Replication and Extension of Households' Demand Flexibility in Response to Variable Hourly Electricity

by

Gaurav Singh Thakur

1st Supervisor:	**Enno Schröder**
Chair and 2nd Supervisor:	**Linda Kamp**
Faculty:	Technology, Policy and Management, Delft

Preface

Two years ago, I made one of the most significant decisions of my life leaving my startup life to pursue a master's degree in an entirely new country. This big change started an amazing journey full of challenges as well as rewarding moments. Moving from full-time work to full-time study was not easy. Getting used to academic life again after years of working felt strange at first. But this time has been very rewarding, giving me not only academic knowledge but also important life skills like resilience, adaptability, and determination.

I was lucky to have the guidance of my thesis committee throughout this process.I would like to sincerely thank Linda Kamp, chair of my committee, whose helpful feedback in our meetings guided me in improving my thesis. My special thanks go to Enno Schröder, my first supervisor and mentor, for his constant support from start to finish. His advice, encouragement and patience ,even with my smallest questions, helped me a lot.

I am deeply thankful to my family for always supporting me. Your encouragement has been my greatest strength during this journey. I also want to thank my friends, both near and far across different time zones, for being there through both the fun times and the hard times, creating many special memories.

Looking back, I know that taking this leap of faith was the right choice. Even with its ups and downs, this experience has helped me grow and given me a better understanding of my goals and passions.

Gaurav Singh Thakur
Delft, August 2025

Abstract

As energy systems transition towards renewable sources, creating challenges for grid stability, household demand-side flexibility has become a cornerstone of modern policy. Dynamic electricity pricing is one of the most prominent tools for activating this flexibility, yet the robustness of its effects and the heterogeneity of responses across socioeconomic groups remain under-explored. The study addresses these gaps by analyzing data from a large-scale randomized controlled trial in Norway, conducted from December 2020 to March 2021, involving 3,746 households.

The study replicates the fixed-effects panel-data model used by Hofmann and Lindberg (2024), to verify the average treatment effect of variable hourly pricing and extends it with interaction terms to test whether household income moderates price responsiveness.

The replication analysis successfully reproduces the original findings, confirming a statistically significant average peak-hour demand reduction of 2.92% (0.085 kWh/h). The subsequent extension reveals a strong and systematic inverse relationship between household income and price elasticity, with low-income households significantly reducing peak consumption by 12.05% (0.221 kWh/h), compared to negligible responses from high-income households, particularly during high-price events for shorter time periods.

These findings confirm that dynamic pricing effectively elicits demand response but is disproportionately driven by financially constrained households. To ensure an effective and equitable energy transition, policymakers must adopt segmented pricing strategies that account for socioeconomic differences. Limitations, including the Norway-specific context, suggest caution in generalizing these results.

Contents

List of Figures

List of Tables

1 Introduction

The transition to renewable energy sources introduces significant challenges for grid stability due to their intermittent nature, making household demand-side flexibility a critical component of modern energy policy. Dynamic electricity pricing, where rates vary based on supply and demand, is a primary tool for encouraging households to reduce consumption during high-price periods, thereby supporting grid reliability and renewable integration. While prior studies (Faruqui et al., 2017; Harding & Sexton, 2017) demonstrate demand reductions during peak prices, the robustness of these findings and their variation across socioeconomic groups remain underexplored. If dynamic pricing effects vary across income levels, policies designed to improve efficiency may unintentionally place a disproportionate burden on financially vulnerable households.

This study aims to address these gaps by replicating and extending the study conducted by Hofmann and Lindberg (2024) as a part of Statnett's iFlex project (Hofmann & Siebenbrunner, 2023),a large-scale field experiment conducted in Norway from December 2020 to March 2021 with 3,746 households. The replication tests the robustness of these results , which report average electricity demand was reduced by 2.92% during hours in which high prices were in effect. This reduction remained consistent across repeated interventions, indicating that there was no significant decline in response over time. Furthermore, the response was more pronounced when the high-price signals were in effect for shorter duration and when the price exceeded a threshold of 15 NOK/ kWh. The authors also reported a 1.57% reduction in electricity use on non experiment days, suggesting behavioral change within households that persisted beyond active treatment periods. No significant rebound effect was observed, meaning that households did not compensate for reduced usage during high-price hours by increasing consumption before or after the price peaks. Instead, a net reduction in electricity consumption was recorded across the entire day during experimental periods. The empirical analysis was based on a fixed effects panel data model, controlling for time-invariant household characteristics and region-specific time shocks.

The original study provided strong evidence of demand flexibility in Norway, yet it did not consider the role of household income. This research therefore has two primary objectives. The first is to assess the robustness of the original findings. By independently developing the analytical code and applying it to the original dataset, it provides a critical test of the reproducibility of the core results. This replication practice is important as it ensures transparency and verifies that the original findings are not dependent on undocumented implementation choices(Colliard et al., 2022).

The second contribution of this study is to expand on this validated approach to examine how household income moderates the responses to dynamic pricing signals. Although the original study found an average demand reduction of 2.92% during peak hours, it is unclear whether this effect exhibits heterogeneity across consumers. Economic theory suggests that lower-income households, facing tighter budget constraints, should exhibit greater price elasticity. However, they may also lack access to flexible loads, such as electric vehicles and smart appliances, which facilitate demand shifting. This creates an empirical tension that our research aims to resolve.

1.1. Research Questions

By answering the following questions, this study strengthens the empirical foundation of dynamic pricing research and provides crucial, policy-relevant evidence on its distributional effects. The findings confirm the results of the original study and reveal that lower-income households are, indeed, more sensitive to price signals, particularly during events when prices are high, but for a short period of time. This highlights the importance of designing policies that account for socioeconomic differences, ensuring that the energy transition is both efficient and equitable.

- Can the main findings of Hofmann and Lindberg (2024) be reproduced using the same data and econometric approach?
- How does household income moderated the responsiveness to dynamic electricity pricing signals ?

2
Literature Review

2.1. Dynamic Pricing Policies

The ongoing decarbonization of the global energy sector, marked by a large-scale integration of intermittent renewables, has rendered traditional, static electricity tariffs increasingly obsolete.To manage the grid instability inherent in a renewables-dominant system, policymakers have converged on demand-side flexibility as a cornerstone of modern energy strategy. Dynamic pricing has emerged as the principal instrument for activating this flexibility, fundamentally reshaping the relationship between utilities and consumers. This transition, however, is not a marginal policy adjustment but a capital-intensive, system-wide intervention with significant economic and social ramifications.

The scale of this intervention underscores the gravity of its evidence base. In the United States, for example, the deployment of the necessary Advanced Metering Infrastructure (AMI) was catalyzed by over $25 billion in combined federal and utility investment, leading to the installation of more than 65 million smart meters by 2015 (Harding & Sexton, 2017). Similarly, policy mandates in the European Union have driven mass adoption, with nations like Italy achieving near-universal smart meter penetration (Kim & Shcherbakova, 2011). Such profound investment is predicated on a key economic assumption: that leveraging demand-side flexibility is a more efficient allocation of capital than constructing and maintaining costly peak power plants, which are utilized for only a small fraction of the year (Azarova et al., 2020).

The reliance on this evidence creates a high-stakes environment where the consequences of acting on flawed or non-robust findings are severe. At a systemic level, an overestimation of demand flexibility could compromise grid stability during periods of peak stress. Economically, it risks the misallocation of billions in infrastructure capital.Socially, and perhaps most critically, it risks imposing inequitable cost burdens on vulnerable households, whose capacity to respond to price signals may be limited. The integrity of the underlying empirical research is therefore not merely an academic ideal but a prerequisite for effective and equitable policy design.

Consequently, the field of applied economics has increasingly adopted Randomized Controlled Trials (RCTs) as the basis for generating reliable policy evidence. By providing a clean, unbiased estimate of causal effects, RCTs are the primary tool used to navigate the risks inherent in large-scale policy implementation. The foundational evidence for dynamic pricing, therefore, is built upon the results of a select number of these pivotal field experiments. The validity, robustness, and generalizability of these cornerstone studies are thus of paramount importance, creating a clear and urgent mandate for their rigorous scientific verification, which forms the primary focus of the review that follows.

2.2. Importance of Replication

To mitigate the risks of policy failure outlined previously, the scientific community relies on a clear and rigorous process for validating empirical findings before they are widely accepted or used to inform policy. A replication study involves repeating an existing experiment or analysis either under identical or modified conditions to assess whether the original results are robust and applicable across different contexts (McMillan, 2017). While the concept is foundational to the scientific method, a recent credibility crisis across many applied fields, including economics, has renewed the focus on its importance. Studies have found to be non-replicable, leading to a necessary re-evaluation of what constitutes trustworthy evidence (Bergh et al., 2017; Camerer et al., 2016). This scholarly mandate for verification is therefore not just a matter of good practice, it is the primary mechanism by which a research field builds a stable and reliable foundation.

The role of replication is crucial and its importance is underscored by the substantial real-world costs of acting on flawed research. Replication serves as the ultimate test of a finding's reliability. When an independent researcher can reproduce a result, it significantly strengthens confidence in the finding's validity (Hensel, 2021).On the other hand, when results can't be repeated, it can help stop the spread of incorrect ideas. For example, some energy efficiency policies were based on mistaken assumptions and didn't consider the rebound effect, which in some cases led to higher overall energy use and worked against sustainability goals (Lijnis Huffenreuter, 2013). Replication also serves a critical corrective function, uncovering unintentional errors or questionable research practices that might otherwise go unnoticed (Block et al., 2023; Duvendack et al., 2017).

Even though replication is clearly important, it often runs into ongoing problems in both research systems and daily practice. Academic journals have traditionally favored new and surprising results, which makes it harder for studies that repeat or check earlier work to get published (Block et al., 2023; Hensel, 2021). This focus on originality can make researchers less likely to do important follow-up studies (Coffman et al., 2017). On top of that, practical issues—like limited access to private data or unclear methods—can make it hard or even impossible to repeat earlier studies (Hensel, 2021). A review by Chang and Li (2015) found that only about one-third of key findings in a group of economics papers could be repeated without help from the original researchers.

These challenges are particularly serious in applied energy economics. Given that dynamic pricing rests on economic theories of elasticity and incentives—concepts that have themselves faced questions of empirical reliability, the need for systematic replication is especially urgent (Ankel-Peters et al., 2025; Colliard et al., 2022). The reported effects of dynamic pricing programs can vary substantially, with peak demand reductions ranging from under 5% to over 20% across different studies (Faruqui et al., 2017). Without replication, it is difficult to ascertain whether this variation stems from genuine contextual differences or from the methodological artifacts that replication is designed to detect. The energy economics community has begun to recognize this gap, with journals starting to call for and create dedicated formats for replication studies (Huebner et al., 2017). This marks an important shift, acknowledging that for a field whose findings directly inform major infrastructure and policy decisions, there is a clear scholarly imperative to first verify its foundational evidence.

Pricing rests on economic theories of elasticity and incentives concepts that have themselves faced questions of empirical reliability so the need for systematic replication is especially urgent (Ankel-Peters et al., 2025; Colliard et al., 2022). The reported effects of dynamic pricing programs can vary substantially, with peak demand reductions ranging from under 5% to over 20% across different studies (Faruqui et al., 2017). Without replication, it is difficult to ascertain whether this variation stems from genuine contextual differences or from the methodological artifacts that replication is designed to detect. The energy economics community has begun to recognize this gap, with journals starting to call for and create dedicated formats for replication studies (Huebner et al., 2017). This marks an important shift, acknowledging that for a field whose findings directly inform major infrastructure and policy decisions, there is a clear scholarly imperative to first verify its foundational evidence.

2.3. The Findings in the Dynamic Pricing Literature

To identify a cornerstone study for replication, it is first necessary to understand the prevailing findings that constitute the evidence base for dynamic pricing. A significant body of research has established a broad consensus that households do respond to time-varying electricity rates, yet the magnitude and nature of this response are highly contingent on a range of factors. While generally effective, meta-analyses consistently show that residential electricity demand is highly price inelastic in the short run, with typical price elasticity estimates around -0.06. This responsiveness increases over time as consumers adjust their habits and technology, with long-run elasticity estimates reaching approximately -0.43 (Labandeira et al., 2017; Zhu et al., 2018).Complementing cross-country meta-evidence, Brazilian household-level panel data (São Paulo, 1998 & 2008) estimate short-run elasticities of about -0.50 for price and $+0.21$ for income, confirming inelastic but non trivial residential responsiveness (Uhr & JÃºlia Gallego Ziero Uhr, 2017).

A complementary strand of work emphasizes behavioral mechanisms that mediate price response. Real-time feedback reduces inattention and improves the timing of actions, while social-comparison messages that benchmark a household against similar peers generate modest but persistent conservation effects, layered onto dynamic prices, these nudges increase event-hour responsiveness (Allcott, 2011; Brewer, 2023; Buckley, 2020; Faruqui et al., 2017).

Empirical evidence since 2010 shows that households do respond to dynamic pricing by lowering usage, but the magnitude of peak demand reduction varies greatly by pricing scheme. Surveying 15 major pilot studies, Faruqui and Sergici (2013) found Time-of-Use (TOU) rates (with predictable daily peak/off-peak prices) typically cut peak-period demand by only about 3–6%, whereas Critical Peak Pricing (CPP) tariffs (with occasional very high prices during critical events) achieved 13–20% peak reductions. When CPP was combined with enabling technologies (like smart thermostats that automate load shedding), peak reductions of 27–44% were observed. These findings were echoed by Newsham and Bowker (2010)'s review of North American programs: they concluded that CPP with automated control technology is the most effective strategy, yielding peak load cuts on the order of 30% or more without undue hardship on customers (given override options), whereas a simple TOU tariff achieves only around 5% peak reduction. In other words, critical-event pricing tends to induce much larger demand response than everyday TOU pricing, especially if paired with technology that relieves households from manual effort. Real-world CPP trials also suggest diminishing returns to ever-higher peak prices for example, a California experiment found that raising the critical price from \$0.50 to \$0.68/kWh did not elicit any additional load drop, indicating a saturation in price responsiveness beyond a point.

Limited rigorous studies directly on Real-Time Pricing (RTP) in homes (with prices varying hourly) and Critical Peak Rebate (CPR/PTR) programs have emerged. Analyses of the Washington D.C. PowerCentsDC pilot (2008–09) which offered households CPP, CPR, or RTP showed all three designs produced substantial peak reductions, with CPP yielding slightly larger cuts than an equivalent CPR rebate and RTP prompting similar usage cuts given the same price levels. A recent meta-analysis by an RFF team likewise finds CPP-only rates tend to deliver the greatest peak demand impacts, outperforming TOU or CPR programs on average. Still, that study noted that evidence on RTP remains sparse in the residential sector, and peak-time rebate outcomes are highly variable with fewer trials to draw on. Broadly, dynamic pricing works on average a 16% reduction in peak demand per participant is reported across 100+ experiments but the exact impact hinges on the pricing format and context.

Beyond the tariff type, specific program design features can crucially shape household response. Experimental studies indicate that households respond more when events are salient and manageable. For instance, the length and timing of peak periods matter: A recent California TOU rollout analysis found consumers saved most energy on the hottest peak-demand afternoons, suggesting that aligning peak pricing windows with truly high-stress hours (e.g., late afternoons) maximizes the response. On the other hand, very long or too-frequent peak events could trigger fatigue a field trial in Norway observed diminishing conservation in later events, though some customers learned to adapt over time. A Korean Peak-Time Rebate experiment (Hwang et al., 2025) used machine learning to track households across multiple events and found high heterogeneity: some customer segments consistently curbed usage, even improving with practice (learning effects), while others showed negligible response. This implies that the frequency of critical events and how households are coached or targeted, can influence whether responses persist or wane.

Advance notification is also believed to impact outcomes, though formal studies are limited. Most CPP programs give day-ahead notice of events; this lead time likely helps households plan load shifts (e.g., precool the house or defer appliances) as evidenced by pre-cooling behavior observed before peak pricing hours in some trials. If events are called with only short notice (or very unpredictably as in RTP), response may be muted unless automation is in place. Indeed, multiple experiments underscore the importance of enabling technology and information to help customers respond on short notice. Brewer (2023) demonstrated that simply installing an in-home display to show real-time usage and prices roughly doubled the price elasticity of households under critical peak pricing, as the continuous feedback made them more aware and engaged. In their RCT, CPP without the device yielded smaller and more inconsistent savings, whereas CPP with an immediate feedback display induced robust conservation during events. These results highlight that salience and convenience via timely information or automation can be as vital as the price levels themselves in dynamic pricing programs.

Dynamic pricing taps into behavioral responses and several studies have probed the human factors and fairness concerns involved. A key insight is that automation dramatically lowers the behavioral barriers to demand response. Bollinger and Hartmann (2017) find that when households had smart thermostats set to react to peak prices, their peak electricity reductions were far greater than those of households relying on manual efforts. In effect, automation removes adjustment costs , enabling consumers to curtail load with minimal inconvenience. By contrast, purely relying on people to notice price signals and take action introduces friction many will ignore or forget short-term price changes if not automated. This aligns with earlier findings that enabling tech or direct load control can boost peak savings substantially (often two- or three-fold) compared to information or price signals alone.A meta-analysis of 63 pilots

(337 treatments, nine countries) found a non-linear “arc” in price responsiveness: a 10% rise in the peak-to-off-peak price ratio cut peak use by 6.5%, with enabling technologies adding 4.6% for an overall 11.1% reduction. Effects tapered at high ratios, indicating behavioural saturation (Faruqui et al., 2017).Behavioral research also suggests consumers might struggle with complex tariffs: evidence from pilot programs with multiple peak periods indicates that if a pricing scheme is too complex, people primarily shift usage out of peak overall rather than finely tuning between tiers. Simpler, more salient designs tend to be more effective in guiding behavior.

Equity considerations are increasingly prominent. Policymakers worry that time-varying rates could disproportionately burden certain groups for example, those who are home all day (such as retirees or low-income families with young children) might have less flexibility to reduce peak usage, or conversely, some vulnerable households might curtail essential consumption to avoid high prices. Empirical evidence on distributional effects is mixed. Some studies have observed heterogeneous responses by income and usage level. In California’s CPP trial, higher-consumption households delivered larger absolute and percentage load reductions, suggesting well-resourced homes (often with central AC) had more discretionary load to shed. Pellini (2021), analysing 30 European countries with methods robust to outliers and structural breaks, found long-run income elasticities of 0.19–0.93 and price elasticities of –0.08 to –0.80, confirming that electricity is a necessity and typically price-inelastic. Higher-income, energy-efficient countries tended to show lower elasticities, suggesting that price signals alone may yield smaller behavioural changes in such contexts. On the other hand, one dynamic pricing experiment reported low-income customers responding as much or more than higher-income customers, indicating strong price sensitivity in financially constrained groups. Harding and Sexton (2017) likewise document that household response varies significantly with demographics and weather meaning customer characteristics influence who wins or loses under dynamic pricing. Because of such variations, recent reviews have flagged the need for more research into how low-income and other vulnerable populations fare with dynamic pricing. If, for instance, low-income households have fewer smart appliances or cannot easily shift consumption, they might see smaller bill savings (or even bill increases) compared to affluent, tech-equipped households an important equity issue.Simionescu and Cifuentes-Faura (2024) show that in the V4 countries (Poland, Hungary, Czech Republic, Slovakia), income inequality significantly increases energy poverty, measured by arrears on utility bills and inability to keep homes warm. Renewable energy use has mixed effects—reducing the share of households unable to heat adequately, but increasing utility bill arrears, likely due to higher consumer prices from renewable surcharges.

In summary, the last decade of experimental studies confirms that residential dynamic pricing can yield meaningful peak demand reductions, especially under CPP or RTP schemes, but program design details are pivotal. Short, infrequent high-price events with adequate notice (or automation) tend to maximize effectiveness without overwhelming consumers. Supportive tools smart thermostats, real-time displays, informative alerts greatly enhance the behavioral response to price signals. However, one-size-fits-all programs may produce uneven outcomes across different households. Dynamic pricing is most effective and fair when paired with measures to engage and protect less-flexible consumers, ensuring that the benefits of demand response can be realized without compromising equity. As the transition to smarter grids continues, ongoing rigorous trials and meta-analyses are helping identify best practices for dynamic pricing program design and mitigation of any unintended social impacts.

Finally, the prevailing understanding is that external context, particularly weather, significantly influences the potential for demand flexibility. Several studies have documented a distinct, non-linear relationship between outdoor temperature and price responsiveness. In regions with significant heating or cooling needs, demand flexibility is often highest during moderately hot or cold weather. This effect diminishes in extreme weather, when heating or cooling becomes a non-negotiable necessity, or in mild weather, when there is little discretionary consumption to shift in the first place (Hofmann & Lindberg, 2024). This body of work collectively provides a detailed picture of how, when, and why households are reported to respond to dynamic prices, establishing a clear set of foundational claims that are ripe for the rigorous verification that replication provides.

2.4. The Hofmann and Lindberg (2024) Study for replication

Within this body of literature, the recent work by Hofmann and Lindberg (2024) stands out as a methodologically rigorous and comprehensive cornerstone, making it an ideal candidate for the verification that this review has argued is essential. The study, a large-scale Randomized Controlled Trial involving 3,746 households over three winter months, was conducted in the Norwegian electricity market a setting of significant international relevance. As the authors note, Norway represents a system, characterized by its high degree of electrification in heating and transport, near-universal smart meter penetration since 2019, and a consumer base already accustomed to variable hourly electricity prices through the Nord Pool market (Hofmann & Lindberg, 2024).

The strategic importance of this context is amplified by the region's trajectory. Recent analyses of the Nordic energy system indicate that the very conditions that make demand flexibility critical are set to intensify. Regional energy outlooks project substantial growth in electricity consumption through 2050, driven by the continued electrification of industry and transport. This rising demand will be met primarily by intermittent renewable sources, which is expected to tighten the region's capacity balance and lead to more frequent and volatile price peaks. This future context elevates the importance of the original study and its findings on how to elicit demand response are not just a snapshot of a past system, but a vital source of insight for managing the more challenging grid conditions of the coming decade.

The methodological design of the study further solidifies its benchmark status. Hofmann and Lindberg (2024) tested a comprehensive set of 14 distinct price signals, allowing for a nuanced analysis of how factors like peak price magnitude and duration affect consumer response. Their explicit goal was to address key unresolved questions in the literature regarding the persistence of response and the impact of external factors like weather. By integrating these multiple layers of analysis into a single, coherent RCT framework, the study provides a state-of-the-art model for understanding household demand response. Consequently, its findings represent a critical set of foundational claims that warrant the rigorous verification undertaken by this thesis.

2.5. Research Gap

This review has constructed a clear, logical case for the research undertaken in this thesis. It began by establishing the high policy stakes of dynamic electricity pricing, a tool being deployed at immense scale with significant financial and social consequences. Against this backdrop, it articulated the scholarly imperative of replication as the primary mechanism for verifying the foundational evidence upon which such critical policies are built. The review then surveyed the prevailing findings within the dynamic pricing literature, identifying the key reported effects related to price elasticity, behavioral factors, and technology. Finally, it positioned the recent, comprehensive, and methodologically rigorous of the original study by as a cornerstone of this evidence base, a benchmark whose relevance is only increasing given the future trajectory of the energy transition.

The logical conclusion of this narrative is that a rigorous, independent replication of this benchmark study is a necessary and valuable scientific contribution.

While verifying the average treatment effect of the original study is a necessary first step, a critical limitation of its analysis and a significant gap in the broader literature is the lack of insight into how demand response varies by socioeconomic status. Understanding this heterogeneity is not merely an academic exercise, it is essential for designing policies that are both effective and equitable. A uniform approach that overlooks differences in income comes with significant risks. It may fall short if it doesn't engage households with high energy use, and more importantly, it could place a heavier burden on lower-income households, worsening existing inequalities.

The literature reveals a clear theoretical and empirical disagreement on this issue. One view argues that lower-income households should respond more to price changes because of stricter budget limits. However, another strong perspective is that these households often lack the means or options to adjust their energy use easily, lacking the high-quality housing, smart appliances, or discretionary consumption needed to meaningfully respond (Hao et al., 2024). This ambiguity is highly problematic for policymakers. If a policy's success relies on the response of vulnerable households defined in some European contexts as the 9% of the population spending over 8–10% of their income on energy (Martinez-Reyes et al., 2025). It risks forcing them to sacrifice essential comfort to avoid punitive costs. Conversely, if these households are unable to respond, they may face increased bill volatility without recourse, effectively being punished by a system they cannot participate in ("Empowering vulnerable consumers in the energy transition - European Commission," n.d.).

The original study, by not stratifying its analysis by income, provides no insight into this critical trade-off. Its rich dataset, however, offers an ideal opportunity to contribute to this unresolved debate. Therefore, contingent upon the successful replication of the original findings, this thesis will pursue a secondary objective.

3 Methodology

This study replicates and extends the empirical study conducted by Hofmann and Lindberg (2024), which assessed household responsiveness to electricity price signals using a randomized controlled trial. While the core experimental design, data sources, and treatment allocation is same as the original setup, the present analysis introduces an additional layer by examining the response across income groups.

The chapter provide a detailed account of the experimental design, recruitment procedures, intervention protocols, and data sources. Particular attention is given to the construction of income categories and the methods used to estimate differential treatment effects across these subgroups.

3.1. Experimental Design

This study analyzes data from the iFlex project, designed to assess household demand flexibility under real-time electricity pricing, a critical tool for reducing peak demand to enhance grid stability and support renewable energy integration. The Statnett's iFlex project was conducted in two phases, this analysis mainly focuses on Phase 2 from December 2020 to March 2021 with full-scale experiment. Phase 1 served as a pilot to evaluate procedures and instrumentation, as reported in Hofmann and Lindberg (2021).

The experiment took place in five Norwegian cities Oslo, Bergen, Trondheim, Tromsø, and Bodø. These cities were chosen to reflect geographical and climatic diversity, helping to improve the external validity of the results.

Table 3.1 provides a summary of the key characteristics of this filtered Phase 2 dataset, outlining the scale and scope of the data used for this analysis.

The experiment used a Randomized Controlled Trial (RCT), a method well-regarded for its strength in establishing clear causal relationships between demand reduction and price signals (Dunning, 2016). In line with standard practice in demand response research (Azarova et al., 2020; Cappers, 2013; Frederiks et al., 2016), randomization helps ensure that any measured effects stem from the intervention only, and not from other confounding factors.

In the treatment group, households received real-time price signals during specific high-price events. These typically occurred in the early morning and late afternoon. Participants were notified a day in advance through SMS or app alerts. Control households, while subject to the same conditions, received no pricing information.

Table 3.1: Summary of the iFlex Experiment's Characteristics

Participating households:	3746
Pricing experiment	
Time period:	16 December 2020 to 25 March 2021
Price signals tested:	14
Number of treatment groups:	11
Sum of experiment days of all treatment groups:	254
Electricity data	
Time period:	1 December 2020 to 26 March 2021
Hours excl. weekend and holidays:	1776
Average consumption:	2.52 kWh/h
Survey data	
Time period:	12 May to 10 June 2020
Number of questions:	43
Responses from project participants:	1733
Responses from the population:	1800

Source: Adapted from Hofmann and Lindberg (2024)

The aim was to test demand-side flexibility under realistic conditions. No specific instructions were given on how to reduce consumption households responded based on the price signals alone. To ensure that the pricing events aligned with actual periods of high electricity use, experimental days were selected dynamically using short-term weather forecasts, especially during colder days to ensure that pricing events coincided with expected high electricity demand, especially related to heating.

A full explanation of the data collection process and access to the complete iFlex dataset is available in (Hofmann & Siebenbrunner, 2023).

3.2. Data

The analysis focuses on Phase 2 of the iFlex project, which ran from 1 December 2020 to 26 March 2021. The actual intervention began on 16 December. Following the original study's methodology, several periods were excluded from the analysis to ensure a consistent comparison of typical consumption patterns. Specifically, weekends and public holidays were removed. Furthermore, the period between December 24, 2020, and December 31, 2020, was excluded as it corresponded to the Christmas school holidays, a time when household routines and energy use are not representative of typical weekday behavior.

Electricity consumption was recorded using smart meters installed in all 3,746 participating households. These meters captured hourly usage, allowing detailed tracking of patterns before, during, and after pricing events. To account for weather-driven changes in demand, especially related to heating, hourly temperature data were included. These came from the Norwegian Meteorological Institute. One weather station was used per city: Oslo, Bergen, Trondheim, Tromsø, and Bodø. The temperature data supported the construction of personalized baselines and served as control variables in the main analysis. Further details are provided in the dataset documentation by Hofmann and Siebenbrunner (2023).

To complement the meter and weather data, a follow-up survey was carried out after the experiment ended. It collected information on household demographics and behaviors, including income, number of residents, housing type, electricity contract type, and ownership of electric vehicles or heat pumps. This survey data is crucial for the second part of the analysis, which examines the heterogeneity of treatment effects across different demographic subgroups, particularly income. Income data were available for 1,617 households; although this represents less than half the sample, it was sufficient for subgroup comparisons. There is no clear evidence that survey respondents differed in ways that would bias the main estimates.

3.3. Recruitment and Randomised Group Assignment

A total of 3,746 households participated in the iFlex experiment. Recruitment took place in late 2020, ahead of the experiment period starting in December. To capture geographic and climatic variation, households were drawn from five Norwegian cities. Three national electricity suppliers handled recruitment, using digital channels to reach customers.

Participation in the experiment was voluntary. Online campaigns hosted by the suppliers invited households to join. To be eligible, households needed a smart electricity meter installed as part of a prior national rollout. They also had to provide informed consent for data sharing and participation. During the experiment, pricing event alerts and other notices were delivered through digital platforms such as SMS or mobile apps.

Incentives were offered to encourage participation. The treatment group could earn financial rewards by reducing electricity use during high-price periods. Control group households, though not eligible for those payments, could still receive vouchers or prize draw entries, depending on their supplier. These incentives helped boost participation and reduce dropout rates.

Households were randomly assigned to either a treatment or control group using stratified randomization within each city. This ensured a balanced distribution across groups based on location and climatefactors known to influence electricity consumption.

Table 3.2: Comparison of electricity consumption between control and treatment groups in each region.

Region	**Mean electricity consumption in kWh/h (number of households)**			
	Control group	**Treatment group**	**Difference**	**p-value**
Bergen	2.52 (111)	2.25 (218)	0.27	0.085
Bodø	2.24 (214)	2.26 (221)	-0.02	0.896
Oslo	2.34 (257)	2.22 (1530)	0.12	0.271
Tromsø	2.48 (266)	2.53 (524)	-0.05	0.665
Trondheim	1.98 (204)	2.04 (201)	-0.06	0.690

****p < 0.001, **p < 0.01, *p < 0.05.*

Note: The p-value of the difference was calculated with Welch's t-test.

While the voluntary nature of the study raises the risk of self-selection bias, randomization limits concerns about systematic group differences. Baseline electricity use, measured as average hourly consumption before the intervention, was compared between groups. As shown in Table 3.2, mean hourly electricity consumption across the five study regions. The differences between treatment and control groups are minimal, ranging from −0.06 to 0.27 kWh/h. None of the differences are statistically significant at the 5% level. The lack of statistical significance supports the assumption that the groups were comparable at baseline.

The combination of randomized group assignment, balance checks, and structured pricing signals provided a strong foundation for estimating causal effects of price-based interventions on household electricity use.

3.3.1. Pricing Signals and Incentives

On dynamically selected event days, which were chosen based on short-term weather forecasts to coincide with periods of high heating demand, households in the treatment group received dynamic electricity price signals. These alerts warned of temporary spikes in electricity prices and were delivered 24 hours in advance by SMS or app notification, giving participants time to adjust their electricity use.

Each pricing event followed one of 14 predefined price profiles. These profiles varied in both peak level and duration. Prices ranged from 2 to 30 NOK/kWh, covering common and rare market situations, including extreme conditions. The events lasted between 2 and 11 hours, often scheduled during typical peak demand windows, such as early mornings or late afternoons.

Figure 3.1 displays the structure and timing of high-price signals sent to different treatment groups. Each row represents the hour of the day, while columns represent the assigned price signal. Orange blocks mark peak hour events. Prices outside the peak periods were generally set to 1 NOK/kWh in all profiles except P0, where the off peak prices are 0 NOK/kWh, allowing for investigation into whether non-peak hour pricing affected peak hour response.At the time, Nordpool's maximum permitted clearing price was 52 NOK/kWh, whereas the average home power price in Norway, including grid tariffs and taxes, was approximately 1 NOK/kWh. The frequency and duration of price signals across groups was determined to cover various climatic conditions.

The design draws from critical peak rebate programs used in earlier studies (Ito et al., 2018). The rebate model was adopted for two main reasons. First, recruitment would likely have failed under a real-time pricing contract with high price spikes. Instead, offering compensation for reduced demand avoided changes to existing tariffs and made participation more appealing. Second, the rebate structure gave researchers room to test a wide range of price levels without affecting the billing systems households were used to.

To measure response, financial rewards were tied to the amount of electricity saved during pricing events. Savings were calculated against a personalized baseline. This baseline used the average hourly electricity use from the ten most recent non-event weekdays. It was also adjusted to account for day type and weather conditions, following the method outlined by Hofmann and Siebenbrunner (2023)

Payments were issued after the event period ended. Depending on the supplier, rewards were credited directly to electricity bills or transferred to participants' bank accounts.

While this reward model was effective in shaping behavior, it may not fully replicate the psychological effects of actual cost savings. Participants received feedback and payment after the event, rather than facing immediate higher prices. This could shape their perception of urgency and value. The possible consequences for how demand flexibility is measured are discussed further in Chapter 5.

Hour	A_10	A_5	B_10	B_5	C	P_10	P_5
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	0.99	0.99	0.99	0.99	0.99	0.99	0.99
3	0.95	0.95	0.95	0.95	0.95	0.95	0.95
4	0.95	0.95	0.95	0.95	0.95	0.95	0.95
5	0.96	0.96	0.96	0.96	0.96	0.96	0.96
6	1.00	1.00	1.00	1.00	1.00	1.00	1.00
7	5.48	2.74	5.48	2.74	2.74	1.00	1.00
8	9.95	4.98	9.95	4.98	4.98	1.02	1.02
9	10.13	5.07	10.13	5.07	5.07	1.04	1.04
10	10.06	5.03	10.06	5.03	5.03	1.03	1.03
11	9.88	4.94	9.88	4.94	4.94	1.02	1.02
12	7.46	3.73	9.95	4.98	3.73	1.01	1.01
13	5.04	2.52	10.13	5.07	2.52	1.06	1.06
14	5.00	2.50	10.00	5.00	2.50	1.06	1.06
15	4.96	2.48	10.06	5.03	2.48	1.09	1.09
16	7.46	3.73	9.88	4.94	6.22	1.20	1.20
17	9.95	4.98	9.95	4.98	9.95	9.95	4.98
18	10.13	5.07	10.13	5.07	10.13	10.05	5.03
19	10.06	5.03	10.06	5.03	10.06	1.20	1.20
20	9.88	4.94	9.88	4.94	9.88	1.10	1.10
21	5.49	2.75	5.54	2.77	5.49	1.10	1.10
22	1.10	1.10	1.10	1.10	1.10	1.10	1.10
23	1.06	1.06	1.06	1.06	1.06	1.06	1.06
24	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Profile_X : X represents the price when the profile (A,B,C,P or P0) is active.
Note : Out of 14 , these profiles are selected to enable clear comparison across same price levels(5 Nok and 10 NOK).

Figure 3.1: Timing and duration of high-price events by treatment subgroup

3.3.2. Key variables

These variables define the structure of the treatment and control conditions. They capture experiment timing, price signals, household characteristics, and environmental controls. They support the estimation of the household-level demand response with adjustments for respective variations. Income-related variables are included to identify differences in price sensitivity between socioeconomic groups.

To examine how demand response varies by income, the analysis includes interaction terms between household income groups and the peak-hour treatment indicator. This allows the model to capture any systematic difference in responsiveness between, for example, low- and high-income households, while still controlling for household and time-specific fixed effects.

Household income was one of the key demographic variables collected through the post-experiment survey. Of the 3,746 participating households, 1,617 provided valid income data, which is about 43% of the sample. While more than half of participants did not report income, the available responses are sufficient for subgroup analysis and are used in several

Table 3.3: Overview of Variables

Variable	**Description**
D_{it}^{exp}	All hours on experiment days for treated households
D_{it}^{peak}	Peak hours on experiment days
D_{it}^{nopeak}	Non-peak hours on experiment days
$D_{it}^{notreat}$	Non-treatment hours after the first experiment day
D_{it}^{x}	Treatment group interacted with time-specific traits
$D_{it}^{psignal}$	Price signals by peak price and profile (e.g., C_5, C_10)
D_{it}^{price}	Dummies for each peak price level (2–30 NOK/kWh)
$D_{it}^{profile}$	Dummies for price profiles (A, B, C, P, P0)
D_{it}^{temp}	Temperature bins based on 24h average
D_{it}^{expday}	Numbered experiment days within treatment groups
D_{it}^{days}	Days without experiments before each experiment day
D_{it}^{hour}	Hour of the day (1 to 24)
D_{i}^{income}	Income group dummy for household i
$D_{i}^{income} \times D_{it}^{psignal}$	Interaction between income group and price signal (e.g., Lower_A_5, High_A_5)

parts of the results. The absence of income data introduces some limitations. However, there is no indication of systematic bias that would affect the validity of the findings.Table 3.4 shows the distribution of income groups across the sample.

Table 3.4: Survey Responses across different income groups by each region

Region	**Lower (271)**	**Middle (670)**	**High (676)**
Bergen	22	59	53
Bodo	22	101	77
Oslo	118	300	346
Tromso	84	143	141
Trondheim	25	67	59

The survey also collected information on household size, home ownership, number of people residing, electricity contract type, and ownership of energy-related technologies such as electric vehicles and heat pumps. These variables were analyzed together with income group data to assess how demographic and infrastructural factors might influence a household's ability to respond to dynamic electricity pricing. Table 3.5 presents these comparisons.

Despite differences in housing and technology ownership, reported awareness of electricity use was similar across income groups. About 65% of households in each group said they actively monitored their electricity consumption. Lower-income households were slightly more likely to follow electricity prices. Among them, 20.7% reported active price tracking, compared to 18.8% in the middle-income group and 17.6% in the high-income group.

3.3.3. Data Analysis

The main outcome variable is the change in electricity use during high-price events. To measure this, a consumption baseline was calculated for each household. The baseline

Table 3.5: Household and Energy Behavior Characteristics by Income Tier

Variable	**Income Group**		
	Lower	**Middle**	**High**
Average household size (no. of people)	2	3	4
Average home size (m^2)	60–79	120–159	160–199
Homeownership rate (%)	64.90	85.20	98.10
EV ownership rate (%)	6.30	13.90	38.00
Actively follow electricity use (%)	65.30	65.50	65.50
Actively follow electricity price fluctuations (%)	20.70	18.80	17.60

reflected average usage in the same hours on the ten most recent non-event weekdays. This method accounts for regular time-of-day patterns, weekday differences, and changes in outdoor temperature. It is consistent with standard practice in demand response research and supports accurate estimation of the treatment effect.

Table 3.6: Data Analysis Models

Model	**Specification**	
Absolute Model	$E_{it} = \beta_0 + \beta_1 D_{it}^{nonpeak} + \beta_2 D_{it}^{peak} + \beta_3 D_{it}^{notreat} + \gamma_{it} + \delta_t + \varepsilon_{it}$	(1)
Absolute with Interactions	$E_{it} = \beta_0 + \beta_1 D_{it}^{nopeak} + \beta_2 D_{it}^{peak} D_{it}^{x} + \beta_3 D_{it}^{notreat} + \gamma_{it} + \delta_t + \varepsilon_{it}$	(2)
Difference Model	$E_{it} = \beta_0 + \beta_1 D_{it}^{nopeak} + \beta_2 D_{it}^{peak} + \beta_3 D_{it}^{notreat} + \beta D_{it}^{peak} D_{it}^{x} + \gamma_{it} + \delta_t + \varepsilon_{it}$	(3)
Experiment-Day Model	$E_{it} = \beta_0 + \beta_1 D_{it}^{notreat} + \beta_2 D_{it}^{exp} D_{it}^{x} + \gamma_{it} + \delta_t + \varepsilon_{it}$	(4)

The fixed effect model controls for household-level differences in electricity demand. It is widely used in recent pricing studies (Azarova et al., 2020; Burkhardt et al., 2019; Ito et al., 2018). Household fixed effects remove time-invariant traits, while time fixed effects account for shared variation. Fixed effects were specified for each hour of the day within each household. This gives 24 fixed effects per household. It accounts for differences in load profiles, such as heating systems (Azarova et al., 2020). γ_{it} controls for household-specific traits, and δ_t adjusts for factors that change over time resulting in 1776 time fixed effect per region. This includes weather, holidays, or broader trends. The model standard errors are calculated for household cluster.

To compare responsiveness,all models used percentage change:

$$\text{Percentage Change} = \frac{\beta}{E_\beta} \times 100$$

where β is the coefficient estimate, and E_β is average electricity consumption during the period. This makes effects easier to compare across households with different base usage (Ito et al., 2018).

Model (3) estimates group-level differences and allows for confounding variables that would create multicollinearity in Model (2). Model (4) is simplified to use a single dummy variable, D_{it}^{exp}, for all hours on experiment days, instead of separate peak and non-peak hours.

All models included time fixed effects to control for shared variation. Standard errors were clustered at the household level to adjust for serial correlation. This improves reliability in repeated-observation data.

4

Results

This section is dedicated to an independent replication and extension of the core findings reported by Hofmann and Lindberg (2024). By systematically testing their main results on demand reduction, price signal characteristics, and temporal persistence, we aim to achieve two goals: first, to verify that their findings are robust and not dependent on specific coding choices and second, to extend into socioeconomic analysis by examining how household income moderates the responsiveness to dynamic pricing signals.

4.1. Replication Results

4.1.1. Demand Reduction in Peak Hours

The replication begins by testing the original study's most critical finding: the average effect of dynamic pricing on peak-hour electricity consumption. The households in the treatment group reduced their electricity use by 2.92% compared to the control groups during designated peak hours, these results are both statistically significant and identical to the estimate reported by Hofmann and Lindberg (2024).

Table 4.1: Absolute (kWh/h) and percentage change (%) in electricity consumption for peak, non-peak, and non-treatment hours across all price signals.

Variable	Change [kWh/h]		Change [%]		Mean Consumption E_β
	Estimate β	CI 95%	Estimate	CI 95%	[kWh/h]
D_peak	-0.085	[-0.117, -0.052]	-2.92	[-3.99, -1.82]	2.82
D_no_peak	-0.058	[-0.089, -0.027]	-2.16	[-3.28, -1.01]	2.63
D_non_treated	-0.039	[-0.063, -0.014]	-1.58	[-2.55, -0.58]	2.4
R^2	0.797				
N	6,652,896				

***p < 0.001, **p < 0.01, *p < 0.05.
Note :Estimates are calculated with model (1) and represent the average across all hours and price signals.

Table 4.1 presents the detailed regression estimates for the treatment effect across peak, non-peak, and non-treatment hours based on model specification (1).

The successful replication of this headline figure is a crucial first step. It serves two primary purposes: first, it validates the independently developed analytical code, ensuring that the methods are consistent with the original research. Second, it provides strong evidence that the observed demand response is a robust phenomenon, not an result of particular, undocumented implementation choices. This solidifies the 2.92% reduction as a reliable baseline, providing a credible foundation for the subsequent analysis of how this effect is distributed across different household income levels.

Furthermore, the model shows smaller but still significant 2.16% (0.058 kWh/h) reductions in non-peak hours and 1.58% (0.039 kWh/h) reduction on non-treatment days. This suggests the presence of a spillover effect, where households not only shift consumption but may also adopt lasting conservation habits that extend beyond the immediate price signals.

Our average peak-hour reduction of 2.92% is directionally consistent with the broader DR literature. Event-based dynamic tariffs frequently show larger peak cuts when the effective peak–to–off-peak price ratio is high—e.g., the Arcturus 2.0 meta-analysis reports ≈13–16% at 5:1–10:1 (Faruqui et al., 2017). Given our winter, heating-dominant context and rebate framing, an attenuated effect size is plausible. Like the Chicago RTP experiment, our estimates indicate conservation during peak hours rather than clear evidence of compensating off-peak increases (Allcott, 2011; Cappers, 2013).

4.1.2. Demand Reduction by Price Signals

The analysis next investigates whether the demand response is uniform or if it varies with the specific price signal. The replication confirms the original study's finding that households respond more strongly to signals that are higher in price and shorter in duration.

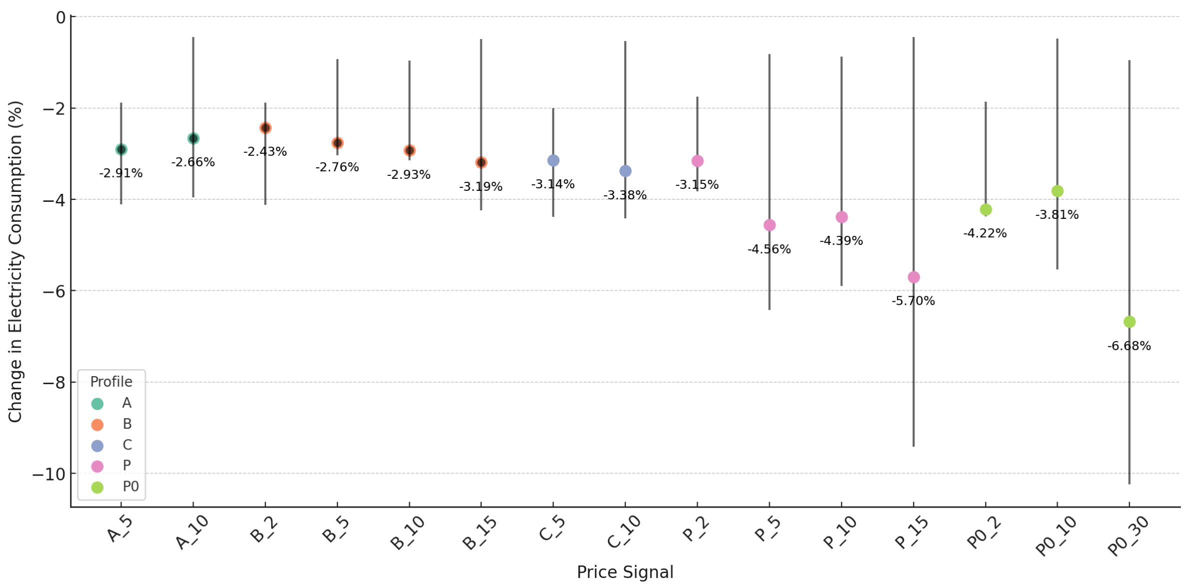

Figure 4.1: Estimated percentage reductions in electricity consumption during peak hours by price signal

Figure 4.1 highlights this relationship. The graph plots the estimated percentage reduction in consumption for each of the unique price signals used in the experiment. A clear pattern emerges where the most intense signals with price greater than 15 NOK/kWh, caused a larger demand reduction than the milder signals.

Table 4.2 provides the detailed regression estimates obtained upon interaction of variable D_{it}^{Psignal} with Model(2). The data confirm that while moderate signals like B_5 with a 5 NOK/kWh peak hour price, yielded a significant 2.76% reduction, the effect increased with price intensity. The P_15 signal with a 15 NOK/kWh peak hour price, led to a 5.70% reduction in demand. The P0_30 signal, produced the largest reduction of 6.68%.

Table 4.2: Absolute and percentage changes in electricity consumption in peak hours for each price signal.

Variable	Change [kWh/h]		Change [%]		Mean Consumption E_β
	Estimate β	CI 95%	Estimate	CI 95%	[kWh/h]
D_no_peak	-0.057***	[-0.089, -0.026]	-2.121***	[-4.27, -1.95]	2.63
D_non_treated	-0.039***	[-0.063, -0.015]	-1.599***	[-3.59, -1.11]	2.40
A_5	-0.082***	[-0.123, -0.041]	-2.906***	[-4.11, -1.81]	2.74
A_10	-0.086***	[-0.14, -0.033]	-2.658***	[-3.96, -0.45]	3.15
B_2	-0.068***	[-0.11, -0.026]	-2.43***	[-4.12, -1.88]	2.73
B_5	-0.081***	[-0.119, -0.043]	-2.764***	[-3.04, -0.93]	2.85
B_10	-0.083***	[-0.119, -0.047]	-2.93***	[-3.15, -0.91]	2.75
B_15	-0.095***	[-0.14, -0.051]	-3.193***	[-4.14, -0.49]	2.88
C_5	-0.083***	[-0.124, -0.042]	-3.14***	[-4.38, -2.01]	2.56
C_10	-0.092***	[-0.136, -0.048]	-3.38***	[-4.34, -0.54]	2.63
P_2	-0.096***	[-0.139, -0.053]	-3.152***	[-3.83, -1.75]	2.95
P_5	-0.154***	[-0.213, -0.095]	-4.564***	[-4.7, -0.82]	3.22
P_10	-0.139***	[-0.183, -0.095]	-4.386***	[-2.87, -0.47]	3.03
P_15	-0.188***	[-0.243, -0.132]	-5.7***	[-1.98, -0.45]	3.11
P0_2	-0.122***	[-0.187, -0.057]	-4.219***	[-4.06, -1.86]	2.77
P0_10	-0.117***	[-0.173, -0.062]	-3.815***	[-2.09, -0.48]	2.95
P0_30	-0.199***	[-0.272, -0.126]	-6.68***	[-3.12, -0.95]	2.78
R^2	0.797				
N	6652896				

***p < 0.001, **p < 0.01, *p < 0.05.
Note :Estimates of the various price signals D_{it}^{Psignal} are calculated with model (2).

The analysis highlights that the price difference between peak and off-peak hours has no influence on response during peak hours. The P0 profiles, which had zero prices during the non-peak hours. It rewarded households only for reduction during the peak hours. Whereas, profile P with 1 NOK/kWh non-peak hour prices, rewarded for reduction across all hours on experiment days.

The price responsiveness reported in meta-studies is consistent with the dose-response we see, which is greater reductions for shorter events with bigger effective differentials (Faruqui et al., 2017). According to theory and previous syntheses, profiles that approximate larger peak–to–off-peak ratios behave more like CPP-style events, whereas profiles that are softer or longer elicit effects that are closer to TOU baselines (Faruqui et al., 2017; Harding & Sexton, 2017).

This suggests that sensitivity to price is a powerful motivator. A high peak hour price clearly encouraged households to shift their consumption, which can be seen as a direct reward for flexibility.

4.1.3. Estimated Effects of Price, Profile, and Temperature

To understand the independent contribution of each primary driver of demand response, the replication analyzes them simultaneously within a single, comprehensive model. The model (3) includes interaction terms for price levels, signal profiles, and temperature bins together. The results confirm that each of these factors remains a significant and distinct driver of household behavior, even when controlling for the others.

Table 4.3 presents the estimated absolute change (in kWh/h) and Table 4.4 shows the percentage change corresponding to each variable D_{it}^{price}, D_{it}^{profile}, D_{it}^{temp} from the output of Model (3).

Table 4.3: Absolute difference of the electricity consumption in peak hours against the respective reference level for various model specifications.

Variable	Coefficient estimate β (kWh/h)					Mean Consumption E_β (kWh/h)
D_non_treated	-0.039**	-0.039**	-0.039**	-0.039**	-0.039**	2.40
D_no_peak	-0.058***	-0.057***	-0.058***	-0.057***	-0.058***	2.63
D_peak	-0.069***	-0.081***	0.016	-0.066***	0.025	2.82
price5	-0.019	—	—	-0.021*	-0.017	2.85
price10	-0.017	—	—	-0.016	-0.013	2.81
price15	-0.106***	—	—	-0.069**	-0.066**	3.10
price30	-0.113***	—	—	-0.090**	-0.089**	2.78
price_profileA	—	-0.004	—	-0.002	-0.002	2.95
price_profileC	—	0.009	—	0.004	-0.006	2.60
price_profileP	—	-0.067**	—	-0.051**	-0.052**	3.00
price_profileP0	—	-0.058*	—	-0.043	-0.048	2.86
temp_category[-20,-15]	—	—	-0.105	—	-0.098	3.10
temp_category[-15,-10]	—	—	-0.102	—	-0.094	3.09
temp_category[-10,-5]	—	—	-0.111**	—	-0.104**	3.10
temp_category[-5,0]	—	—	-0.108**	—	-0.104**	2.85
temp_category[0,5]	—	—	-0.077*	—	-0.074*	2.32
Price	X	–	–	X	X	
Profile	–	X	–	X	X	
Temperature	–	–	X	–	X	

***p < 0.001, **p < 0.01, *p < 0.05.
Note :The estimates are calculated for different versions of model (3), including various combinations of confounding variables of peak price level D_{it}^{price}, price profile D_{it}^{profile}, and outdoor temperature D_{it}^{temp}. The reference levels are for the price level: 2 NOK/kWh (Price 2), for the price profile: Profile B, and for the outdoor temperature: temperature range of 5 to 10 °C (Temperature (5,10]).

The results clearly isolate the powerful effect of high prices. Even when controlling for signal profile and weather, the Price 30 signal is associated with a demand reduction of -0.089 kWh/h relative to the reference price. This demonstrates that the financial incentive from high prices is a primary driver of conservation, independent of other factors.

To contextualize magnitudes, typical residential short-run price elasticities cluster around -0.05 to -0.10, with steeper long-run adjustments as habits and capital stocks change (Labandeira et al., 2017; Zhu et al., 2018). EU panel evidence likewise suggests modest short-run responsiveness and stronger longer-run adaptation in some settings (Jin & Kim, 2022).

Table 4.4: Percentage difference of the electricity consumption in peak hours against the respective reference level for various model specifications.

Variable	Coefficient Estimate (% change)				
D_peak:price5	-0.65			-0.73*	-0.59
D_peak:price10	-0.58			-0.57	-0.46
D_peak:price15	-3.66***			-2.38**	-2.31
D_peak:price30	-3.96***			-3.14**	-3.08**
D_peak:price_profileA		-0.14		0.07	0.07
D_peak:price_profileC		-0.31		-0.15	-0.23
D_peak:price_profileP		-2.31***		-1.74*	-1.68**
D_peak:price_profileP0		-2*		-1.38	-1.38
D_peak:temp_category[-20,-15]			-3.61		-3.38
D_peak:temp_category[-15,-10]			-3.51		-3.24
D_peak:temp_category[-10,-5]			-3.82**		-3.62**
D_peak:temp_category[-5,0]			-3.7**		-3.58**
D_peak:temp_category[0,5]			-2.65*		-2.55*

***p < 0.001, **p < 0.01, *p < 0.05.
Note :The percentage values are calculated based on the results in Table 4.3.

The Profile P reduced consumption by an additional 1.68% (-0.052 kWh/h) compared to the reference Profile B, holding price and temperature constant. This confirms that the certain signal profiles contribute significantly to demand flexibility. Finally, the temperature effects persist. The largest impact is seen in the (-10, -5] °C range, which is associated with a -0.104 kWh/h reduction compared to the mild weather reference bin.

The results aligns closely with the original study's results and confirms its robustness. The analysis shows that price, signal profiles, and weather are not just correlated but are independent, separable drivers of demand response. This provides strong evidence that a multi-faceted approach, considering all factors, is necessary for accurately predicting and shaping household energy behavior.

4.1.4. Impact of external factors: Outdoor temperature

The analysis next examines how the response is influenced by external temperature, a critical factor for a country like Norway with high heating demand. To analyse this effect, the interaction of average outdoor temperature intervals of the last 24 h D_{it}^{temp} with model specification (2) was used. The replication confirms a distinct relationship: household demand reduction is highest during colder weather and diminishes when conditions become either milder or more severe.

Figure 4.2 plots the estimated percentage reduction in peak-hour consumption against different temperatures between -20°c to 5°c, showing that the significant demand reductions are concentrated in the -10°C to 0°C range.

Table 4.5 presents the detailed regression estimates that quantify this pattern. The demand reductions are largest and most statistically significant in the (-10, -5] and (-5, 0] temperature bins, where peak consumption fell by 3.01% and 3.17%, respectively.

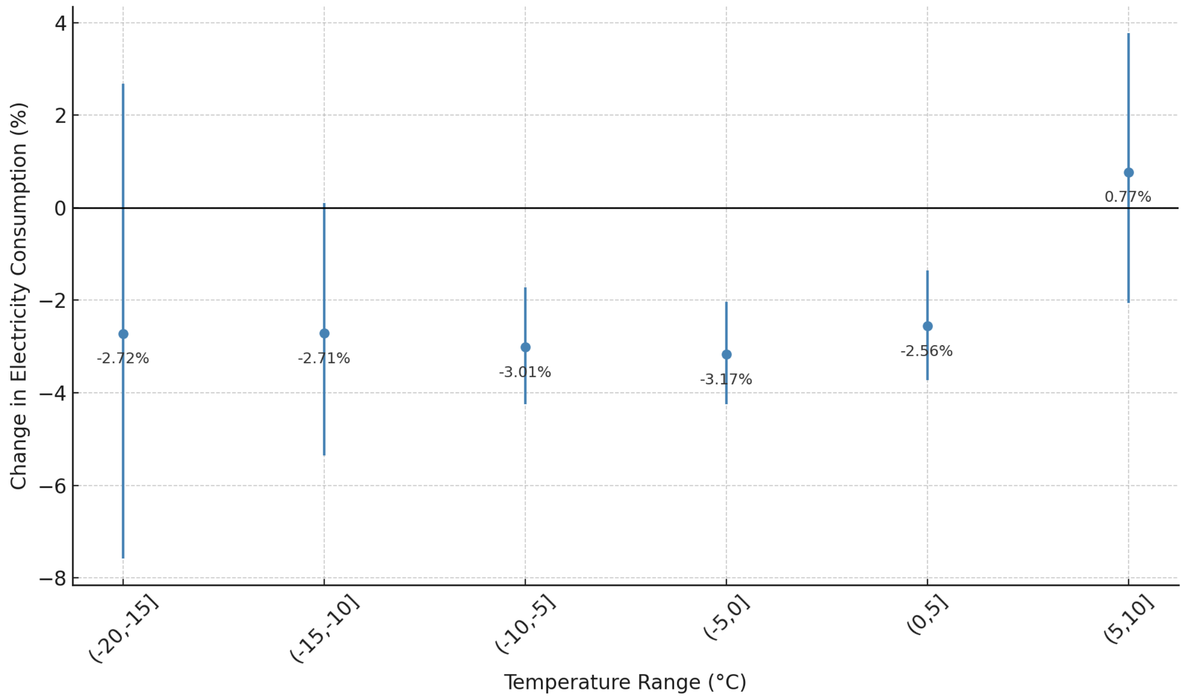

Figure 4.2: Estimated reduction in electricity use during peak hours by outdoor temperature

Table 4.5: Absolute and percentage changes in electricity consumption in peak hours for different outdoor temperature ranges.

Variable	Change [kWh/h]		Change [%]		Mean Consumption E_β
	Estimate β	CI 95%	Estimate	CI 95%	[kWh/h]
D_no_peak	-0.058***	[-0.09, -0.027]	-2.158***	[-3.309, -1.016]	2.63
D_non_treated	-0.039***	[-0.063, -0.014]	-1.599***	[-2.558, -0.679]	2.40
Temp_(-20, -15]	-0.089	[-0.261, 0.083]	-2.72	[-7.582, 2.679]	3.18
Temp_(-15, -10]	-0.086	[-0.175, 0.003]	-2.71	[-5.354, 0.097]	3.09
Temp_(-10, -5]	-0.096***	[-0.137, -0.054]	-3.014***	[-4.247, -1.718]	3.09
Temp_(-5, 0]	-0.093***	[-0.126, -0.059]	-3.169***	[-4.245, -2.034]	2.84
Temp_(0, 5]	-0.061***	[-0.09, -0.032]	-2.56***	[-3.732, -1.36]	2.32
Temp_(5, 10]	0.016	[-0.044, 0.076]	0.77	[-2.065, 3.78]	2.09
R^2	0.797				
N	6652896				

***p < 0.001, **p < 0.01, *p < 0.05.
Note :Estimates of the effect of outdoor temperature D_{it}^{temp} are calculated with model (2).

This finding supports a clear behavioral interpretation linked to heating patterns. In moderately cold weather, household heating is active but still allows for discretionary adjustments, such as lowering a thermostat in response to a price signal. This flexibility is the likely driver of the strong response in this temperature range. In lines with the original study's results, this flexibility is constrained at the extremes. In severe cold (below -10°C), the price response loses statistical significance. This is due the limited observations corresponding to these temperatures. In milder weather between 5°C and 10°C, the effect is also insignificant, most likely because low heating demand leaves little consumption to be shifted in the first place. This confirms that the effectiveness of dynamic pricing is highly dependent on environmental context.Reduction at lower outdoor temperatures is expected in heating-dominant systems: when thermal comfort becomes a constraint, discretionary flexibility tightens and price-only signals are less effective (Harding & Sexton, 2017; Hofmann & Lindberg, 2024). Our pattern is therefore consistent with climate-contingent limits to DR identified in prior work.

4.1.5. Response Fatigue and Cumulative Exposure

A critical question for the viability of dynamic pricing is whether households exhibit response fatigue after repeated exposure to price signals. This section analyzes the persistence of the treatment effect by measuring its magnitude based on the cumulative number of treatment days each household has experienced. The interaction of dummy variable representing consecutive number of experiment days D_{it}^{expday} with the model (2) was used for this analysis.

Figure 4.3 plots the average peak-hour demand reduction across experiment days. The ”Expday 6-10” represents the average effect during a household’s 6th through 10th day of exposure to a price signal. Replication Table 4.6 confirms the original study’s finding. The response was strongest at 3.39% (o.1 kWh/h) reduction on the second to fifth experiment day, the effect remains statistically significant and stable around 2.6% from the eleventh experiment day on wards. The persistence of a significant reduction even after more than 15 days of exposure suggests a combination of response fatigue and behavioral adaptation. This suggests that participants adopted to the price signals and adjusted their consumption habits, reducing electricity use regardless of the number of experiment days.

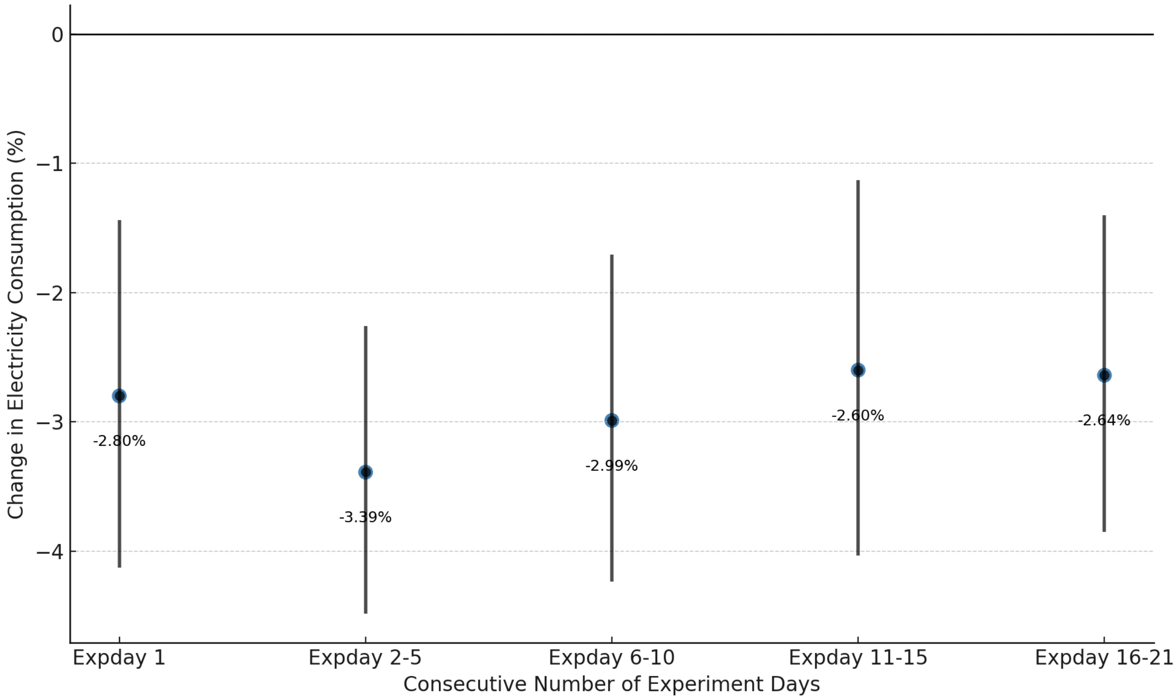

Figure 4.3: Average hourly price response with a 95% confidence interval in dependency on the consecutive number of the experiment days. Model results can be found in Table 4.6.

Table 4.6: Absolute and percentage changes in electricity consumption in peak hours for different consecutive numbers of the experiment day.

Variable	Change [kWh/h]		Change [%]		Mean Consumption E_β
	Estimate β	CI 95%	Estimate	CI 95%	[kWh/h]
D_no_peak	-0.057***	[-0.089, -0.026]	-2.12***	[-3.273, -0.979]	2.63
D_non_treated	-0.039***	[-0.063, -0.014]	-1.6***	[-2.558, -0.58]	2.40
Expday_1	-0.071***	[-0.106, -0.036]	-2.8***	[-4.13, -1.442]	2.46
Expday_2-5	-0.1***	[-0.134, -0.066]	-3.39***	[-4.486, -2.261]	2.85
Expday_6-10	-0.094***	[-0.135, -0.053]	-2.99***	[-4.238, -1.708]	3.05
Expday_11-15	-0.082***	[-0.129, -0.035]	-2.6***	[-4.035, -1.128]	3.07
Expday_16-21	-0.065***	[-0.096, -0.034]	-2.64***	[-3.854, -1.4]	2.39
R^2	0.797				
N	6652896				

***p < 0.001, **p < 0.01, *p < 0.05.
Note :Estimates of the effect of the consecutive number of experiment day D_{it}^{expday} are calculated with model (2).

4.1.6. Behavioral Persistence

In addition to response fatigue, it is important to understand if the learned behavior persists over time or decays when price signals are absent. This section analyzes how the response to a price signal is affected by the number of ”gap days” the time elapsed since the previous experiment day.

Figure 4.4 plots the demand reduction on a treatment day based on how many days have passed since the last experiment, clearly highlighting that the reduction is largest (3.45%) when there has been a gap of 7-19 days.

This finding suggests that spacing out the price signals may prevent habituation and make each new signal more significant. The fact that the response does not weaken and in fact strengthens after a pause of up to 19 days is a powerful finding. It implies that continuous, daily interventions may not be necessary to maintain a state of readiness and responsiveness in households, which is important for designing less intrusive and more sustainable demand response programs.

Apart from this, the Table 4.1 shows that the households reduced their consumption by 1.57% on non experiment days. This reduction shows that these habits are maintained for prolonged durations. The analysis shows that, far from decaying, the effect is actually strongest after a longer pause, suggesting that occasional signals are sufficient to maintain high responsiveness.

Table 4.7 provides the detailed regression estimates. On days immediately following another treatment day (a ”0-day gap”), the demand reduction was 2.65%. This effect increases to 2.95% for gaps of 1-6 days and peaks at 3.45% for gaps of 7-19 days.

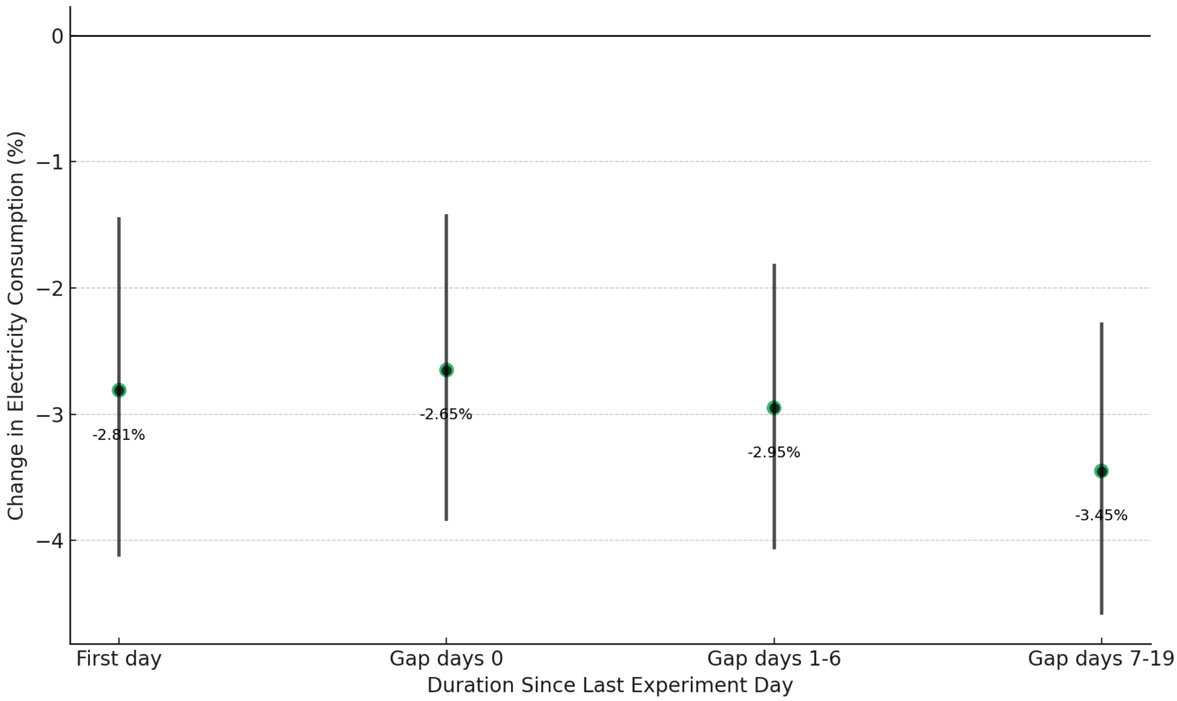

Figure 4.4: Temperature-dependent average hourly price response with a 95% confidence interval. Model results can be found in Table 4.7

Table 4.7: Absolute and percentage changes in electricity consumption in peak hours for different duration since the last experiment day.

Variable	Change [kWh/h]		Change [%]		Mean Consumption E_β
	Estimate β	CI 95%	Estimate	CI 95%	[kWh/h]
D_no_peak	-0.058***	[-0.089, -0.027]	-2.16***	[-3.273, -1.016]	2.63
D_non_treated	-0.039***	[-0.063, -0.014]	-1.58***	[-2.558, -0.58]	2.40
First_day	-0.071***	[-0.106, -0.036]	-2.81***	[-4.13, -1.442]	2.46
Days_0	-0.08***	[-0.117, -0.042]	-2.65***	[-3.845, -1.415]	2.93
Days_1-6	-0.087***	[-0.122, -0.053]	-2.95***	[-4.069, -1.809]	2.88
Days_7-19	-0.092***	[-0.124, -0.061]	-3.45***	[-4.59, -2.275]	2.58
R^2	0.797				
N	6652896				

***p < 0.001, **p < 0.01, *p < 0.05.
Note :Estimates of the effect of the duration since the last experiment day D_{it}^{days} are calculated with model (2).

The small but positive non-experiment-day reduction suggests limited persistence, consistent with evidence that repeated salient cues and incentives can seed modest habit formation (Burkhardt et al., 2019; Ito et al., 2018). At the same time, econometric studies often attribute stronger longer-run adaptation to households with greater ability to adjust capital and routines, a theme we revisit in the heterogeneity analysis (Huang et al., 2024; Jin & Kim, 2022).

4.2. Income as a Moderator of Demand Response

Having established the robustness of the average treatment effect in the replication analysis, the study now examines how these responses vary by household income. This section moves beyond the average effect to test the hypothesis that income acts as a key moderator of demand-side flexibility. The analysis uncovers significant distributional patterns.

4.2.1. Effect of Income on Average Demand Response

The analysis first tests whether the average response to a price signal is dependent on a household's income level. The results show a strong moderating effect. While all income groups respond to dynamic pricing, the magnitude of this response is inversely related to income.

Figure 4.5 shows the average peak-hour demand reduction across six different income brackets included in the survey responses. The visual evidence strongly suggests a moderating relationship, with the treatment effect being largest for the lowest-income households and diminishing as income rises.

Table 4.8 provides the detailed regression estimates for each income bracket. The moderating effect of income is clear and statistically significant. Households earning below 300,000 NOK/year reduced their peak demand by a striking 12.05%. This effect is substantially larger than that of the high income group (1,500,000 NOK or more), which showed a statistically insignificant reduction. Even the second highest income groups show a small reduction of 1.54% (0.55 kWh/h)

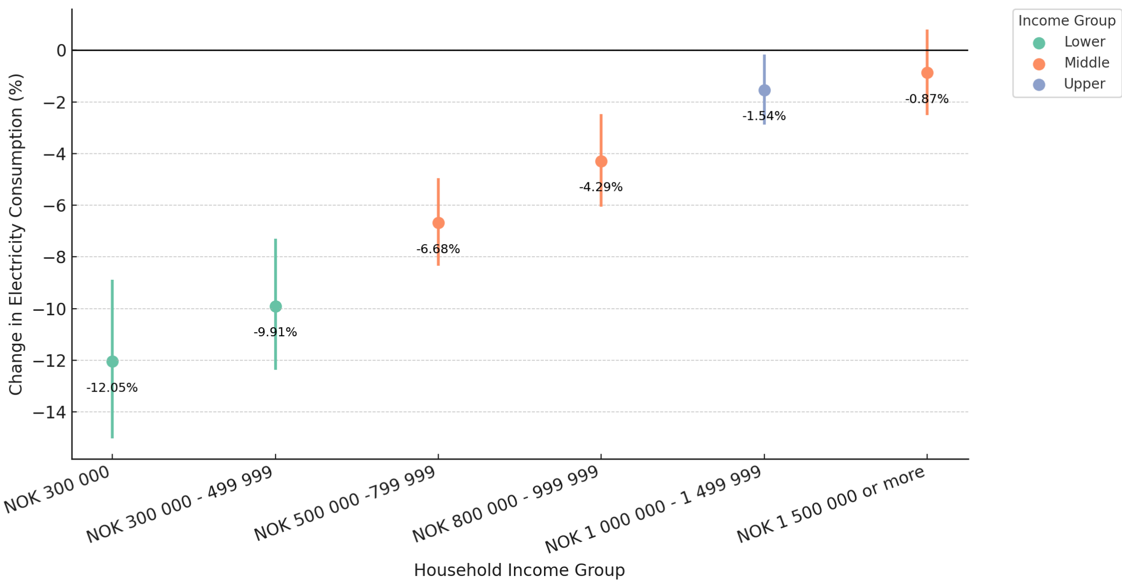

Figure 4.5: The average hourly price response during peak hours across different income groups along with the 95% confidence interval. Model results can be found in Table 4.8

This finding aligns with economic theory, suggesting that households with tighter budget constraints exhibit greater price elasticity (Huang et al., 2024; Jin & Kim, 2022; Reiss & White, 2005) However, it is important to interpret the point estimates with caution. The wide confidence intervals, particularly for the highest income brackets, overlap with those of adjacent groups. While the difference between the lowest and highest-income groups is robust, the data do not allow for strong conclusions about differences between adjacent income categories.

Table 4.8: Absolute and Percentage Change in Peak Hours for Each Income Group

Variable	Change [kWh/h]		Change [%]		Mean Consumption E_β
	Estimate β	CI 95%	Estimate	CI 95%	[kWh/h]
D_no_peak	-0.058***	[-0.089, -0.027]	2.16***	[-3.272, -1.016]	2.63
D_non_treated	-0.038***	[-0.063, -0.014]	-1.58***	[-2.558, -0.58]	2.40
Below_NOK_300_000	-0.221***	[-0.285, -0.157]	-12.05***	[-15.025, -8.076]	1.61
NOK_300,000 - 499,999	-0.213***	[-0.273, -0.152]	-9.91***	[-12.374, -7.29]	1.93
NOK_500,000 - 799,999	-0.164***	[-0.208, -0.12]	-6.69***	[-8.133, -4.91]	2.30
NOK_800,000 - 999,999	-0.124***	[-0.178, -0.07]	-4.29***	[-6.05, -2.47]	2.76
NOK_1,000,000 - 1,499,999	-0.055*	[-0.103, 0.032]	-1.54*	[-2.87, -0.17]	3.52
NOK_1,500,000 or more	-0.065	[-0.103, 0.032]	-0.87	[-2.505, 0.805]	4.01
R^2	0.794				
N	6652896				

***p < 0.001, **p < 0.01, *p < 0.05.
Note :The estimates of the effect of income D_i^{income} are calculated with model (2). This table does not show data for the households that did not respond.

4.2.2. The Interaction of Income and Price Profiles

To further explore how responsiveness varies with signal price the six income brackets were aggregated into three broader categories: 'Low Income' (households earning less than 500,000 NOK), 'Middle Income' (500,000 to 999,999 NOK), and 'High Income' (1,000,000 NOK or more). This aggregation allows for a clearer analysis of the interaction between income level and price profile characteristics.

This study shows that household income plays a major role in determining how people react to sharp increases in electricity prices, especially when prices are high for a short duration. The results clearly show that, on such occasions, low-income households are much more responsive reducing comcumption by 14.55% (0.352 kWh/h) to high-price signals than high-income households which do not significantly reduce their consumption.

Table 4.9 provides the full statistical breakdown of this interaction of these income subgroups and price profiles from Model (2) .It indicate that lower and middle-income households exhibit high price sensitivity, consistently demonstrating statistically significant reductions in peak hour electricity consumption across all price profiles.

The Lower income group consistently sensitive to the high prices,showing statistically significant reductions in peak hour electricity consumption across all detailed price profiles . The reductions were often substantial, with Price Profile P_15 showing the large percentage decrease (-14.11%) and B_15 also demonstrating significant impact (-14.79%). This indicates that lower-income households are highly responsive to various peak pricing strategies, irrespective of the duration. This is likely due to a greater need to manage utility costs. Their mean consumption remained the lowest among all groups.

Similar to the lower-income group, middle-income households showed consistent and statistically significant reductions across all detailed price profiles. Their responsiveness generally fell between that of the high and lower-income groups, confirming their sensitivity to peak pricing. For instance, P0_30 led to a -8.52% reduction and P_15 to a -8.76%

reduction.This highlighting their higher senstivity to high prices for short duration as compared to lower income group with stable reductions across almost all profiles.

The high-income group exhibited a more nuanced response to peak pricing. While most price profiles did not yield statistically significant reductions, Price Profile P0 did. Further analysis with detailed price levels revealed that the effectiveness of P0 for high-income households was primarily concentrated at specific price levels (P_10 and P0_10), which showed statistically significant reductions (-3.50% and -4.58% respectively). However, the highest price level within this profile (P0_30) did not result in a statistically significant change. This suggests that high-income households may be less sensitive to smaller price variations and that there might be an optimal price point beyond which further increases do not lead to additional reductions. Their mean consumption remained the highest.

These results strongly support the notion that peak pricing mechanisms are effective in reducing electricity consumption, with varying degrees of impact across income groups. The detailed analysis of price levels within profiles refines our understanding, particularly for high-income consumers, where effectiveness is concentrated at specific price points. The findings underscore the importance of designing policies carefully calibrated pricing strategies that consider the distinct price sensitivities and response patterns of diverse income demographics to optimize energy conservation efforts during peak hours.

Table 4.9: Changes in Electricity Consumption by Income Category and Price Profile

Income Category	Price Profile	Estimate β	CI 95%	Estimate (%)	CI 95%	Mean Consumption E_β (kWh/h)
Variable						
	D_no_peak	-0.058***	[-0.089, -0.026]	-2.16***	[-3.866, -0.55]	2.63
	D_non_treated	-0.039**	[-0.063, -0.014]	-1.58**	[-2.611, -0.363]	2.40
Lower	A_5	-0.130***	[-0.204, -0.056]	-5.93***	[-10.761, -3.204]	2.06
	A_10	-0.187**	[-0.315, -0.06]	-7.45**	[-14.236, -3.065]	2.33
	B_2	-0.180***	[-0.233, -0.126]	-9.39***	[-10.577, -6.012]	1.74
	B_5	-0.262***	[-0.341, -0.184]	-13.00***	[-14.815, -8.579]	1.76
	B_10	-0.185***	[-0.251, -0.12]	-8.89***	[-12.638, -6.469]	1.90
	B_15	-0.294***	[-0.384, -0.203]	-14.79***	[-17.946, -10.363]	1.69
	C_5	-0.101*	[-0.187, -0.016]	-4.92*	[-7.807, -0.719]	1.96
	C_10	-0.142**	[-0.238, -0.047]	-6.73**	[-10.494, -2.263]	1.97
	P_2	-0.152**	[-0.254, -0.05]	-6.52**	[-11.631, -2.526]	2.18
	P_5	-0.303***	[-0.431, -0.175]	-13.59***	[-18.218, -8.294]	1.93
	P_10	-0.287***	[-0.375, -0.198]	-12.37***	[-14.673, -8.324]	2.03
	P_15	-0.363***	[-0.474, -0.251]	-14.11***	[-19.741, -11.524]	2.21
	P0_2	-0.245**	[-0.392, -0.099]	-11.26**	[-11.43, -3.156]	1.93
	P0_10	-0.238***	[-0.338, -0.137]	-10.97***	[-14.041, -6.21]	1.93
	P0_30	-0.352***	[-0.524, -0.181]	-14.55***	[-15.886, -6.124]	2.07
Middle	A_5	-0.066*	[-0.117, -0.015]	-2.33*	[-3.808, -0.505]	2.77
	A_10	-0.128**	[-0.212, -0.044]	-4.05**	[-7.068, -1.554]	3.04
	B_2	-0.083***	[-0.120, -0.045]	-2.80***	[-4.405, -1.699]	2.87
	B_5	-0.106***	[-0.154, -0.058]	-3.45***	[-5.69, -2.222]	2.96
	B_10	-0.121***	[-0.166, -0.076]	-4.15***	[-5.47, -2.581]	2.79
	B_15	-0.123***	[-0.183, -0.063]	-4.00***	[-5.829, -2.087]	2.96
	C_5	-0.077*	[-0.138, -0.017]	-2.94*	[-5.254, -0.678]	2.55
	C_10	-0.115**	[-0.185, -0.046]	-4.24**	[-6.201, -1.617]	2.60
	P_2	-0.156***	[-0.232, -0.079]	-5.94***	[-6.803, -2.425]	2.46
	P_5	-0.177***	[-0.281, -0.072]	-5.59***	[-8.702, -2.384]	2.99
	P_10	-0.187***	[-0.255, -0.119]	-6.12***	[-9.379, -4.607]	2.80
	P_15	-0.239***	[-0.330, -0.148]	-8.76***	[-9.949, -4.721]	2.49
	P0_2	-0.201***	[-0.312, -0.091]	-6.39***	[-9.261, -2.891]	2.95
	P0_10	-0.198***	[-0.290, -0.106]	-5.88***	[-8.126, -3.132]	3.18
	P0_30	-0.305***	[-0.433, -0.177]	-8.52***	[-14.516, -6.491]	3.28
High	A_5	0.068	[-0.114, 0.249]	1.52	[-2.865, 6.886]	4.52
	A_10	0.139	[-0.081, 0.36]	2.95	[-2.054, 10.276]	4.87
	B_2	-0.020	[-0.103, 0.063]	-0.51	[-2.455, 1.563]	3.85
	B_5	-0.089	[-0.192, 0.015]	-2.28	[-4.215, 0.345]	3.80
	B_10	-0.052	[-0.15, 0.045]	-1.34	[-3.751, 1.183]	3.86
	B_15	-0.052	[-0.177, 0.073]	-1.33	[-4.446, 1.957]	3.87
	C_5	-0.039	[-0.168, 0.09]	-0.88	[-4.402, 2.529]	4.36
	C_10	-0.137	[-0.354, 0.079]	-3.25	[-8.186, 2.03]	4.09
	P_2	-0.071	[-0.214, 0.071]	-1.80	[-4.397, 1.549]	3.88
	P_5	-0.039	[-0.235, 0.157]	-0.93	[-5.348, 3.923]	4.21
	P_10	-0.144*	[-0.276, -0.013]	-3.50*	[-6.641, -0.334]	3.97
	P_15	0.039	[-0.196, 0.275]	1.09	[-4.446, 6.985]	3.65
	P0_2	-0.201	[-0.442, 0.04]	-4.61	[-15.958, 1.748]	4.16
	P0_10	-0.223*	[-0.424, -0.023]	-4.58*	[-8.431, -0.497]	4.65
	P0_30	-0.049	[-0.256, 0.159]	-1.04	[-11.057, 8.367]	4.60
R^2		0.794				
N		6,652,896				

***p < 0.001, **p < 0.01, *p < 0.05.

Note :The estimates of the effect of interaction between D_i^{income} X D_{it}^{Psignal} are calculated with model (2). This table does not show data for the households that did not respond.

5 Discussion

The empirical analysis yields two primary conclusions. The replication exercise first establishes the robustness of the foundational findings in demand response research. Building on this, the study's main contribution is the clear evidence that household response to dynamic pricing is not uniform, but is instead significantly moderated by socioeconomic status.

The successful replication of the main finding from original study, an average peak-hour demand reduction of 2.92%, solidifies confidence in the evidence base for dynamic pricing. In a field where the credibility of empirical results is paramount (McMillan, 2017), this independent verification confirms that the effects are a robust and reliable and not a part of the original study's specific analytical choices. This validation provides a solid foundation from which to explore more nuanced questions.

The central contribution of this study, is the clear evidence that household income is a powerful moderator of demand response. The analysis reveals a strong, inverse relationship between income and price responsiveness. This effect is most pronounced among the lowest-income households, which reduced their peak-hour consumption drastically. This response is more than four times the sample average. Conversely, the demand reduction from the highest-income households was not statistically significant, indicating that the aggregate effect of dynamic pricing is driven disproportionately by the actions of the most financially constrained consumers.

This result provides compelling evidence for a core principle of economic theory: households facing tighter budget constraints exhibit greater price elasticity. For a low-income family, a sudden spike in electricity prices represents a significant financial consideration, creating a strong incentive to mitigate the cost through behavioral change.

This finding directly addresses the empirical tension identified in the literature review. While some research posits that low-income households may not be flexible due to a lack of enabling technologies, the results here suggest this factor is secondary. The dominant form of flexibility in this experiment appears to be behavioral rather than technological. This interpretation is supported by the demographic data, which shows that while electric vehicle ownership is skewed towards high-income households, the capacity to manually adjust heating is near-universal. The availability of this low-tech and accessible means of response appears to be the primary channel through which low-income households exercise their demand flexibility.

5.1. Limitations

While this study provides critical insights into the distributional effects of dynamic pricing, its findings should be understood within the context of its specific methodological and contextual boundaries. These limitations do not undermine the validity of the results but rather define their scope. Acknowledging these boundaries is essential for a nuanced interpretation and serves to highlight several important future research options.

The primary limitations of this research can be grouped into two categories: those related to the context and generalizability of the experiment, and those inherent in its methodological design of the experiment.

First, the study's context is highly specific. The experiment was conducted in Norway, a country with a unique energy profile characterized by near-universal electrification of heating, high consumer awareness of energy issues, and a cold winter climate. Consequently, the findings may not be directly generalizable to regions with different climatic conditions, energy mixes, or consumer behaviors. For instance, in a warmer climate where peak demand is driven by air conditioning, the patterns of household flexibility could be substantially different. The focus on the winter season also means these results may not reflect household responsiveness during other times of the year.

Second, the experiment methodology has inherent constraints. The study relied on a sample of households that volunteered to participate, which introduces a potential for self-selection bias. These participants may be more engaged with energy consumption or more motivated to respond to incentives than the general population, potentially inflating the magnitude of the observed effects. Furthermore, while the income analysis yielded significant results, it was based on survey data from the iFlex experiment provided by only 43% of the sample. Although no systematic bias was detected between respondents and non-respondents, the incomplete nature of this data warrants caution when drawing firm conclusions about income-related heterogeneity. Finally, the experiment utilized a rebate-based incentive structure rather than exposing households to actual high prices on their bills(Hofmann & Lindberg, 2024). This design, while effective for recruitment, may not fully replicate the psychological impact of facing a direct financial penalty, which could influence behavior in different ways.

5.2. Directions for Future Research

The limitations outlined above give rise to a clear agenda for future research. Each boundary of this study represents an opportunity to build a more comprehensive and universally applicable understanding of demand response.

The most critical step is to address the question of generalizability. Future work should aim to replicate this study's methodology in diverse geographical and climatic contexts. Conducting similar experiments in warmer regions, in countries with different housing stocks, or in markets with less pre-existing consumer engagement would be invaluable for testing whether the observed income-based heterogeneity is a universal phenomenon or a context-specific one. Additionally, longitudinal studies that track household behavior over multiple years and seasons are needed to understand the long-term persistence of these responses and to investigate whether effects like "response fatigue," a known concern in behavioral interventions, eventually emerge.

Future research should also seek to deepen the analysis of household heterogeneity. While this study focused on income, a household's capacity and willingness to be flexible is likely influenced by a host of other factors. Researchers could expand the analytical model to include variables such as education level, housing tenure (renting vs. owning), household composition, and psychographic factors like environmental attitudes or risk aversion. This would allow for the development of more sophisticated behavioral models that move beyond simple economic indicators.

Finally, the findings of this thesis call for research into alternative and more equitable policy designs. Rather than treating all consumers as a monolith, future experiments should test the effectiveness of targeted interventions. For example, a study could directly compare a price-only signal with a technology-support model for high-income households to empirically measure the most effective way to unlock their flexibility. Furthermore, the increasing sophistication of smart home technology presents an opportunity to explore the potential of AI-driven automation. Future research could investigate how personalized, AI-powered systems that learn a household's unique consumption patterns and preferences can optimize energy use without requiring active user intervention. Such systems could prove particularly effective for engaging otherwise unresponsive households and could be designed with equity in mind, ensuring that automation benefits all segments of the population. By pursuing these avenues, the research community can help policymakers design programs that are not only effective but also fair, ensuring a just transition for all.

6
Conclusion

This investigation was undertaken to achieve two primary objectives. The first was to conduct a rigorous, independent replication of the core findings reported by Hofmann and Lindberg (2024) concerning household demand response to variable electricity prices. The second was to extend this analysis to test the hypothesis that household income acts as a significant moderator of these behavioral responses.

The replication analysis successfully reproduced the original study's main finding. An average peak-hour demand reduction of 2.92% was estimated, a result that is statistically significant and identical to the previously reported effect. Along with other findings this outcome validates the robustness of the original findings and confirms the analytical framework used for the subsequent extension.

The primary contribution of this thesis comes from the extension, which examined the heterogeneity of treatment effects across income groups. The analysis uncovered a strong, statistically significant, and somewhat inverse relationship between household income and price-induced demand reduction. The effect was most pronounced in the lowest-income group, which exhibited the highest peak-hour consumption decrease, upto 15% for certain price profiles. In contrast, the demand response of the highest-income households was not statistically significant for majority of price profiles.

These findings lead to two main conclusions. First, the replication confirms that dynamic pricing effectively reduces household electricity consumption during peak hours in the Norwegian experimental context. Second, this effect is not uniform across income groups. It is disproportionately driven by lower-income households, highlighting important distributional differences in price responsiveness. The aggregate demand reduction observed in the experiment is disproportionately driven by the price sensitivity of the most financially constrained households.

The practical implication of this conclusion is that uniform dynamic pricing policies for all socioeconomic groups, are likely to produce regressive outcomes and may be inefficient at capturing the full demand-side flexibility potential, particularly from high-income segments. Therefore, future policy design and program implementation should incorporate this observed heterogeneity to create more effective and equitable energy systems.

6.1. Practical and Policy Implications

The discovery of such profound heterogeneity in demand response has significant practical implications for the design and implementation of energy policy. The findings serve as a critical warning against the adoption of ”one-size-fits-all” dynamic pricing schemes, which appear to be the default approach in many jurisdictions.

For policymakers, the most pressing implication is the inherent risk of regressive outcomes. A uniform dynamic pricing policy, when applied to a population with varied price sensitivity, will disproportionately place the burden of grid stability on the households least able to afford it. The success of the program in this study was driven almost entirely by the significant actions of low- and middle-income households, while high-income households were largely unresponsive. To ensure a just energy transition, policymakers must therefore move beyond averages and design frameworks that account for this heterogeneity, potentially through tiered pricing structures to financial safeguard low-income households.

From a practical design perspective, these findings suggest that a uniform strategy is not only inequitable but also inefficient, as it fails to unlock the vast and largely untapped demand flexibility potential of high-income households. Their lack of response should not be interpreted as an unwillingness to contribute, but rather as an insensitivity to the price signal alone, for which the financial savings are likely too marginal to warrant conscious effort. This points to a clear opportunity for a more targeted, dual-pronged program design. An approach that successfully promotes the flexibility across population segments becomes more urgent. For low and middle-income households, dynamic price signals appear to be the key motivator of behavioral change. This suggests that programs should focus on delivering simple and salient price-based signals to effectively engage these groups. For high-income customers, however, the focus should shift from price to convenience and automation. By offering enabling technology like subsidized smart thermostats or automated EV charging schedulers, utilities can make participation frictionless and promote demand flexibility from the very group that a price-only signal fails to engage. By adopting such a segmented approach, the energy sector can design demand response programs that are not only more effective and efficient but also fundamentally more equitable.

A

Appendix

A.1. Disclaimer

This report benefited from the use of AI-based tools to support language editing and clarity. These tools were used to improve readability and consistency without altering the substance of the analysis.

In addition, ”Writing Empirical Research Report” by Galvan and Pyrczak was used to guide the structure and style of academic argumentation of the report.

References

Allcott, H. (2011). Rethinking real-time electricity pricing. *Resource and Energy Economics*, *33*(4), 820–842. https://doi.org/10.1016/j.reseneeco.2011.06.003

Ankel-Peters, J., Brodeur, A., Dreber, A., Johannesson, M., Neubauer, F., & Rose, J. (2025). A protocol for structured robustness reproductions and replicability assessments. *Q Open*, *5*(3), qoaf004. https://doi.org/10.1093/qopen/qoaf004

Azarova, V., Cohen, J. J., Kollmann, A., & Reichl, J. (2020). Reducing household electricity consumption during evening peak demand times: Evidence from a field experiment. *Energy Policy*, *144*, 111657. https://doi.org/10.1016/j.enpol.2020.111657

Bergh, D. D., Sharp, B. M., Aguinis, H., & Li, M. (2017). Is there a credibility crisis in strategic management research? evidence on the reproducibility of study findings. *Strategic Organization*, *15*(3), 423–436. https://doi.org/10.1177/1476127017701076

Block, J. H., Fisch, C., Kanwal, N., Lorenzen, S., & Schulze, A. (2023). Replication studies in top management journals: An empirical investigation of prevalence, types, outcomes, and impact. *Management Review Quarterly*, *73*(3), 1109–1134. https://doi.org/10.1007/s11301-022-00269-6

Bollinger, B. K., & Hartmann, W. R. (2017). Information vs automation and implications for dynamic pricing. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2543149

Brewer, D. (2023). Household responses to winter heating costs: Implications for energy pricing policies and demand-side alternatives. *Energy Policy*, *177*, 113550. https://doi.org/10.1016/j.enpol.2023.113550

Buckley, P. (2020). Prices, information and nudges for residential electricity conservation: A meta-analysis. *Ecological Economics*, *172*, 106635. https://doi.org/10.1016/j.ecolecon.2020.106635

Burkhardt, J., Gillingham, K., & Kopalle, P. (2019, February). *Experimental evidence on the effect of information and pricing on residential electricity consumption* (w25576). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w25576

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*(6280), 1433–1436. https://doi.org/10.1126/science.aaf0918

Cappers, P. (2013). Quantifying the impacts of timebased rates, enabling technology, and other treatments in consumer behavior studies: Protocols and guidelines. Retrieved August 3, 2025, from https://escholarship.org/uc/item/9wr4t18w

Chang, A. C., & Li, P. (2015). Is economics research replicable? sixty published papers from thirteen journals say "usually not". Retrieved August 3, 2025, from https://www.federalreserve.gov/econres/feds/is-economics-research-replicable-sixty-published-papers-from-thirteen-journals-say-quotusually-notquot.htm

Coffman, L. C., Niederle, M., & Wilson, A. J. (2017). A proposal to organize and promote replications. *American Economic Review*, *107*(5), 41–45. https://doi.org/10.1257/aer.p20171122

Colliard, J.-E., Hurlin, C., & Pérignon, C. (2022). The economics of research reproducibility. https://doi.org/10.2139/ssrn.3418896

Dunning, T. (2016). Transparency, replication, and cumulative learning: What experiments alone cannot achieve. *Annual Review of Political Science*, *19*(1), 541–563. https://doi.org/10.1146/annurev-polisci-042114-015939

Duvendack, M., Palmer-Jones, R., & Reed, W. R. (2017). What is meant by “replication” and why does it encounter resistance in economics? *American Economic Review*, *107*(5), 46–51. https://doi.org/10.1257/aer.p20171031

Empowering vulnerable consumers in the energy transition - european commission. (n.d.). Retrieved August 3, 2025, from https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/empowering-vulnerable-consumers-energy-transition-2022-06-30_en

Faruqui, A., & Sergici, S. (2013). Arcturus: International evidence on dynamic pricing. *The Electricity Journal*, *26*(7), 55–65. https://doi.org/10.1016/j.tej.2013.07.007

Faruqui, A., Sergici, S., & Warner, C. (2017). Arcturus 2.0 : A meta-analysis of time-varying rates for electricity. *The Electricity Journal*, *30*(10), 64–72. https://doi.org/10.1016/j.tej.2017.11.003

Frederiks, E. R., Stenner, K., Hobman, E. V., & Fischle, M. (2016). Evaluating energy behavior change programs using randomized controlled trials: Best practice guidelines for policymakers. *Energy Research & Social Science*, *22*, 147–164. https://doi.org/10.1016/j.erss.2016.08.020

Hao, C. H., Wesseh, P. K., Wang, J., Abudu, H., Dogah, K. E., Okorie, D. I., & Osei Opoku, E. E. (2024). Dynamic pricing in consumer-centric electricity markets: A systematic review and thematic analysis. *Energy Strategy Reviews*, *52*, 101349. https://doi.org/10.1016/j.esr.2024.101349

Harding, M., & Sexton, S. (2017). Household response to time-varying electricity prices. *Annual Review of Resource Economics*, *9*(1), 337–359. https://doi.org/10.1146/annurev-resource-100516-053437

Hensel, P. G. (2021). Reproducibility and replicability crisis: How management compares to psychology and economics – a systematic review of literature. *European Management Journal*, *39*(5), 577–594. https://doi.org/10.1016/j.emj.2021.01.002

Hofmann, M., & Lindberg, K. B. (2021). Do households react to variable power prices? - results from a norwegian pricing experiment. *2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, 1–6. https://doi.org/10.1109/ISGTEurope52324.2021.9639976

Hofmann, M., & Lindberg, K. B. (2024). Evidence of households’ demand flexibility in response to variable hourly electricity prices – results from a comprehensive field experiment in norway. *Energy Policy*, *184*, 113821. https://doi.org/10.1016/j.enpol.2023.113821

Hofmann, M., & Siebenbrunner, T. (2023). A rich dataset of hourly residential electricity consumption data and survey answers from the iFlex dynamic pricing experiment. *Data in Brief*, *50*, 109571. https://doi.org/10.1016/j.dib.2023.109571

Huang, W., Li, H., & Li, Z. (2024). A comprehensive study to estimate income and price elasticities of household electricity consumption using auto-metrics. *Heliyon*, *10*(8), e28656. https://doi.org/10.1016/j.heliyon.2024.e28656

Huebner, G., Nicolson, M., Fell, M., Kennard, H., Elam, S., Hanmer, C., Johnson, C., & Shipworth, D. (2017). Are we heading towards a replicability crisis in energy efficiency research? a toolkit for improving the quality, transparency and replicability of energy efficiency impact evaluations.

Hwang, J., Kim, H., Caesary, D., & Eom, J. (2025). Data-driven assessment of the varied effects of the peak time rebate on household electricity usage. *Applied Energy*, *396*, 126329. https://doi.org/10.1016/j.apenergy.2025.126329

Ito, K., Ida, T., & Tanaka, M. (2018). Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy*, *10*(1), 240–267. https://doi.org/10.1257/pol.20160093

Jin, T., & Kim, J. (2022). The elasticity of residential electricity demand and the rebound effect in 18 european union countries. *Energy Sources, Part B: Economics, Planning, and Policy*, *17*(1), 2053896. https://doi.org/10.1080/15567249.2022.2053896

Kim, J.-H., & Shcherbakova, A. (2011). Common failures of demand response. *Energy*, *36*(2), 873–880. https://doi.org/10.1016/j.energy.2010.12.027

Labandeira, X., Labeaga, J. M., & López-Otero, X. (2017). A meta-analysis on the price elasticity of energy demand. *Energy Policy*, *102*, 549–568. https://doi.org/10.1016/j.enpol.2017.01.002

Lijnis Huffenreuter, R. (2013). The economic effects of energy efficiency. Retrieved August 3, 2025, from https://thesis.eur.nl/pub/14780

Martinez-Reyes, A., Wolters, S., Luijk, O., Okur, Ö., & Hoppe, T. (2025). Who is vulnerable in regional energy transitions? an intersectional energy justice analysis of the rotterdam-the hague region. *Energy Research & Social Science*, *119*, 103859. https://doi.org/10.1016/j.erss.2024.103859

McMillan, D. (2017). Replication studies. *Cogent Economics & Finance*, *5*(1), 1410940. https://doi.org/10.1080/23322039.2017.1410940

Newsham, G. R., & Bowker, B. G. (2010). The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review. *Energy Policy*, *38*(7), 3289–3296. https://doi.org/10.1016/j.enpol.2010.01.027

Pellini, E. (2021). Estimating income and price elasticities of residential electricity demand with autometrics. *Energy Economics*, *101*, 105411. https://doi.org/10.1016/j.eneco.2021.105411

Reiss, P. C., & White, M. W. (2005). Household electricity demand, revisited. *The Review of Economic Studies*, *72*(3), 853–883. Retrieved August 3, 2025, from http://www.jstor.org/stable/3700676

Simionescu, M., & Cifuentes-Faura, J. (2024). Evaluating the relationship between income inequality, renewable energy and energy poverty in the v4 countries. *Energy Research & Social Science*, *115*, 103640. https://doi.org/10.1016/j.erss.2024.103640

Uhr, D. d. A. P., & JÃºlia Gallego Ziero Uhr, A. L. S. C. (2017). Estimation of price and income elasticities for the brazilian household electricity demand. *Working Papers, Department of Economics*. Retrieved August 14, 2025, from https://ideas.repec.org//p/spa/wpaper/2017wpecon12.html

Zhu, X., Li, L., Zhou, K., Zhang, X., & Yang, S. (2018). A meta-analysis on the price elasticity and income elasticity of residential electricity demand. *Journal of Cleaner Production*, *201*, 169–177. https://doi.org/10.1016/j.jclepro.2018.08.027