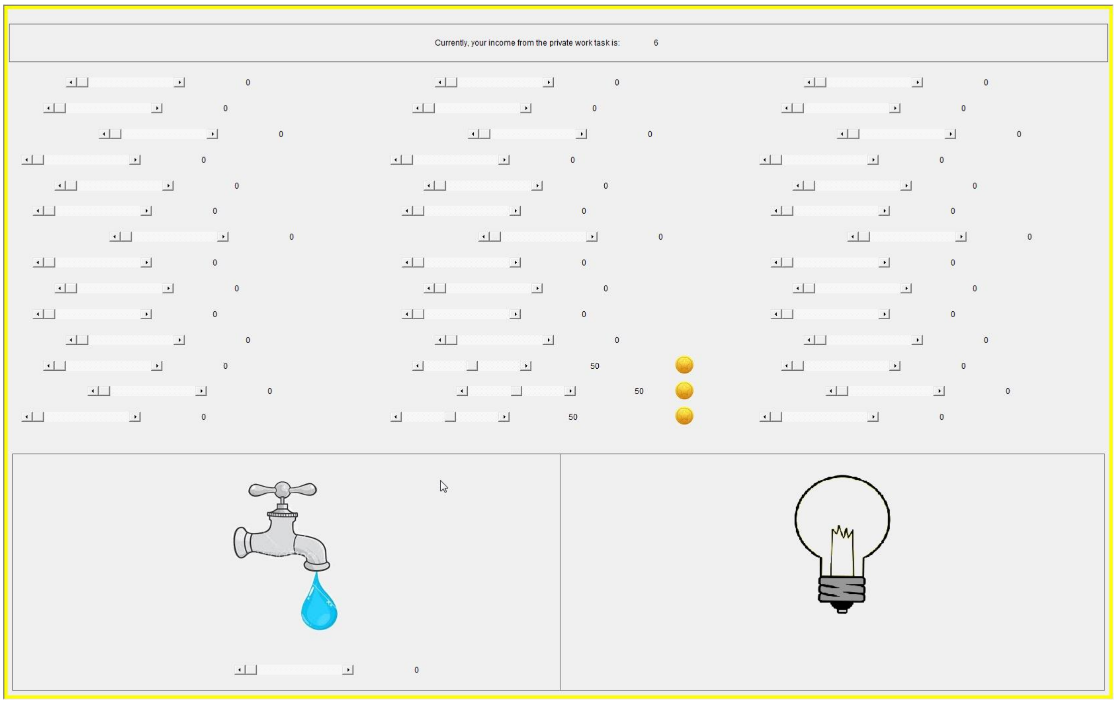


Figure 5.2: Screenshot of the decision environment of the Real-Effort Public Good Game (with three completed private tasks and the untreated/water task available)

Subjects were assigned to groups of four, and the incentive structure of the experiment corresponds to a public good game with a marginal per capita return of 0.5, i.e. for every public task completed each subject in a group earns half the amount they receive for a private task. Hence the game is in line with the general public good framework in which the marginal earnings for each subject from the private task is higher than for the public tasks. At the same time the marginal earnings for the group are always higher for the public tasks, than from the private task. The total earnings of each subject i is based on private and public tasks they themselves and their group members have completed:

$$\pi_i = 2 \cdot x_i^P + 1 \cdot (x_i^T + \Sigma_{j=1}^{3} x_j^T) + 1 \cdot (x_i^U + \Sigma_{j=1}^{3} x_j^U), \tag{5.1}$$

with x_i^P the number of private tasks completed; x_i^T the number of treated tasks completed; x_i^U the number of untreated tasks completed; x_j^T the number of treated tasks completed by another member of the group; and x_j^U the number of untreated tasks completed by another member of the group.

5.2.1 Treatments

Subjects are assigned to a control or social reference treatment, with treatments assigned between subjects and at session level. The evaluation screen shown to subjects after each period differs between the two treatments. Both control and treatment group receive feedback on their pay-off in the previous period, the total number of private and public tasks they completed, and how many public tasks their other group members completed. (see Figure 5.3, left side) In addition to this information, subjects assigned to the social reference treatment received information on treated tasks completed by others in their session through the communication of the session average and the average in the most active group. (see Figure 5.3, whole screen)

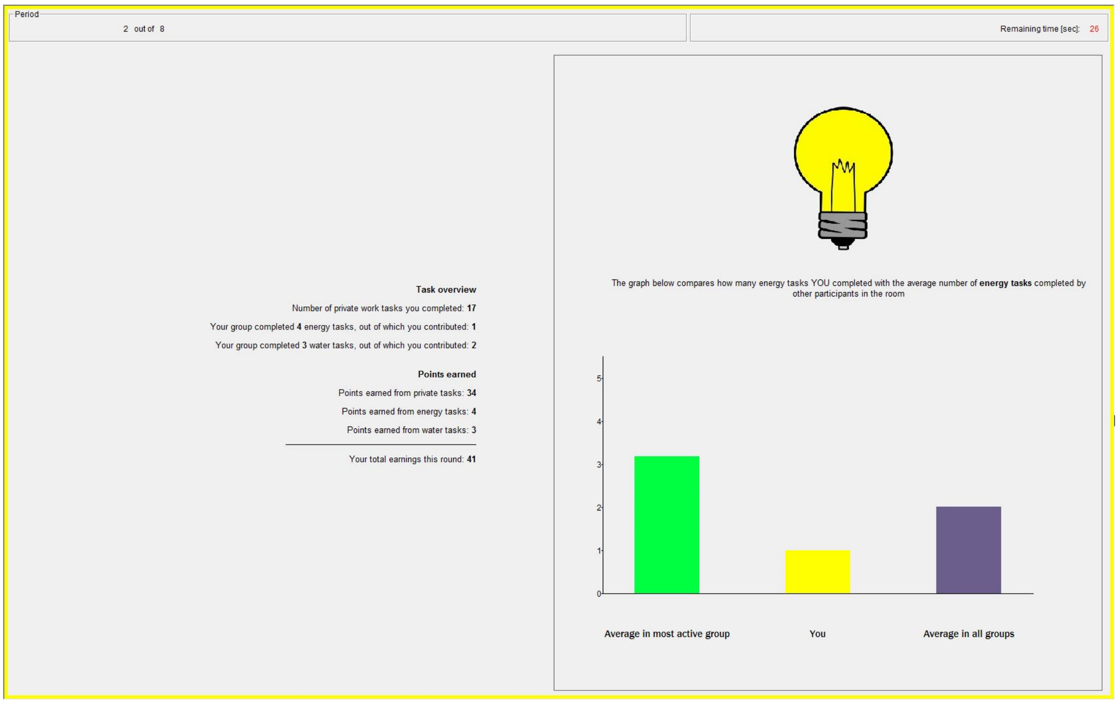


Figure 5.3: Screenshot of the evaluation screen for the treatment group

5.2.2 Implementation

The experiment was conducted in 4 experimental sessions of approximately 75 minutes, with a total of 92 subjects. The experiment took place at the Laboratory for Experimental Economics (LEE) at the faculty for Social Sciences at the University of Copenhagen in February 2017. The experiment was programmed in z-Tree version 3.4.7 (Fischbacher, 2007). Lab assistants followed a script to guide participants through the experiment, with further instructions provided to the subjects in written form. The lab script describes how to run the experiment in z-Tree, provides announcements to read out to participants at the different stages of the experiment to ensure replicability of the experiment. The lab script and participant instructions are provided in Appendices B and C.

Subjects were paid based on their decisions in each period of the experiment. During the experiment participants earned points, which were converted into Danish Kroner (DKK) at a rate of 6:1. Average earnings across control and treatment group was DKK 135.96 (around €18.25). Participants were mostly students at Copenhagen University, and were recruited through the 'Online Recruitment System for Economic Experiments' (ORSEE, Greiner 2015). Subjects played in groups of four, which made it necessary to turn away subjects if the number of participants showing up were not divisible by four. These subjects were selected on a volunteer basis or by random selection, and compensated with a DKK 50 show-up fee.

Each experimental session consisted of four parts: (1) instructions to the experiment including control questions; (2) 8 periods of the Real-Effort Public Good game; (3) a questionnaire consisting of a risk attitude question, a set of ten questions on environmental behaviour, and some background information; and (4) payment. Differences between control and treatment groups were limited to the evaluation screen in part 2.

5.3 Hypotheses

Our experiment builds a novel experimental framework based on previous research for investigating spillovers in a laboratory setting where all aspects of the decision environment can be controlled. We use the incentive structure of a standard public good game, as it captures the dilemma at the heart of collective action problems, where the private and social optimum diverge. For example, reducing consumption of energy or water through efficiency gains or curtailment is often considered a public good (Hasson et al., 2010; Allcott, 2011; Heitzig et al., 2011), as it can reduce resource scarcity or pollution affecting everyone in a neighbourhood, a region, or even globally.

The Nash equilibrium prediction for a public good game is for every player to contribute nothing to the public good, i.e. free-riding. Despite this prediction it is commonly found that subjects contribute a sizeable part of their endowment to the public good, with initial contributions often found in the range of 40-60 percent of the total endowment (Ledyard, 1995; Cox and Sadiraj, 2007)). In repeated games with multiple periods contributions usually deteriorate, especially in the final round of the game.

The basic setup of our experiment differs from the usual public good game in that subject do not receive an endowment for free, but have to work for it. Studies that investigate the effect of working for an endowment do not find any differences from contribution patterns in other public good games (Cherry et al., 2005; Muehlbacher and Kirchler, 2009). However, the majority of studies that investigate the effect of effort to earn the endowment on contributions in a public good game conduct these two activities in subsequent stages of the experiment, i.e. subjects first have to conduct some type of real-effort task earning them the endowment, and in a second step decide how to use this endowment. In our experiment, investing effort and contributing occurs simultaneously and continuously, by deciding which slider tasks to work on. The closest experiment we found in the literature was conducted by Dutcher et al. (2015). In this experiment subjects transcribed database entries, and while they were doing this task they could flip a switch to decide if their current earnings are going towards their private or a public account. Contribution patterns were again similar to results from other public good experiments.

Given the public good incentive structure and results from related literature we formulate our first hypothesis:

Hypothesis 1 *The Real Effort Public Good game produces contribution patterns in line with regular public good games.*

Social norms can guide individual behaviour, because individuals often rely on rules of thumb (also termed decision heuristics) in order to decide on an appropriate action or behaviour (Kallgren et al., 2000). Behavioural insights demonstrate that individuals evaluate their behaviour in a social context rather than in isolation, and people tend to do what is socially approved (Cialdini, 2003). This insight can be used to induce individuals to re-evaluate the costs and benefits of their decisions by directing attention to the behaviours of others. However, social norms are not useful to direct behaviours in all situations. Kallgren et al. (2000) argue that a social norm affects a subject's behaviour if the social norm is focal for the behaviour of the subject and from a group that the subject identifies with.

In the field, social references have been found to be effective at reducing consumption in domains such as water or energy use (Tiefenbeck et al., 2013; Ferraro and Price, 2013; Allcott and Rogers, 2014; Costa and Kahn, 2013). In these domains behavioural tools are used in addition to pricing mechanisms such as carbon taxes or energy efficiency subsidies. Instead of changing the actual cost and benefits through market-based mechanisms, the aim is to change the perceived costs and benefits (Madrian, 2014). By communicating a social norm in a salient way interventions exploit people's decision heuristic to look at peer behaviour to achieve pro-environmental behaviour.

In our experiment we provide a social reference by informing subjects about the contribution behaviour of others in the same experimental session. Since the majority of the participants are students, this implies a strong identification as peer group. Furthermore, the behaviour is elicited with low temporal lag, i.e. directly following a period. The social reference is provided in a way that compares a subject's contribution to the treated public task to the average number of treated tasks completed in the same session/room, and to the average number of treated tasks completed in the 'most active' group. Thus, we provide a descriptive social reference (Cialdini et al., 1991). While the injunctive social reference does give a higher benchmark, we avoid language giving a positive framing (e.g. 'best', 'highest') of the benchmark, to avoid experimenter demand effects (Zizzo, 2008).

Based on this evidence we formulate our second hypothesis:

Hypothesis 2 *The social reference intervention increases contributions to the treated task in the social reference treatment compared to the control treatment.*

Finally, we turn to how contributions in the second, untreated public good will be influenced by the social reference treatment applied to the treated public good. Dolan and Galizzi (2015) describe the general idea of a spillover to be like dropping a pebble into a pond (intervention), causing a splash (intended effect) and ripples on the surface (spillover). A spillover can be in line with the initial, intended effect, but could also contradict the purpose of the policy. Based on Thøgersen (1999) we generally define a behavioural spillover as a change in a non-targeted behaviour caused by an intervention aimed at changing another behaviour. In line with Truelove et al. (2014) we make a simple distinction between positive and negative spillovers, where the former indicates a spillover in the same direction as the primary effect on the treated behaviour, and a negative spillover indicating an effect in the opposite direction. Following Dolan and Galizzi (2015) we exclude spillovers that are automatic responses, as for example income effects (see also Tiefenbeck et al., 2013). In the experiment automatic responses are prevented by not having overlapping public tasks, i.e. there is no direct competition between the public tasks.

Psychological mechanisms behind negative spillovers are theorised to be forms of moral licensing and single action bias. (Truelove et al., 2014) Moral licensing gives an individual a good feeling about themselves when engaging in a behaviour perceived as morally good and this creates the feeling that this good deed justifies transgressions in another behaviour domains. The single action bias is based in worry about the circumstances of a decision (e.g. desire to do something about climate change), and this worry is alleviated by acting in a more social or environmental way. However, acting in the 'right' way makes the individual worry less, which reduces the chance of acting pro-social or pro-environmental in a subsequent decision. Truelove et al. (2014) specifically raise the point that social norms may induce negative spillovers, as they might invoke negative feelings about pro-environmental behaviour in individuals that do not have a pro-environmental mindset.

In Dolan and Galizzi (2015) the process of a positive spillover is explained by individuals having a preference for consistency, meaning that people have a preference for behaving in a manner consistent with prior actions and beliefs. Therefore, positive spillovers are more likely to occur for behaviours that share a common motive relevant for a person's view of themselves, e.g. their perception as pro-environmental individual. In our experiment the two public good tasks are closely linked by a framing that present the tasks as reducing water consumption and saving energy. At the same time the tasks are identical in terms of cost and benefit.

Based on the literature we do not see a clear indication on whether to expect positive or negative spillovers, which leads us to state our third hypothesis as:

Hypothesis 3 *The social reference intervention targeting the treated task generates a spillover, increasing or decreasing contributions in the untreated task in the treatment* *group compared to the control group.*

5.4 Results

We first look at general results of the experiment and will look at treatment effects in the following subsections. Figure 5.4 shows a boxplot of the total number of slider tasks (i.e. both private and public tasks) completed by the test subjects in the control and treatment groups across all 8 periods. Over the first five periods the average number of tasks in both groups increase by about 10 sliders on average, indicating a learning effect. In the last three periods the average number of tasks does not increase any more. The mean number of sliders completed in both groups in the first period is 42.3, while in the last period it reaches 51.9. This corresponds to an 18.5 percent increase. Gill and Prowse (2011) and Benndorf et al. (2014) find a similar degree of learning for the slider task, with 16.6 and 16.7 percent over 9 periods.

The boxplots in Figure 5.4 also reveal that subjects differ a lot in terms of their performance. In later periods top performers were able to complete more than 70 sliders, while a few low performers completed less than 20 sliders even in the last periods. In terms of the incentives to complete public or private tasks this should not make a difference, as these remain in the same proportion. However, low performers might see completing a task as more costly, and thus have to sacrifice more effort in relative terms to contribute to the public good. Even if this is the case, we would assume that this effect is prevalent for both control and treatment group. A more critical issue with low performers is that the public tasks are only shown for a limited time window of 10 seconds. This means that subjects that complete less than 15 tasks over the 150 seconds of the experiment will have a hard time completing these, and might get frustrated when trying them. We address this in a sensitivity analysis (see Tables 5.5 and 5.6 in the Appendix), but the qualitative results remain unchanged.

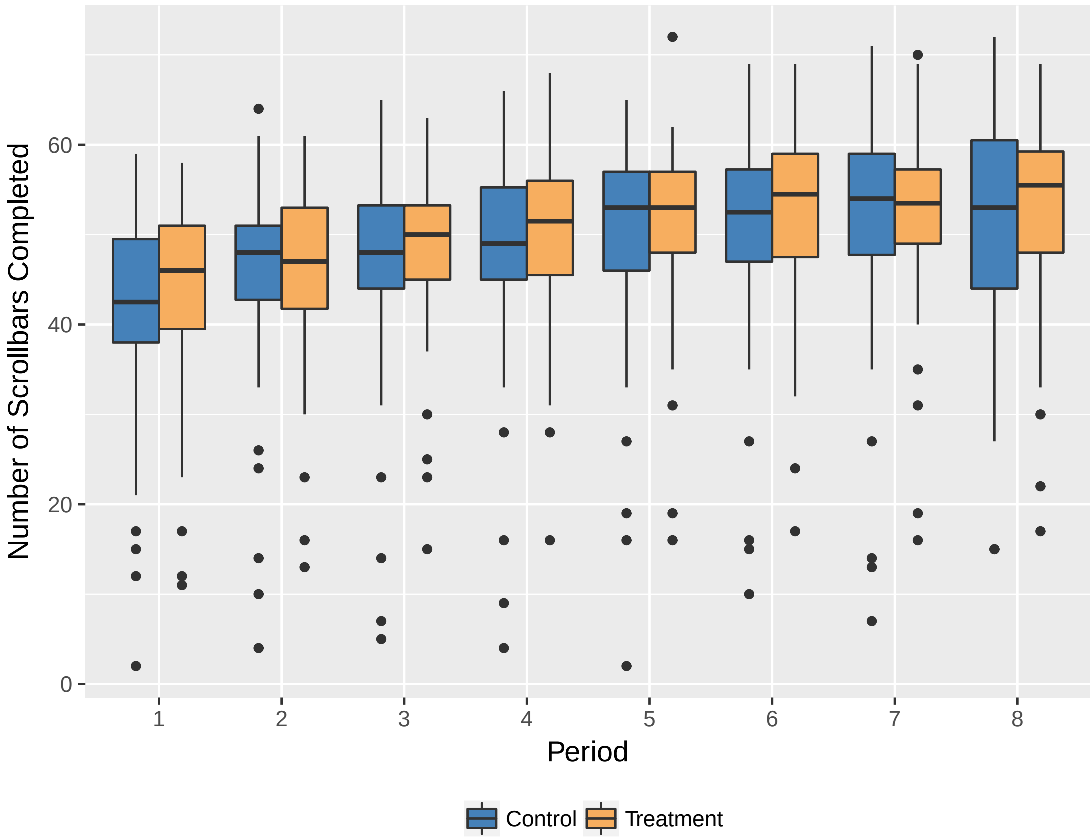
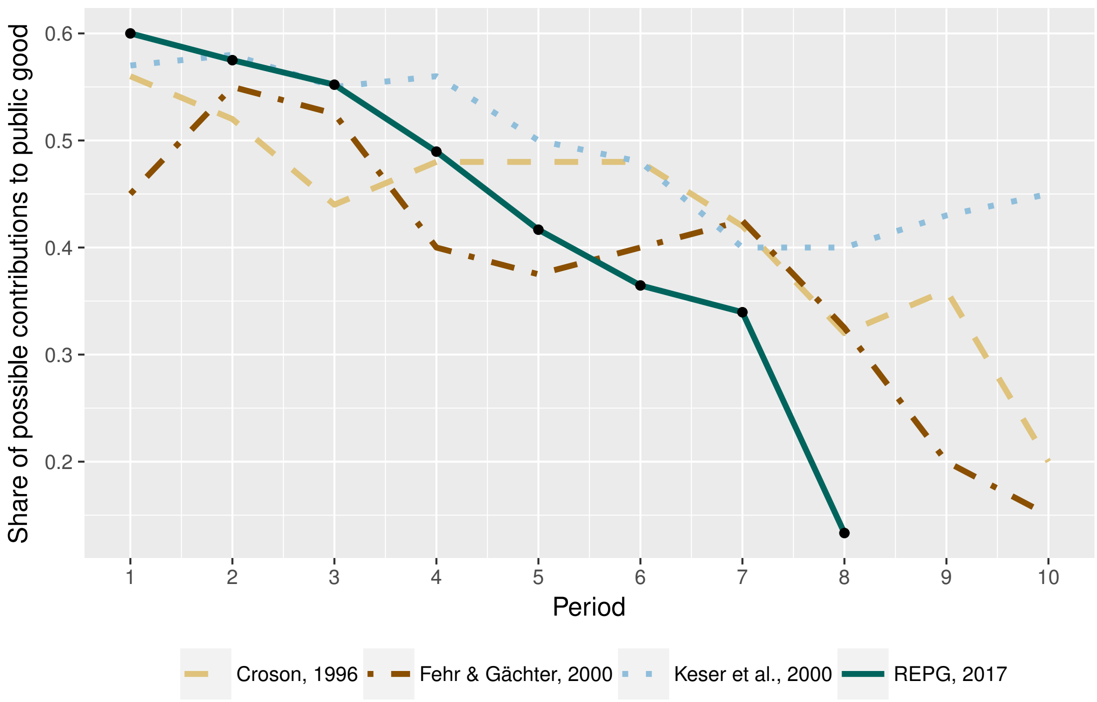


Figure 5.4: Boxplots for Total Number of Slider Tasks Completed Per Subject

Figure 5.5 plots the contributions to both public goods in the control group of our Real Effort Public Good (REPG) experiment together with contribution values obtained from public good experiments in the literature that also apply a partner matching and similar marginal per capita return. (Croson 1996, Fehr and Gächter 2000, Keser and van Winden 2000) Contributions start out at a similar level, with about 60 percent of the public tasks completed by subjects in the control group. This is slightly higher than what we see in the literature, but is generally in the same range. Across periods we see declining contribution patterns in both our experiment and in the experimental values from the literature. There is a similar overall trend in our REPG experiment, but contributions in our experiment decrease monotonously, while in the other experiments



Notes: Values from the literature have been transcribed from graphs; Fehr and Gächter (2000) uses a slightly lower MPCR of 0.4; Keser and van Winden (2000) did not finish after 10 periods.

Figure 5.5: Contributions to Public Good in Different Public Good Games

there are occasional up-ticks in contributions. We hypothesise that this might be related to the randomization in the order of public task appearance, making it harder to plan and follow strategic contribution patterns. Contributions in the final round reach their minimum, which is in line with the literature values. Only Keser and van Winden (2000) has significantly higher contributions in the last period shown in Figure 5.5, which is due to their experiment running for a total of 25 periods. We conclude that the contribution pattern in the REPG experiment are comparable to the ones found in regular public good games.

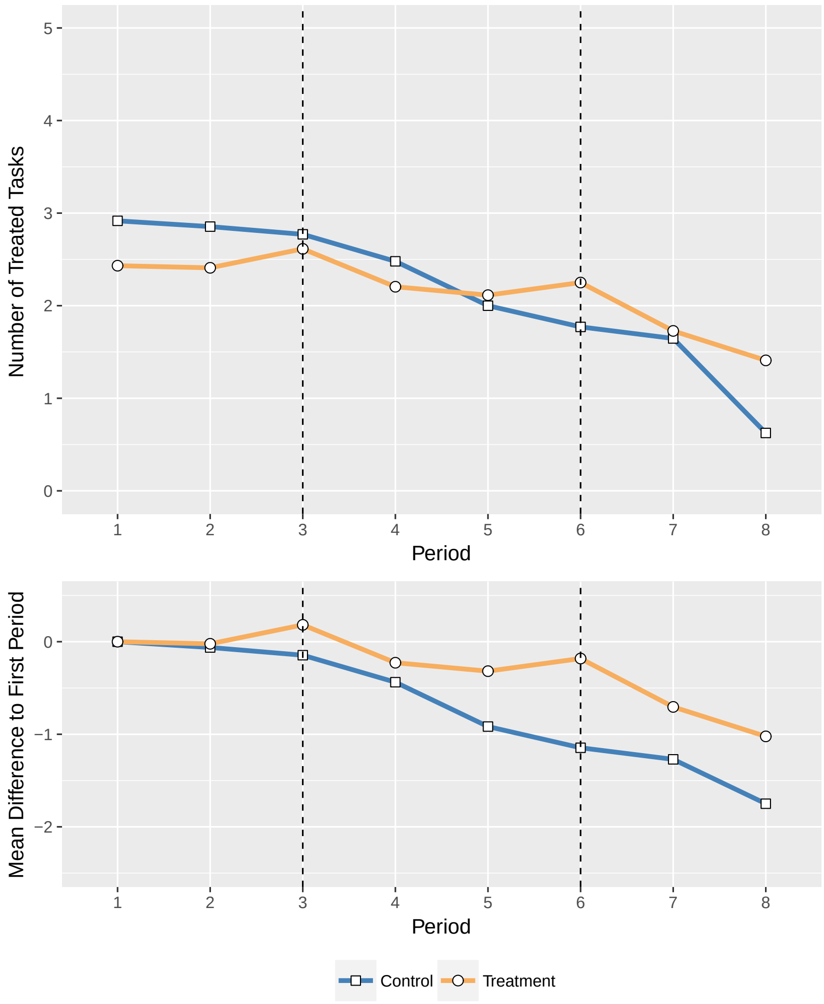
5.4.1 Social Reference Treatment Effect

To investigate the direct effect of the social reference treatment we compare contributions to the treated public good in the two treatments. The first plot in Figure 5.6 shows the mean number of treated tasks completed in the control and treatment groups.[1] The dotted lines delineate the periods which were directly affected by the intervention, i.e. the evaluation screen directly before these periods gave feedback on the number of treated tasks completed by others in the same session. Contributions in the control group appear to start at a slightly higher level, with on average half a task more completed. However, this difference is not significant at the 10%-level. Regarding the development of the contributions there is a clear difference between the treatments, with the control group decreasing monotonously, while the treatment group shows two small peaks in periods 3 and 6, and finally finishing higher than the control group.

The second plot in Figure 5.6 uses the first period task completion in the control and treatment group as index, to better illustrate the differences in the development of the contributions to the treated public good. Again we see the two peaks in periods 3 and 6, with the peak in period 3 raising contributions above the initial level. However, this does not change the general trend of declining contributions over time, as is commonly observed in public good experiments. When looking over the whole course of the experiment the control group decreased contributions to the treated task by around 1.75, while the treatment group only reduced contributions by about 1 task compared to the index in period 1.

Since the observations after the first period are not independent, the comparison of means only gives us an indication that the treatment group completed slightly more of the treated public tasks. To address this we run ordinary least squares (OLS) regressions in pooled and fixed effects specifications to get a more robust and precise estimate of

[1] The control group in period 8 is only based on one control session, because there was a glitch with the public tasks in the other session. This explains the inconsistency between the upper and lower graph in Figure 5.6 and 5.7 for period 8.



Note: For period 8 only one control session is reported, due to a glitch in the first session of the experiment, which explains the slight inconsistency between the two graphs in the last period.

Figure 5.6: Development of contributions to the treated public task across periods (the dotted lines delineate the periods directly affected by the social reference treatment)

the effect of the social reference treatment. As dependent variable we use the number of treated tasks completed by the subjects in a period.

Table 5.1: Pooled OLS and Fixed Effects regression results for the treated public task

	Dependent variable:		
	Number of treated tasks completed		
	(1)	(2)	(3)
SocialReference	0.127	0.146	0.434^{**}
	(0.292)	(0.313)	(0.202)
Period	-0.153^{***}	-0.157^{***}	-0.215^{***}
	(0.052)	(0.051)	(0.048)
ContributionByOthers	0.167^{***}	0.169^{***}	0.060^{*}
	(0.032)	(0.031)	(0.035)
MeanEnvironmentScore		-0.240	
		(0.219)	
Risk		0.057	
		(0.100)	
GenderMale		-0.033	
		(0.308)	
Constant	1.651^{***}	1.830^{**}	
	(0.411)	(0.878)	
Subject fixed effects?	No	No	Yes
Observations	620	613	620
R^2	0.170	0.180	0.134
F Statistic	41.941^{***} (df = 3; 616)	22.189^{***} (df = 6; 606)	27.032^{***} (df = 3; 525)

*p<0.1; **p<0.05; ***p<0.01

Note: Heteroscedasticity robust standard errors in parenthesis

Column (1) of Table 5.1 shows the results of a simple OLS regression with a dummy for the social reference intervention that is 1 for the treatment group from period 3 to 8, when they are or have been subject to the treatment evaluation screen; a period variable to capture the time trend; and the number of treated tasks completed by others in the group in the previous period. The latter captures the influence the different group members have on each others contributions, assuming that the main effect is based on contributions in the most recent period. The Period variable has a negative sign, as expected based on the diminishing public task contributions seen in Figure 5.6. Treated public tasks completed by others in the group has a positive sign, which may be a sign of some degree of conditional cooperation as motivation to complete public tasks. The social reference dummy is not significant in this specification, which is not too surprising, given that the number of treated tasks completed did look fairly close in Figure 5.6.

In column (2) we add variables based on the questionnaire participants answered after the experiment, and a gender dummy. The MeanEnvironmentScore variable is derived from ten responses about behaviours related to environmental themes (e.g. recycling, saving energy), where participants indicated how much they engage in the behaviour on a five point scale ranging from Never to Very Often. This qualitative scale was translated into points from 0-4, with the average over all answers used as the MeanEnvironmentScore variable. Subjects had the choice to not answer a question, which was then excluded from the average calculation. One subject refused to give any answers and had to be excluded in this specification. The Risk variable is based on a simple risk elicitation question, where participants are asked to indicate how risk taking they judge themselves to be on a scale of 1 to 10. None of the added variables are significant in this specification, but it is still a bit surprising to see a negative sign on the mean environmental score variable.

Column (3) presents the results from an OLS regression with subject level fixed effects, which accounts for the correlation across periods for each subject. As the estimator subtracts the mean differences from both dependent and independent variables, this will also address any issues with differences in the initial contribution levels in the control and treatment groups. In this specification all the included variables are significant, at least at the 10%-level. Based on the coefficient on the social reference dummy we see that the treatment increases the number of treated tasks completed by about 0.4. In line with the graphical analysis this effect does not reverse the time trend, which shows that contributions to the treated public good go down by an average of 0.2 each period. Due to the mean differencing we cannot include the time invariant questionnaire answers in this specification.[2]

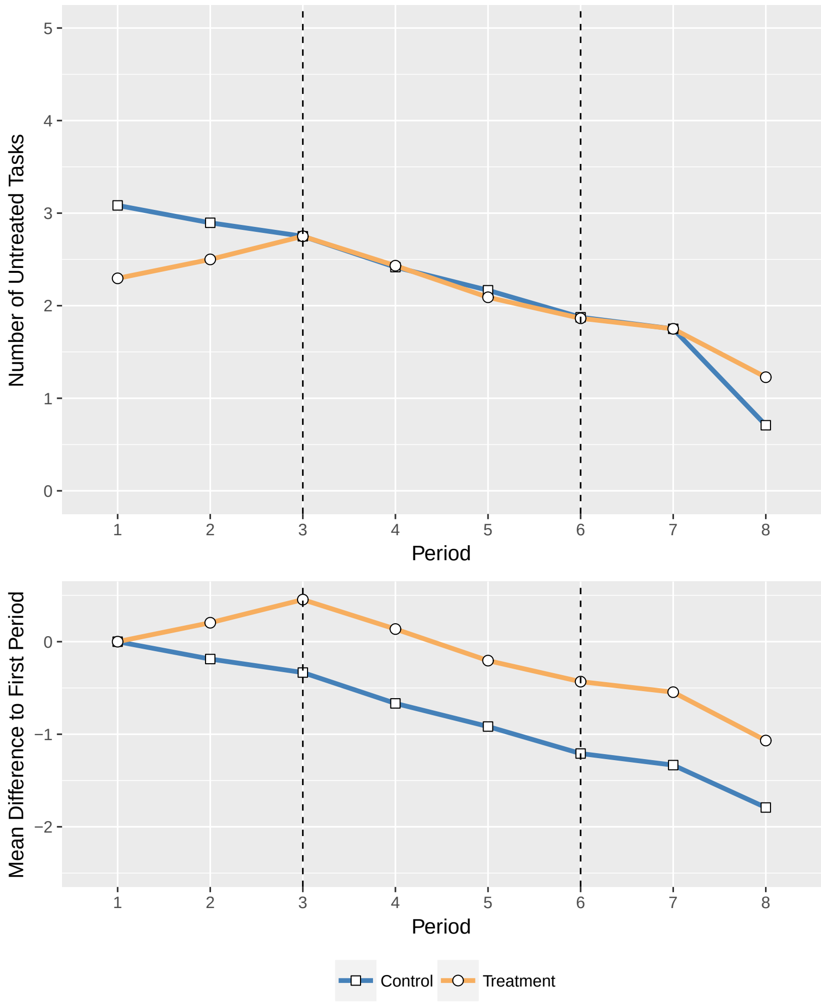
5.4.2 Evidence for a Spill-over Effect

Having established a direct effect of the social reference, we investigate a potential spillover on the untreated public tasks. Figure 5.7 shows the mean contributions to the untreated public task in the upper graph, and the mean contributions compared to the first period as index in the lower graph. Also for the untreated task the initial contribution level appears to be lower in the treatment group, which is significant at the 10%-level. Although, the social reference intervention did not take effect, there is a slight increase in contributions in the second period for the treatment group.

After the social reference intervention targeting the treated task is shown to the treatment group after period 2, this increasing trend continues to a peak in period 3. After this the development of the two groups is nearly identical. Hence, it appears like any effect of the social reference on the untreated Task had only a singular effect, without changing the rate of decay in contributions. From the graphical investigation it appears that mean contributions are more similar for the untreated task (in periods 3-8 that were actually affected by the treatment), which may suggest the absence or only weak spillover effects. Again, we run some OLS regressions to investigate the spillover effect further.

The specifications presented in Table 5.2 use the number of untreated tasks completed in a period by each subject, and the same independent variables as we reported in Table 5.1. The results reported in columns (1) and (2) are very similar to the results from the OLS regressions for the treated public task. The social reference treatment has small and non-significant coefficients, with only the time trend and contribution of other group members correlating with the number of completed tasks with expected signs. Also here

[2] We also ran a Random Effects estimation with the questionnaire variables included, but they remained insignificant and a Hausman test revealed that the model is inconsistent, hence we do not include it here.



Note: For period 8 only one control session is reported, due to a glitch in the first session of the experiment, which explains the slight inconsistency between the two graphs in the last period.

Figure 5.7: Development of contributions to the untreated public task across periods (the dotted lines delineate the periods directly affected by the social reference treatment)

Table 5.2: Pooled OLS and Fixed Effects regression results for the untreated public task

	Dependent variable:		
	Number of untreated tasks completed		
	(1)	(2)	(3)
SocialReference	0.059	0.066	0.369
	(0.283)	(0.308)	(0.240)
Period	-0.169^{***}	-0.173^{***}	-0.241^{***}
	(0.050)	(0.050)	(0.048)
ContributionByOthers	0.159^{***}	0.159^{***}	0.040
	(0.033)	(0.032)	(0.034)
MeanEnvironmentScore		-0.213	
		(0.217)	
Risk		0.056	
		(0.096)	
GenderMale		-0.212	
		(0.302)	
Constant	1.821^{***}	2.041^{**}	
	(0.409)	(0.904)	
Subject fixed effects?	No	No	Yes
Observations	620	613	620
R^2	0.164	0.173	0.145
F Statistic	40.304^{***} (df = 3; 616)	21.199^{***} (df = 6; 606)	29.666^{***} (df = 3; 525)

*p<0.1; **p<0.05; ***p<0.01

Note: Heteroscedasticity robust standard errors in parenthesis

we do not see a significant effect of the questionnaire variables in column (2).

In the results from the fixed effects regression in column (3) of Table 5.2 all variables have the same sign as in the fixed effects regression for the treated task in Table 5.1, but neither the treatment dummy nor the contribution of other group members are highly significant. However, the coefficients for both variables are just barely failing the 10%-level (p = 0.124 for the social reference dummy; p = 0.122 for the lagged contributions of other group members). Thus, we still consider this as evidence for the existence of a diminished positive spillover from the social reference intervention on the treated task into the domain of the untreated task. The spillover is positive, as it promotes behaviour in line with the intended treatment effect on the treated public good, but the effect is diminished both in magnitude and in terms of statistical significance.

5.4.3 Variance in Public Task Completion

We have shown that the social reference intervention increased contributions in both the treated and untreated task, without compensating the overall diminishing trend in contributions. In the following we investigate heterogeneity in the treatment effects, i.e. does the social reference intervention affect all subjects equally, or just some? To investigate this, we carry out F-tests to compare the variance between the control and treatment group in each period (see Tables 5.3 and 5.4).

Table 5.3: Variance in Treated Task Completion (with session 1 excluded)

Period	1	2	3	4	5	6	7	8
Control	4.592	4.418	4	4.027	3.130	3.101	2.911	1.462
Treatment	4.716	4.340	4.754	4.818	4.568	5.215	4.389	3.596
Difference	-0.124	0.078	-0.754	-0.790	-1.438	-2.114	-1.478	-2.134*

*Significant at 95%-level

Table 5.4: Variance in Untreated Task Completion (with session 1 excluded)

Period	1	2	3	4	5	6	7	8
Control	4.348	3.607	3.955	3.819	3.129	3.114	2.723	1.172
Treatment	4.725	4.767	3.959	4.391	4.643	4.353	4.750	3.715
Difference	-0.377	-1.161	-0.005	-0.572	-1.514	-1.239	-2.027	-2.542*

*Significant at 95%-level

While differences in variance are not significant throughout most of the game, it becomes significant in the final period for both the treated and the untreated task. Thus, it appears that in the final round the intervention does not affect all subjects equally. Instead some subjects appear to reduce contributions a lot, while others maintain high contribution levels. Thus, we see some indication that the social reference intervention cushions the end of game effect in some subjects. It is also interesting to see that in period 3 – when the intervention takes effect – the variance in the completion of treated tasks starts diverging between the control and treatment group, after being very close before. Although, this difference is not significant, this may be an indication for heterogeneous treatment effects of the intervention. This hypothesis is supported by Costa and Kahn (2013), who found significant heterogeneity with regards to the effect of a social reference type intervention in the field.

5.5 Conclusion

This study presented a framework that allows for the detection of spillovers in an experiment with real effort tasks and an underlying public good incentive structure. Results from this experiment are threefold. First, we show that using our novel experimental design produces results that are comparable to a standard public good game. This result validates our design, and informs us, that subjects understood the decision environment to a degree similar to other forms of the public good game. Second, we demonstrate that our social reference intervention successfully increases contributions to the targeted public good. This result confirms our hypothesis and is again well in line with literature investigating the influence of communicating social norms. Finally, our results indicate that the social reference intervention produces a positive spillover, increasing contributions to the untreated public good. This result is less significant, with a significance level just below the 10%-level, which is driven by a diminished effect size of the intervention on the untreated task.

Our experiment was set up in a way to favour the occurrence of spillovers, as the tasks were only marginally differentiated through framing with a common environmental motive. This makes it somewhat obvious for subjects to apply the social reference from the treated task to the untreated task. Nevertheless, we see merit in this result as a baseline from which to generate a systematic analysis framework by varying, e.g.: the similarity of the motive of the two public goods; the public good tasks; group composition in the public goods; or the type of intervention that is applied to the treated public good. Such experiments would greatly increase our knowledge of the occurrence of spillovers.

While we advise caution when translating lab results into advice for policy makers, our results support general notions from the literature that spillover effects should receive more attention when designing and evaluating policy. Our result of a positive spillover in closely related tasks with a similar motive may open up ways to design policy that generates positive spillovers, by making these linkages between activities salient to consumers. However, further research is needed to support this claim both in the lab and the field.

Bibliography

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

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

Andreoni, J., 1988. Why free ride?: Strategies and learning in public goods experiments. Journal of public Economics 37 (3), 291–304.

Benndorf, V., Rau, H. A., Soelch, C., 2014. Minimizing learning behavior in experiments with repeated real-effort tasks. Tech. rep., Available at SSRN: https://ssrn.com/abstract=2503029 (last accessed 26/04/2017).

Catlin, J. R., Wang, Y., 2013. Recycling gone bad: When the option to recycle increases resource consumption. Journal of Consumer Psychology 23 (1), 122–127.

Cherry, T. L., Kroll, S., Shogren, J. F., 2005. The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. Journal of Economic Behavior & Organization 57, 357–365.

Cialdini, R. B., 2003. Crafting normative messages to protect the environment. Current directions in psychological science 12 (4), 105–109.

Cialdini, R. B., Kallgren, C. A., Reno, R. R., 1991. A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. Advances in experimental social psychology 24 (20), 1–243.

Costa, D. L., Kahn, M. E., 2013. Energy conservation “nudges” and environmentalist ideology: evidence from a randomized residential electricity field experiment. Journal of the European Economic Association 11 (3), 680–702.

Cox, J. C., Sadiraj, V., 2007. On modeling voluntary contributions to public goods. Public Finance Review 35 (2), 311–332.

Croson, R. T., 1996. Partners and strangers revisited. Economics Letters 53 (1), 25 – 32.

Dolan, P., Galizzi, M. M., 2015. Like ripples on a pond: Behavioral spillovers and their implications for research and policy. Journal of Economic Psychology 47, 1–16.

Dutcher, E. G., Salmon, T. C., Saral, K. J., 2015. Is ”real” effort more real? Tech. rep., Available at: https://mpra.ub.uni-muenchen.de/68394/ (last accessed 03/05/2017).

Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. The American Economic Review 90 (4), 980–994.

Ferraro, P. J., Price, M. K., 2013. Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment. Review of Economics and Statistics 95 (1), 64–73.

Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental economics 10 (2), 171–178.

Frey, B. S., Meier, S., 2004. Social comparisons and pro-social behavior: Testing ”conditional cooperation” in a field experiment. The American Economic Review 94 (5), 1717–1722.

Gill, D., Prowse, V., 2012. A structural analysis of disappointment aversion in a real effort competition. The American economic review 102 (1), 469–503.

Gill, D., Prowse, V. L., 2011. A novel computerized real effort task based on sliders. Tech. rep., IZA Discussion Paper 5801.

Gneezy, U., Rey-Biel, P., 2015. What is an economic theory that can inform experiments. In: Fréchette, G. R., Schotter, A. (Eds.), Handbook of Experimental Economic Methodology. Oxford University Press.

Goldstein, N. J., Cialdini, R. B., Griskevicius, V., 2008. A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. Journal of consumer Research 35 (3), 472–482.

Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with orsee. Journal of the Economic Science Association 1 (1), 114–125.

Hasson, R., Åsa Löfgren, Visser, M., 2010. Climate change in a public goods game: Investment decision in mitigation versus adaptation. Ecological Economics 70, 331–338.

Heitzig, J., Lessmann, K., Zou, Y., 2011. Self-enforcing strategies to deter free-riding in the climate change mitigation game and other repeated public good games. Proceedings of the National Academy of Sciences 108 (38), 15739–15744.

IEA, 2008. Promoting energy efficiency investments: Case studies in the residential sectors. Tech. rep., International Energy Agency.

Kallgren, C. A., Reno, R. R., Cialdini, R. B., 2000. A focus theory of normative conduct: When norms do and do not affect behavior. Personality and social psychology bulletin 26 (8), 1002–1012.

Keser, C., van Winden, F., 2000. Conditional cooperation and voluntary contributions to public goods. Scandinavian Journal of Economics 102 (1), 23–39.

Lanzini, P., Thøgersen, J., 2014. Behavioural spillover in the environmental domain: An intervention study. Journal of Environmental Psychology 40, 381–390.

Ledyard, J. O., 1995. Public goods: A survey of experimental research. In: Kagel, J. H., Roth, A. E. (Eds.), The handbook of experimental economics. Princeton university press Princeton, NJ.

Littleford, C., Ryley, T. J., Firth, S. K., 2014. Context, control and the spillover of energy use behaviours between office and home settings. Journal of Environmental Psychology 40, 157–166.

Madrian, B. C., 2014. Applying insights from behavioral economics to policy design. Tech. rep., National Bureau of Economic Research.

Mazar, N., Zhong, C.-B., 2010. Do green products make us better people? Psychological science 21 (2), 494–498.

Muehlbacher, S., Kirchler, E., 2009. Origin of endowments in public good games: The impact of effort on contributions. Journal of Neuroscience Psychology and Economics 2 (1), 59–67.

Poortinga, W., Whitmarsh, L., Suffolk, C., 2013. The introduction of a single-use carrier bag charge in wales: Attitude change and behavioural spillover effects. Journal of Environmental Psychology 36, 240–247.

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

Sunstein, C. R., Thaler, R. H., 2003. Libertarian paternalism is not an oxymoron. The University of Chicago Law Review 70 (4), 1159–1202.

Thøgersen, J., 1999. Spillover processes in the development of a sustainable consumption pattern. Journal of Economic Psychology 20 (1), 53 – 81.
URL http://www.sciencedirect.com/science/article/pii/S0167487098000439

Tiefenbeck, V., Staake, T., Roth, K., Sachs, O., 2013. For better or for worse? empirical evidence of moral licensing in a behavioral energy conservation campaign. Energy Policy 57, 160–171.

Truelove, H. B., Carrico, A. R., Weber, E. U., Raimi, K. T., Vandenbergh, M. P., 2014. Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. Global Environmental Change 29, 127–138.

Zizzo, D. J., 2008. Experimenter demand effects in economic experiments. Tech. rep., Available at SSRN: https://ssrn.com/abstract=1163863 (last accessed 04/05/2017).

Appendix

5.A Sensitivity Analysis

Table 5.5: Pooled OLS and Fixed Effects regression results for the treated task without subjects that completed less than 15 tasks in any period

	Dependent variable:		
	Number of treated tasks completed		
	(1)	(2)	(3)
SocialReference	0.083	0.037	0.416**
	(0.300)	(0.323)	(0.207)
Period	−0.158***	−0.160***	−0.211***
	(0.053)	(0.053)	(0.049)
ContributionByOthers	0.170***	0.172***	0.081**
	(0.034)	(0.033)	(0.036)
MeanEnvironmentScore		−0.321	
		(0.217)	
Risk		0.124	
		(0.102)	
GenderMale		0.027	
		(0.314)	
Constant	1.732***	1.626*	
	(0.431)	(0.890)	
Subject fixed effects?	No	No	Yes
Observations	586	579	586

*p<0.1; **p<0.05; ***p<0.01

Note: Heteroscedasticity robust standard errors in parenthesis

Table 5.6: Pooled OLS and Fixed Effects regression results for the untreated task without subjects that completed less than 15 tasks in any period

	Dependent variable:		
	Number of untreated tasks completed		
	(1)	(2)	(3)
SocialReference	0.039	−0.016	0.358
	(0.292)	(0.308)	(0.251)
Period	-0.177^{***}	-0.180^{***}	-0.241^{***}
	(0.052)	(0.050)	(0.050)
ContributionByOthers	0.160^{***}	0.159^{***}	0.057
	(0.035)	(0.032)	(0.035)
MeanEnvironmentScore		−0.280	
		(0.217)	
Risk		0.112	
		(0.096)	
GenderMale		−0.174	
		(0.302)	
Constant	1.905^{***}	1.912^{**}	
	(0.431)	(0.904)	
Subject fixed effects?	No	No	Yes
Observations	586	579	586

*p<0.1; **p<0.05; ***p<0.01

Note: Heteroscedasticity robust standard errors in parenthesis

5.B Script for Lab Assistants

1 INSTRUCTIONS FOR RUNNING THE SPILLOVER EXPERIMENT IN THE LAB

1.1 BEFORE PEOPLE ARRIVE:

- Make sure there is a sufficient amount of money to pay all subjects
- Make sure that all subjects have a keyboard and a mouse which is working
- Check paper in printer
- Print instructions
- Familarize yourself with the instructions

Then:

A. Start a server
B. Start 28 computers in the lab
C. Start Z-tree version 3.4.7, and open the z-leafs on each computer, by:
   a. computer\Z:\zTree\zTree_3.4.7.
   b. Run the file: ztree_experiment_(update_zleaf).bat
   c. Run the file ztree_experiment_(start_ztree).bat
D. Please distribute a copy of the instructions and a pen at each computer.

1.2 BEFORE SUBJECTS ENTER THE LAB:

1. Meet subjects outside the lab. Bring participation list and plastic cards numbered 1-28 (corresponding to the computer numbers)
2. Count number of subjects. If the number is below 28 and not divisible by 4 then wait a few minutes to see if anybody else shows up.
3. Welcome people in a standard greeting, e.g.:

   “Dear all, welcome to this economic experiment. We are ready to begin in a few seconds. Before you enter the lab, please pick a card. Each card has a number between 1-28, and this number corresponds to your computer at which you will sit. When we go in to the laboratory please find the computer which matches your number. Please do not speak when we enter the laboratory and please switch off your cellphone Next to your computer you will find the instructions for today's experiment. Please read these instructions carefully. The experiment will begin when all people have read the instructions and answered a few questions about them. (#)

   Now I will shuffle (*) the cards and then come and pick one.”

   (#): People are playing in groups of 4 so there should be a number of subject divisible by four (20, 24, 28) . If there are too many people please add:

   “Unfortunately we have to send XX people home with their show up fee of XX kr, are there any volunteers?”

   If there are no volunteers then follow the standard procedure and add in XX number of cards with X written on it. Please say:

“Since there are no volunteers, we will add these cards with an X to the deck of cards. If you draw a card with X on it, then please go to the next room, where we will give you your show-up fee”

(*): Please shuffle the card while saying that you do it.

1.3 IN THE LAB:

1.

When people are seated please welcome people again and ask them to read the instructions. Also say that everybody are playing the same game and have the same instructions. It is not necessary to read out the full instructions but please show the instructions and mention that there are four pages. Say the following:
“Dear all, we are now ready to begin, but first, please listen carefully. In this experiment you earn money. Please read the instructions carefully. There are 4 pages of instructions, and they are at your desk next to your computer. You will all play the same game, and you all have the same instructions. After the experiment itself you will be asked to answer a questionnaire.
We will inform you when the experiment is over and pay you your earnings in the room next door. You get 1 kroner for every 6 points you earn during the experiment“

2.

Please start performing these steps while the subjects are reading the instructions. All programs are located in the following folder: *Z:\Experiments\146. Spillover experiment\c. Program (zTree (please update according to the lab folder system)*, and:
The programs are ordered and named 1-4. This is the order in which the programs are to be run. Regarding program 3:

I:
If this is a control session, please run the file that’s called “3_Experiment_CONTROL”.
If this is a treatment session, please run the file that's called "3_Experiment_TREATMENT_DYN"
In the program you have to specify the number of subjects in the “background”, and number of groups.

Then

A. Open the clients table and choose “Shuffle clients”.
B. Run program number 1 (the control questions) state in Z-tree: Number of participants, group=1
C. The answers for the control questions are: **5, 5** (see page 2), **0, 30** (see page 3, section 2.1 and 2.2), **40, 40, 40** (see page 3) and **20, 13, 33** (see page 4). If some are having troubles answering then please help by showing them the relevant information and if necessary help them calculating the answers. However, subjects must type in the right answer themselves.

   After all answered the control questions say out loud:

   “You have now all read the instructions and answered the control questions. There will now be a practice round where you can get familiar with the task but you will not earn point. After this trial round the real experiment will begin, and you will earn points.
   If you have any questions then raise your hand and we will come to your computer.

We will now start the trial round.

D. Run programme number 2 (the practice round) state in Z-tree: Number of participants, group=1
E. Run programme number 3 (the experiment OR control), state in Z-tree: the number of participants **AND,** number of groups =Participants/4). Choose partner matching.
After this programme has run, please say:
"You have now finished the first part of the experiment. The next part of the experiment is a questionnaire. The questions will be shown on your screen shortly. Please note, that your choices in the previous phases will have no consequences for this phase."

F. Run the questionnaire.

1.4 After the session

Pay subjects according to the payment file. Please say:

"The experiment is now finished and you are ready to receive your payment. (*)in the room next door. Please come to the room next door one by one".

(*) Or however the payment is done.

Copy all the files produced by Z-tree to the relevant session folder, located in:

Z:\Experiments\146. Spillover experiment\b. Files by zTree codes

5.C Instructions for Participants

Welcome

You are now taking part in an economic experiment undertaken by researchers from the University of Copenhagen.

By participating in this experiment you can earn money. How much money you earn depends on your decisions, and the decisions of other participants. The following instructions will explain you how. Please read the instructions carefully. You will have to answer control questions to check that you have understood the instructions, and these must be answered correctly before the experiment starts.

During the experiment you are not allowed to communicate with others in the room. If you have any question, just raise your hand and an assistant will come to your desk.

During the experiment your earnings will be calculated in points. These points will be converted to Danish kroner at the following rate:

6 points = 1 DKK

At the end of the experiment your total earnings will be paid to you in cash.

The experiment has two parts: the experiment itself and a questionnaire. The two parts are independent of each other and your behavior in one does not influence the other part in any way. The following instructions explain the experiment. The questionnaire will be explained on your screen.

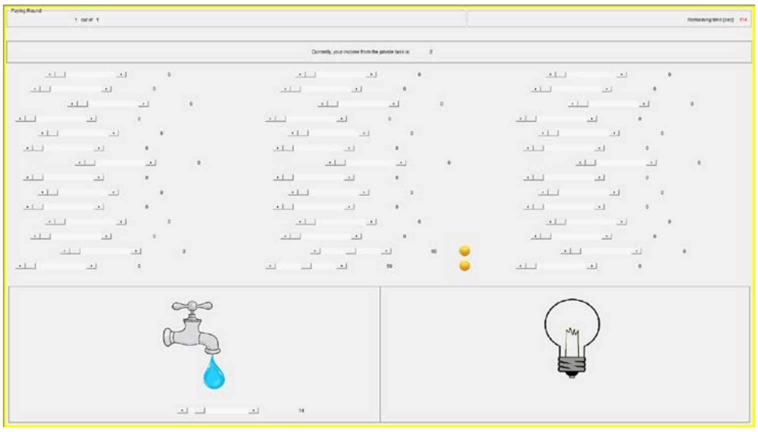
1. Instructions for the Experiment

In this game you will play **8 periods of 3 minutes**. In the beginning of the experiment you will be randomly assigned to a **group of 4 members**, that is, besides you there are 3 other group members. You will be in the same group in every period of the game. Everybody will be anonymous, meaning that you will not know who is in your group, and they will not know you. Each member in your group faces the same situation: You will see a screen with a set of 42 tasks looking like this:

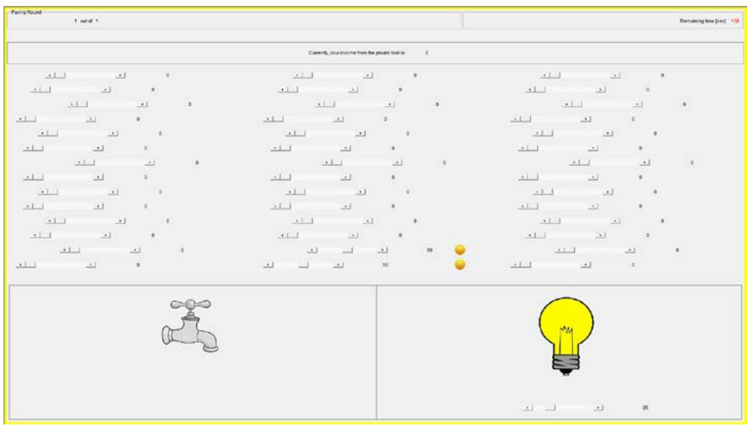
Your task is to move the slider from the left side to the middle of the bar, to a value of exactly 50. Every slider that has been moved to a value of "50" by the end of the period will count as a completed task.

In each period you will see a screen as the ones below. The tasks in the upper part of your screen are called **private tasks**, which will only earn points for you. In the lower part of your screen the **water and energy tasks** appear occasionally. These tasks will earn points to all members of your group (including you). During each round it will be possible to complete **5 water tasks** (in the bottom-left corner) and **5 energy tasks** (in the bottom-right corner).

In this screen you can either do a private task, or the water group task:



In this screen you can either do a private task, or the energy group task:



2. Your earnings

2.1 Private work

When you complete a **private task** in the upper part of the screen a token will appear and you earn **2 points**. That is:

*Earnings from private tasks = sum of private tasks completed * 2*

This task is always available to you. If you finish all the private tasks on one screen, a set of 42 new sliders will appear.

2.2 Water and energy group tasks

In addition to your private work, you can also choose to complete the water and energy tasks that appear during each round. The water task is available when a water drop appears and an energy task is available when the lightbulb lights up. When completed, the water drop will disappear, or the lightbulb will turn off. The same will happen after 10 seconds if you do not complete the group tasks, but you will not earn any points.

5 of each group task will appear on your screen in random order and at random points in time throughout each period**.**

You and each group member will earn **1 point for every water and energy task completed in the group**, that is:

Earnings from group tasks = sum of water tasks completed by group + sum of energy tasks completed by group

Since you are 4 people in your group and each of you face 5 water and 5 energy tasks your total earnings from the group task can reach a maximum of (4*5) + (4*5) = 40, if all group member completes all group tasks.

2.3 Total earnings in each period

Your total earnings in each period are equal to your earnings from your private task plus your earnings from the group tasks. That is

Total earnings = earnings from private work + earnings from group tasks you completed *+ earnings from group tasks completed by others in your group*

After each period you will see a feedback screen reporting your earnings from each round. These earnings will be stored for you and your number of points will determine your final payment.

3. Examples

- If **you complete** 33 private tasks (each giving 2 points), 2 energy tasks and 3 water tasks, and **each of the other people** in your group complete 2 energy task and 1 water task, then your earnings will be:

Earnings from private work	$33 * 2$
+ Earnings from group tasks you completed	$2 + 3$
+ Earnings from group tasks completed by others in your group	$(3 * 2) + (3 * 1)$
Total earnings	80

OR

- If **you complete** 25 private tasks (each giving 2 points), 5 water tasks and 5 energy tasks, and the **other people** in your group **only** complete private tasks, then you will earn:

Earnings from private work	$25 * 2$
+ Earnings from group tasks you completed	$5 + 5$
+ Earnings from group tasks completed by others in your group	$(3 * 0) + (3 * 0)$
Total earnings	60

Once the 8th period is over, you will have to complete a questionnaire in which you are asked to answer questions about yourself.