Essays in Environmental Economics and International Trade

by

Alastair Edward Wilson Fraser

B.Sc., The University of Victoria, 2007
M.Sc., The University of Alberta, 2010
M.A., Queen's University, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Essays in Environmental Economics and International Trade

submitted by Alastair Edward Wilson Fraser in partial fulfillment of the requirements for

the degree of Doctorate of Philosophy

in Economics

Examining Committee:

Prof Brian Copeland, Vancouver School of Economics
Co-supervisor

Prof Werner Antweiler, Sauder School of Business
Co-supervisor

Prof Marit Rehavi, Vancouver School of Economics
University Examiner

Prof Richard Barichello, Land and Food Systems
University Examiner

Prof Nicholas Rivers, University of Ottawa
External Examiner

Additional Supervisory Committee Members:

Prof Carol McAusland, Land and Food Systems
Supervisory Committee Member

Prof Tomasz Swiecki, Vancouver School of Economics
Supervisory Committee Member

Abstract

In Chapter 1, I study how households respond to financial rewards offered for achieving electricity conservation targets. Using an event-study empirical approach, I estimate the short-run and long-run changes in electricity use. I find that electricity use declines as households join the program and attempt to achieve their conservation targets, but rebounds close to pre-program levels as households leave the program. This suggests that households do not make changes that result in persistent electricity conservation, and that the ongoing incentive of the financial rewards is necessary for causing long-run lower electricity use.

In Chapter 2, I exploit a discontinuity in the probability that households re-enroll in the same energy conservation program. This provides direct evidence on what determines households' re-enrollment decisions, and permits me to use a fuzzy regression discontinuity empirical strategy to estimate the treatment effect of re-enrolling. I find households' decisions whether to re-enroll are sensitive to their success or failure in achieving their conservation target but, conditional on their success, are largely independent of the level of conservation they achieve or their pre-determined characteristics. As a result, households do not make re-enrollment decisions that are consistent with the incentive structure of the reward program. Importantly for many incentive programs, this suggests households are using simple heuristics in making decisions rather than responding to the detailed information provided to them.

In Chapter 3, I show that trade models incorporating multiple transport modes have imposed strong—and potentially unrealistic—restrictions on substitution patterns across mode-specific trade flows. In particular, I show that different models have implicitly assumed bilateral trade by air and sea to be both complements and substitutes, and this assumption has significant quantitative and qualitative implications for counterfactual trade patterns. Using freight costs for U.S. imports I estimate that the elasticity of substitution between modes. I find no evidence that transport modes are substitutes and limited evidence they are complements.

Lay Summary

Using energy efficiently is an important component of reducing carbon emissions. This dissertation focuses on two aspects of this challenge. I first consider how households conserve energy in an electricity conservation program that offers them financial rewards. I find that while such rewards cause households to reduce their electricity use, their electricity use rebounds as they leave the program. In addition, households are found to rely on simple rules in making decisions whether to participate - rather than respond as is typically expected from economic theory. The dissertation then shows how prior models of international trade impose unrealistic restrictions on the choice of air vs. ocean transportation, and that there is little potential for substitution between them. This suggests carbon emissions from international aviation and ocean trade can be separately regulated without unintended consequences due to interactions between them.

Preface

This dissertation is original, unpublished, independent work by the author, Alastair Fraser.
The research in Chapters 1 and 2 was approved by the UBC Behavioural Research Ethics Board under the project title *Household Energy Consumption* with Certificate Number H14-03073.

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgements

A dissertation is never an individual undertaking, but the product of many supportive and thoughtful voices. I am deeply grateful to Werner Antweiler, Brian Copeland, Carol McAusland, and Tomasz Swiecki for their advice, encouragement, and patience at all stages of this thesis and throughout the PhD. They have been generous with their time and continue to inspire me to be a better economist and researcher. I am particularly indebted to Joshua Gottlieb, Thomas Lemeiux, Vadim Marmer, and Kevin Milligan who, despite not being on my thesis committee, were willing to frequently provide immensely helpful advice on the details and big picture view of my work, and to Patrick Baylis, Nicole Fortin, David Green, and Sumeet Gulati for insightful research and career advice. I also recognize the assistance of the department staff and in particular Maureen Chin, who has ever-cheerfully shepherded me through the graduate school process and into meeting deadlines I would have otherwise missed. I would like to single out the regular attendees of the environmental reading group, Richard Barichello and James Vercammen, and the trade group, Vanessa Alviarez, Matilde Bombardini, Keith Head, Hiro Kasahara, and John Ries, for listening to many iterations of this work and providing helpful suggestions at all stages. Too many other professors of the UBC community to name have provided advice in seminars, the hallways, and discussions in their offices: thank you. Beyond UBC, Geoff Steeves, Mark Freeman, and Andrew Leach have been instrumental in inspiring this pursuit of cautious knowledge. They have provided not only mentorship in research, but also deeply appreciated guidance in those important questions of life and what to do with it. Geoff Steeves taught me my first steps of physics research and then provided much advice over the years since. Mark Freeman, whose lessons in what careful thinking and research should and can be are ones I continue to use regularly. I will be forever grateful to have his counsel and support in thinking through my difficult decision to leave physics. And Andrew Leach, for his willingness to assist my curious but largely uninformed first steps into economics as well as in providing an example of how economics can be an applied discipline focused on pressing problems. I also count myself extremely lucky to have had the support of such close and talented friends in my small cohort: Joao da Fonseca, Brad Hackinen, Nouri Najjar, Jose Pescador, and Iain Snoddy. And beyond it: Tom Cornwall, Alex Hemingway, Jeff Hicks, Neil Lloyd, Timi Molnar, and Gaelle Simard-Duplain. Finally and most importantly, to my parents Noni and Stirling, and my brother David, for their love and guidance. Thank you.

To my family.

Introduction

Climate change is among the central challenges facing society. Important in responding to this challenge is improving the effective use of energy. This motivates the two topics studied in this dissertation: how households respond to financial rewards offered for electricity conservation, and the choice of transportation method within international trade.

Electricity and heat production is the single largest source of carbon emissions, within which the largest sector is buildings (IPCC, 2014). This has spurred efforts to conserve energy use in buildings through a wide variety of incentive programs, efforts to inform consumers, and direct regulations. Governments and utility companies have many reasons beyond carbon emissions for desiring changes in the use of energy, including other market failures of information asymmetry, transaction costs, principle agent problems, other pollution externalities, and the political costs of raising the price of energy. This dissertation evaluates an energy efficiency program to learn whether and how it is effective, and for the insights into how people respond to the general incentives that it provides.

In Chapter 1 and 2, I use a confidential dataset of households' electricity use to study how consumers respond to financial rewards offered for achieving electricity conservation targets. Households in this program repeatedly choose whether to attempt annual conservation targets; this allows me to track households' extensive margin participation decisions and their intensive margin electricity conservation efforts. This feature provides several insights into how households respond to the financial rewards and use information in making decisions, while the potential for self-selection makes it a challenge to attribute causal treatment effects. To address these challenges, I employ two separate and complementary empirical strategies. In Chapter 1, I use an event study empirical approach to estimate the short-run and long-run changes in electricity use from households participating in the program. Comparing the pattern of estimated reductions over time and across different households provides insights into their self-selection decisions and the changes made within the home to conserve energy. In Chapter 2, I use the same dataset to examine a different margin of households' responses—the extensive margin decision whether to re-enroll in the energy conservation program. By exploiting a discontinuity in the probability that households re-enroll in the program, I provide direct evidence on households' re-enrollment decisions. I then use a fuzzy regression discontinuity empirical strategy to estimate the treatment effect from re-enrolling. This approach complements the event study estimates from Chapter 1, and consistent with them, finds that an additional conservation challenge causes lower electricity use.

These chapters make two separate and important contributions. First, I find that electricity use declines as households join the program and attempt to achieve their successive conservation targets, but rebounds close to pre-program levels as they leave the program. This suggests that households do not make changes, such as physical investments or lasting habits, that result in persistent electricity conservation, and that the ongoing incentive of the financial rewards is necessary for causing long-run lower electricity use. Second, I find households' decisions whether to re-enroll are sensitive to their success or failure in achieving their conservation target but, conditional on their success, are largely independent of the level of conservation they achieve. As a result, households do not make re-enrollment decisions that are consistent with the incentive structure of the reward program. Importantly for many incentive programs, this suggests households are using simple heuristics in making decisions rather than incorporating the detailed information that is provided to them.

Chapter 3 turns to different topic: the choice of transportation mode in international trade. Transportation modes differ substantially in freight cost, delivery time, their use across countries and industries, and environmental impacts. With air freight releasing approximately one hundred times the carbon emissions per tonne-km as ocean freight, the choice of transport mode is particularly important to the direct effects of international trade on climate change. These differences have been considered through a variety of trade models that incorporate multiple transportation modes, as well as exploited for identification of trade elasticities, the value of delivery time, and the importance of distance to trade. This chapter first shows how several tractable models of trade start from different theoretical motivations yet impose the same reduced form predictions for counterfactual trade patterns. This common framework imposes strong, and potentially unrealistic, restrictions on substitution patterns of trade across modes and countries. In particular, I find models have implicitly treated bilateral trade by air and sea as both substitutes and complements, and that this has significant quantitative and qualitative implications for counterfactual trade patterns. To evaluate whether these models accurately approximate real trade flows, this chapter then asks whether the empirical evidence on mode choice is consistent with air and ocean transport being complements or substitutes. By exploiting idiosyncratic variations in freight rates, I estimate the degree of substitution across transport modes within bilateral trade. I find little evidence for any substitution between transport modes for products primarily carried by air, and some evidence that imports typically carried by ocean transport are complements with air transport. I then revisit the observation—used to motivate the potential for substitution between modes—that many products arrive by both transport modes. I find that this fact can be explained in part by heterogeneous product quality within detailed product categories. In addition, this unobserved quality is highly correlated with the choice of transport mode across countries, suggesting that a countries specialization in high vs. low quality products may be a major and largely unrecognized determinant of the choice of mode.

Chapter 1

The Intensive Margin of Electricity Conservation

1.1 Introduction

Governments and electrical utility companies use a wide variety of incentive and information programs to reduce electricity use and improve energy efficiency. The success of these programs has been mixed. Programs can underperform when consumers do not respond to price schedules and information as predicted. Their cost-effectiveness can be reduced when consumers are rewarded for changes that would have occurred in the absence of the program. How consumers use heuristics and deviate from neoclassical models of consumer choice is still poorly understood, and this lack of understanding leads to ad-hoc program designs. Even when energy is conserved or efficiency improved the long-run persistence of changes is often uncertain but important for program effectiveness. In addition, many incentive programs are voluntary. This leaves them vulnerable to selection biases that can make it difficult to evaluate their cost-effectiveness. As a result of these limitations, the widespread deployment of energy efficiency and conservation programs risks generating few and expensive reductions in energy use. These are challenges beyond energy use; ensuring that programs cost-effectively deliver anticipated results requires improved models of how incentives and information affect decisions and the careful evaluation of existing programs.

In this chapter, I analyze a novel program that incentivizes energy conservation through repeated financial rewards. BC Hydro's Team Power Smart program offers households the opportunity to undertake annual electricity conservation "challenges." Households that successfully reduce their annual electricity use by 10% relative to their use in the previous 12 months receive a $75 financial reward. This reward is worth 10% of the average household's annual electricity bill. After each annual conservation challenge, all households have the option of participating in an additional conservation challenge. Importantly, subsequent conservation challenges require an additional 10% reduction in electricity use relative to the previous year and are available to households regardless of their past challenge success or failure.

Using an event study empirical strategy, I estimate the short-run reductions and long-run persistence of changes in electricity use associated with participation in the Team Power Smart program. This provides substantial insight into households' intensive margin effort and extensive margin participation decisions. I find that an initial conservation challenge is associated with an immediate 4.3% average reduction in electricity use which lasts throughout the twelve months of the first challenge. Comparing the electricity conservation across households I show that there is selection into subsequent challenges based on the level of reductions in electricity use achieved by households. In addition, electricity use continues to decline among households that re-enroll and rebounds close to pre-program levels as households end their participation in the program. This rebound occurs primarily in the months leading up to a household's exit from the program and suggests conservation effort stops well before households officially end their participation. This shows that these households tend to make only short-run adjustments rather than permanent investments or create persistent habits, and that the ongoing incentive of additional financial rewards is important for causing long-run lower electricity use.

There are relatively few papers that have studied programs offering financial rewards for meeting energy conservation targets.[1] Ito (2015) evaluates the financial rewards for energy conservation offered under the California 20/20 program. This program, implemented in response to a crisis in electricity supply, was mandatory for all eligible households and gave customers a 20% rebate on their summer electricity bill if they reduced electricity use by 20% compared to the previous year. Ito (2015) found that the program generated persistent reductions in electricity use for inland customers, but no reductions — short run or long run — for households in coastal climactic zones. He attributes this heterogeneous effect to the higher temperatures, lower income, and increased use of air conditioning in inland regions. Gerard and Costa (2015) study a suite of mandatory incentives introduced in response to an electricity supply crisis in Brazil. They find the incentives, including fines for overconsumption and rebates for reductions, caused both a large short-run and then smaller but persistent long-run reduction in energy use. Dolan and Metcalfe (2015) undertake a randomized controlled trial (RCT) and find large financial rewards cause a large conservation over the two months of their treatment period. They find persistent effects over the two post-program months they observe.

This chapter differs from the above work in two important ways. First, I find that electricity use rebounds as households leave the program and do not find evidence of long-run persistent reductions. In addition, I do not find evidence of significant heterogeneous effects across household types; reductions come primarily from changes in non-heating electricity use. This, along with the heterogeneous treatment effects found by Ito (2015), suggest that whether households respond to conservation targets with persistent reductions may depend on the way electricity is used within a home. Alternatively, the context in which a financial reward program is offered — during an electricity crisis as in Ito (2015) and Gerard and Costa (2015), or routine electricity use as in Team Power Smart studied in this thesis — may be important to whether households respond with short-run reductions or persistent changes. This chapter's finding of continued reductions in electricity use when households re-enroll, and a rebound when they don't, is similar to the "action and backsliding" found by Allcott and Rogers (2014) in households' responses to repeated home energy reports. Allcott and Rogers (2014) show these home energy reports mailed to households can cost-effectively cause reductions in energy use. Importantly, they also show that electricity use initially rebounds after the arrival of home energy reports, but repeated reports can cause persistent changes in electricity use.

Second, voluntary conservation programs are characterized by potential self-selection into participa-

[1] Several papers, primarily from the psychology literature, have undertaken randomized control trials of financial rewards (Mizobuchi and Takeuchi, 2012; Midden et al., 1983; McClelland and Cook, 1980; Winett et al., 1978). However, these consider very short timeframes, small sample sizes, and unrepresentative electricity users and are of limited use in understanding how households respond to financial rewards.

tion. This poses different difficulties for identifying causal program effects than in mandatory programs and different considerations for the program cost-effectiveness. The cost-effectiveness of voluntary programs may be overstated if households self-select into the program to receive credit for reductions in electricity use that are not additional to what would have occurred in the program absence. This 'additionality' problem is well known and can be substantial. For example, Boomhower and Davis (2014) use several eligibility thresholds in a program offering subsidies for replacing appliances to find that half of the participants would have invested in the energy-efficient technology even without the subsidy. Alternatively, self-selection can increase a voluntary programs cost-effectiveness by raising the share of participant households that respond to the incentive. This is particularly important for electricity conservation as household electricity use, even after controlling for weather changes, exhibits large idiosyncratic variations over time. As a result, even households that do not respond to the incentive may still receive rewards for sufficiently large idiosyncratic reductions in electricity use, and this share can be substantial in mandatory programs. These two offsetting types of self-selection make it difficult to predict the cost-effectiveness of voluntary programs and necessitates their ex-post evaluation.

The remainder of this chapter is organized as follows. Section 1.2 describes the institutional setting, design of the Team Power Smart electricity conservation program, and data. Section 1.3 gives an overview of the two empirical approaches used in this dissertation and then describes the event study empirical strategy of this chapter. Section 1.4 presents the empirical results, and I conclude in Section 1.5.

1.2 Institutional Setting, Program Design, and Data

1.2.1 The Team Power Smart Program

BC Hydro is Canada's second largest integrated electrical utility company and serves 1.7 million residential customers covering 95% of the population in British Columbia (BCH, 2014). BC Hydro is owned and has a mandate set by the provincial government of British Columbia. The B.C. government, through its Clean Energy Act, has required that BC Hydro "[meet] at least 66 per cent of the expected increase in demand through conservation and efficiency by 2020" (BCH, 2014). As part of their efforts to achieve this mandate and in response to the Conservation Potential Review, BC Hydro launched a new conservation initiative — a program targeting ongoing behaviours called Team Power Smart.[2] Team Power Smart is a voluntary program promoted and summed up by BC Hydro with "*Looking to save money on your electricity bills? Become a member of Team Power Smart and challenge yourself to reduce your home's electricity use by 10% in the next year. If you're successful, you can earn a [$75] reward.*"[3] This electricity conservation challenge requires households to reduce their aggregate electricity use over the 12-month challenge by 10% relative to a conservation target. Each household's

[2] It is possible for households to join Team Power Smart to view their electricity use online without undertaking a conservation challenge. For simplicity, I will use Team Power Smart to refer to those households which also undertake a conservation challenge. I do not observe households which registered online without undertaking a conservation challenge.

[3] BC Hydro Team Power Smart website landing page. The reward value for challenges studied in this thesis is $75 which was reduced to $50 in September 2014. I exclude households undertaking a challenge under the new $50 reward value. Accessed June 2017.

conservation target is their own annual electricity use over the preceding 12 months, adjusted for changes in heating degree days to help prevent households from being unduly penalized or rewarded for changes in weather. Beginning a challenge requires a minimal time cost of registering online. Households can start a challenge in any month of the year as long as they have 12 months of electricity use in their current home to establish their target. Online signup ensures that all participants can view their progress towards their conservation target through the BC Hydro website and access a variety of tips and suggestions for reducing their electricity use. The online account provides households with feedback on both their monthly and cumulative progress towards their annual 10% conservation target; examples of this are shown in Figures 1.2a and 1.2b. Because it is the aggregate annual conservation that matters for success in a challenge, households can miss their 10% target in any month and still pass the challenge.

Upon completing the 12 months of the challenge BC Hydro undertakes a final evaluation of household's cumulative conservation. BC Hydro applies the final weather adjustment, accounts for bi-monthly billing and any idiosyncratic factors, and evaluates whether the household passed or failed their challenge. While the conservation target advertised to customers is 10%, BC Hydro evaluates final success or failure against a 9.5% conservation threshold. Households that reduced their electricity use by greater than or equal to 9.5% below their target pass their challenge while the rest fail. BC Hydro notifies all customers of their success or failure and gives successful households the choice of a rebate through either a cheque or credit applied to their account.

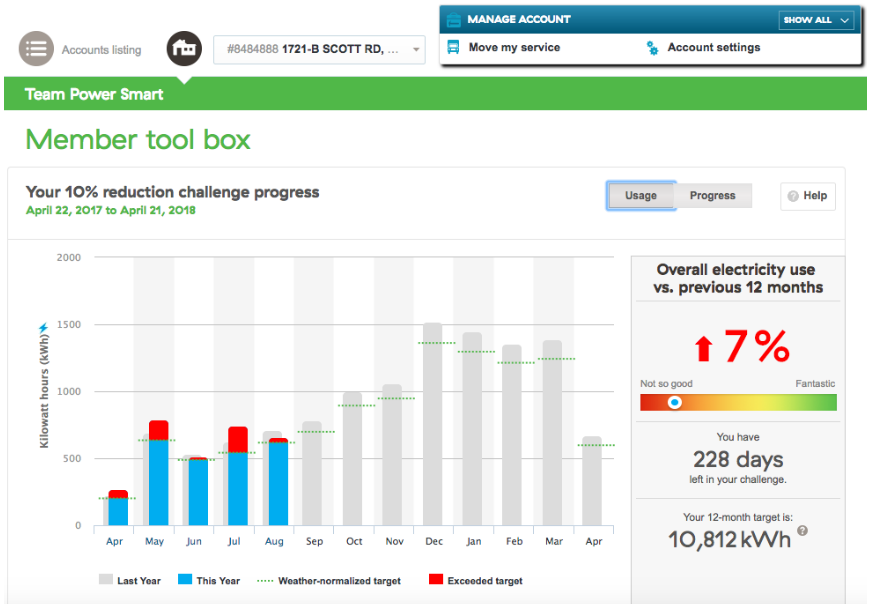
1.2.2 Structure of Additional Conservation Challenges

A novel feature of Team Power Smart is that all households have the option of re-enrolling in additional annual conservation challenges. Upon completion of each challenge both households that pass and fail their challenge are given the same option to start a subsequent conservation challenge for another $75 rebate. This process is summarized in Figure 1.2. Each subsequent challenge follows the same process as the initial conservation challenge; households have a goal of another 10% conservation target measured and weather adjusted relative to their previous 12 months of electricity use. The new reduction target is independent of whether the prior 12 months contained a challenge or not, and independent of whether the prior challenge was a success or failure. The baseline for a household immediately starting an additional challenge would be the 12 months of the just completed challenge, while a household waiting 4 months before starting their next challenge would have a baseline set by the average of their last 8 months of their previous challenge and the 4 months of the gap prior to starting their next challenge.

Because each additional 10% conservation challenge is evaluated relative to the prior 12 months, the reduction in electricity use achieved by a household during a challenge affects their incentives on when and whether to undertake a subsequent challenge. For example, a household that achieves a 20% reduction during their first challenge and immediately undertakes a second challenge will have to reduce their emissions by a further 10%, for a cumulative 28% reduction, to pass their next challenge and obtain the rebate. However a household that passes the first challenge with a 10% reduction will only have to achieve a cumulative 19% reduction for the same second rebate. Under the reasonable assumption of increasing marginal costs to electricity conservation, the greater the conservation achieved during a

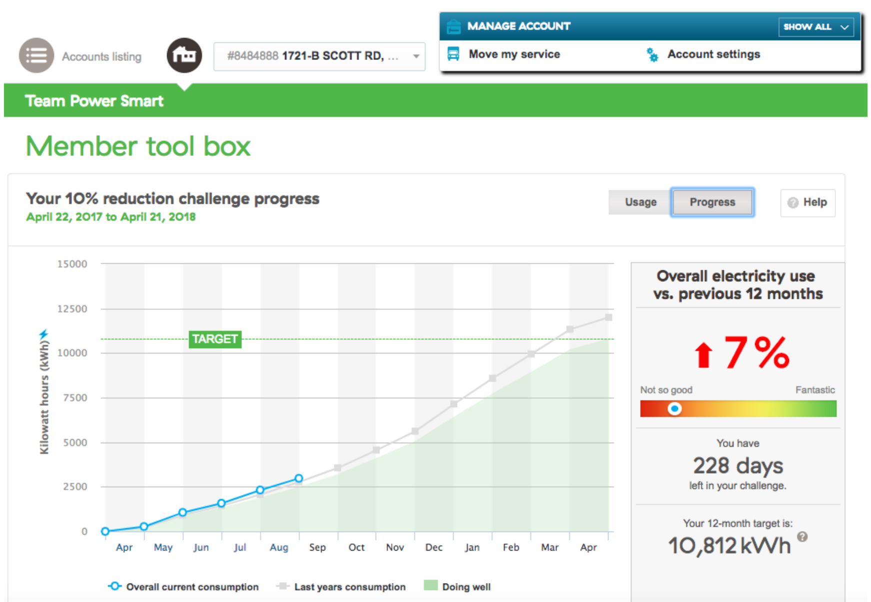
Figure 1.1: **BC Hydro's Online *Member Tool Box***

(a) **Monthly Challenge Progress**



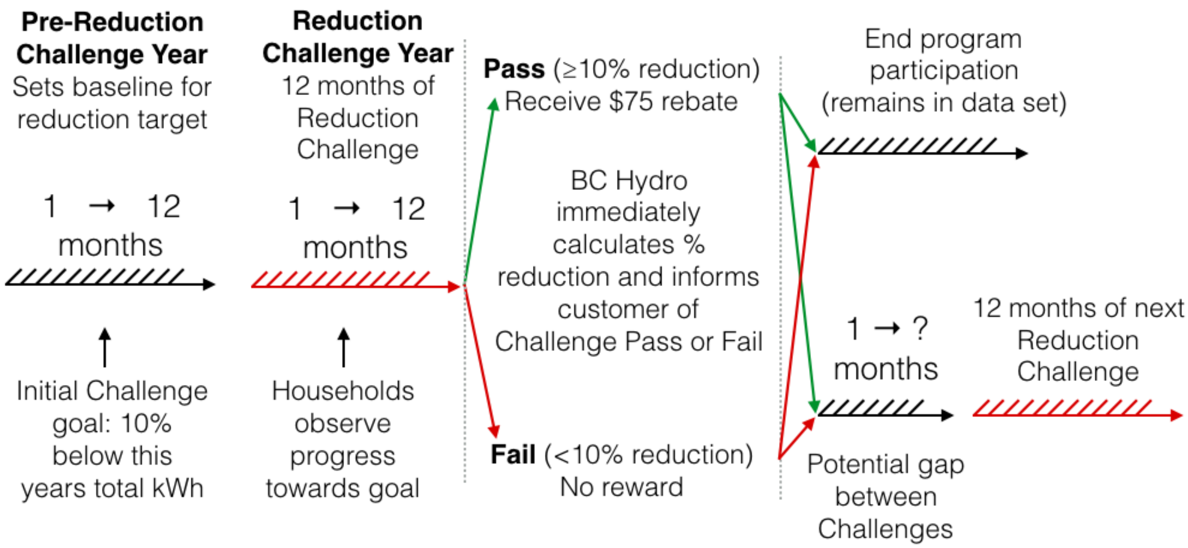
Notes: BC Hydro provides households participating in a conservation challenge with an online portal showing their electricity use and progress towards their target. The online portal includes information on monthly electricity use compared to the same month the previous year and their 10% conservation target.

(b) **Cumulative Challenge Progress**



Notes: In addition to monthly electricity use the online portal displays a household's cumulative progress towards their annual 10% conservation target.

Figure 1.2: **Decisions in a Conservation Challenge**



Notes: This figure summarizes the options available to a household upon completing a conservation challenge. Every conservation challenge follows the same process independent of how many prior challenges were undertaken or whether they were successful.

challenge the greater the incentive to postpone a subsequent challenge or leave the program.

1.2.3 Data and Household Characteristics

Under a non-disclosure agreement with BC Hydro I obtained an anonymized sample of monthly electricity billing records for 10,000 Team Power Smart program participants and 20,000 non-participants from January 2006 to December 2015.[4] The panel includes customers' Team Power Smart program participation history including the number of conservation challenges, each challenges start and end date, whether the challenge was successful, and the building and heating type of the household. BC Hydro also provided the weather-adjusted annual electricity use totals used to evaluate a household's success against the 9.5% conservation threshold. Individual household characteristics from BC Assessment including building type, number of bedrooms, assessed value, floor space, and the postal codes Forward Sortation Area were merged with the BCH Hydro panel. Removing duplicate accounts, erroneous data, and dropping households with electricity use more than 5 standard deviations from the mean left a sample of 9,817 households participating in Team Power Smart.

The sample of households provided by BC Hydro was a random sample of participant households from the British Columbia Lower Mainland region which covers 60% of the province's population (BCStats, 2016).[5] Temperatures range from a summer average of 18°C to winters averaging 4°C (ECCC, 2017). Electricity use in British Columbia peaks in the winter due to the widespread use of electricity for heating and the limited use of air conditioning in the summer, and BC Hydro estimates that 46% of residential electricity use in British Columbia comes from electric heating.

[4]I define participant households as those which participate in Team Power Smart prior to the panel end in December 2015, and non-participant households as those which do not participate. The sample was selected from households which did not move over the panel period.

[5] Another 18% of the province's population lives in regions with similar coastal climatic zones.

Table 1.1 compares the household characteristics of program participants to non-participants. The principal difference is that participants in Team Power Smart are more likely to live in apartments or townhouses compared to single family dwellings and are more likely to use non-electric heating. BC Hydro classifies households into heating categories based on surveys of residents and information on the building where the meter is installed. Non-Electric are households that heat primarily from sources other than electricity. Electric are households that heat primarily from electricity, and Unknown are unclassified households. Importantly, BC Hydro does not classify households into heating categories based on their observed electricity use. Differences in the composition of participant and non-participant household types cannot be attributed to self-selection into the program. This is because BC Hydro engages in a range of advertising for Team Power Smart that will differently affect households' awareness of the program and thus their likelihood of becoming participants. Figure 1.3 shows the average electricity use for these households over the panel. Non-participant households have significantly higher electricity use, particularly in the winter months, and electricity use is declining over the period studied among both participants and non-participants. The difference in electricity use between participants and non-participants is almost entirely a composition effect; after controlling for building and heating type the average electricity use among participant households is 1.1% higher (p-value 0.08) than non-participants during the pre-program year of 2006. Average monthly electricity bills among participant households are $62 - making the rebate reward of $75 equivalent to 10% of a household's annual electricity bill in addition to their bill savings.

1.2.4 Outcomes During Multiple Conservation Challenges

Table 1.2 summarizes the decisions and outcomes of participant households across multiple challenges. During the initial three challenges approximately 59% of households decide to re-enroll in an additional challenge.[6] Consistent with an increasing difficulty of achieving additional reductions in electricity use, the unconditional probability of passing a conservation challenge declines with additional challenges. Households are more likely to re-enroll in another challenge if they pass, rather than fail, their current challenge. In contrast, households are less likely to pass their next challenge if they passed their previous challenge. This pattern matches the incentive structure previously discussed; passing a challenge requires achieving the 9.5% conservation target, which makes passing the next challenge harder.

As households choose the start date of subsequent conservation challenges they could strategically establish a new baseline before undertaking their next challenge. Households could in theory increase their electricity use (or stop any ongoing efforts to reduce their electricity use) to create a new higher baseline that would make their subsequent conservation challenge easier to achieve. Figure 1.4 shows no obvious evidence of this; most households, if they continue to additional challenges, begin their next challenge in the first 3 months immediately after completing their prior challenge and there is no obvious bunching at 12 months.

The option to undertake a subsequent challenge does not expire; households can sign up for another challenge immediately or postpone indefinitely. Figure 1.5 shows the distribution of start dates for challenges one through four. New households continually enroll in TPS throughout the panel and

[6] The probability declines with higher challenges in part mechanically due to the limited panel length.

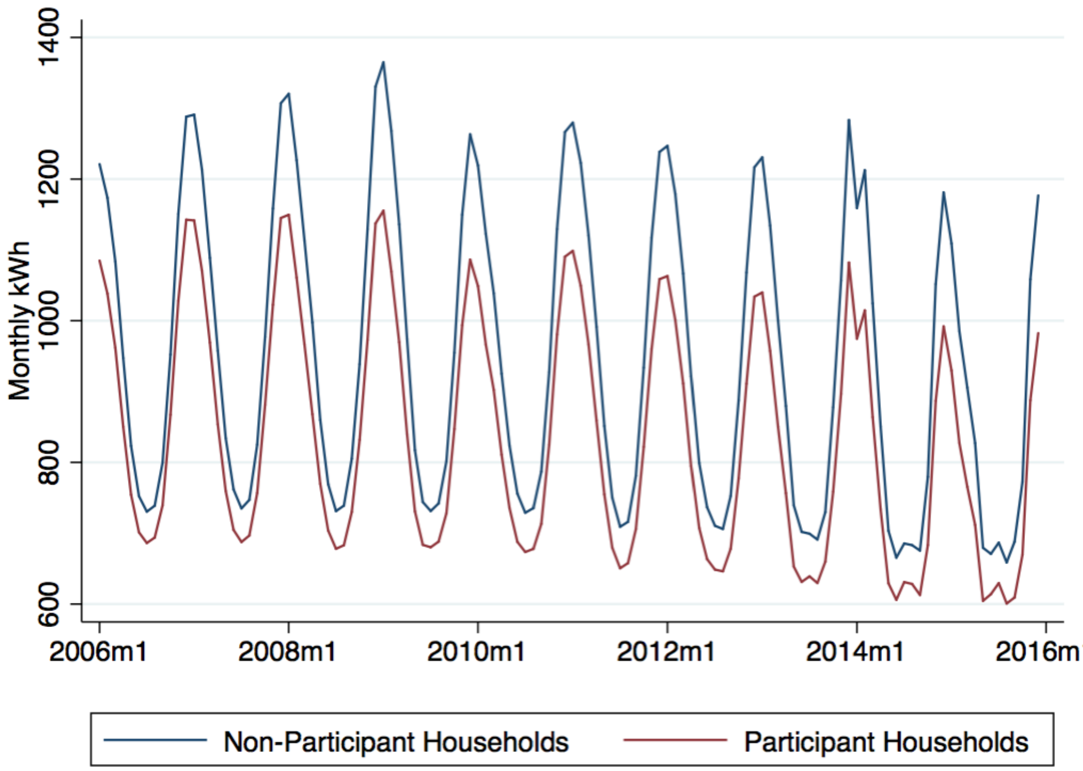
Table 1.1: **Participant and Non-Participant Characteristics**

	Participants		Non-Participants	
	N	%	N	%
Building Type				
1 Story Single Family Dwelling	3,796	39	8,764	46
2 Story Single Family Dwelling	2,716	28	5,088	26
1.5 Story Single Family Dwelling	400	4	977	5
Apartment	1,412	14	1,813	9
Townhouse	1,202	12	1,677	9
Other	291	3	931	5
Heating Type				
Non-Electric	5,599	57	9,687	50
Electric	2,874	29	7,294	38
Unknown	1,344	14	2,269	12
Bedrooms				
0	12	0	12	0
1	505	5	703	4
2	1,597	17	2,745	14
3	3,626	34	6,709	35
4	2,632	23	4,569	24
5 or more	1,814	18	4,512	23
Total HH's	9,817	100	19,250	100

	Participants		All Non-Participants	
	Mean	SD	Mean	SD
kWh	884	568	972	636
Average Monthly Bill	$62		$69	
Value ($1,000)	$664	$467	$721	$575
Floor Area (Square Feet)	2025	934	2123	997

Notes: This table shows the building characteristics of participant households and non-participant households. These households are chosen from a random selection of British Columbia lower-mainland households, which is primarily an urban and suburban area concentrated around the Vancouver metropolitan area.

Figure 1.3: **Time Trends in Electricity Use**



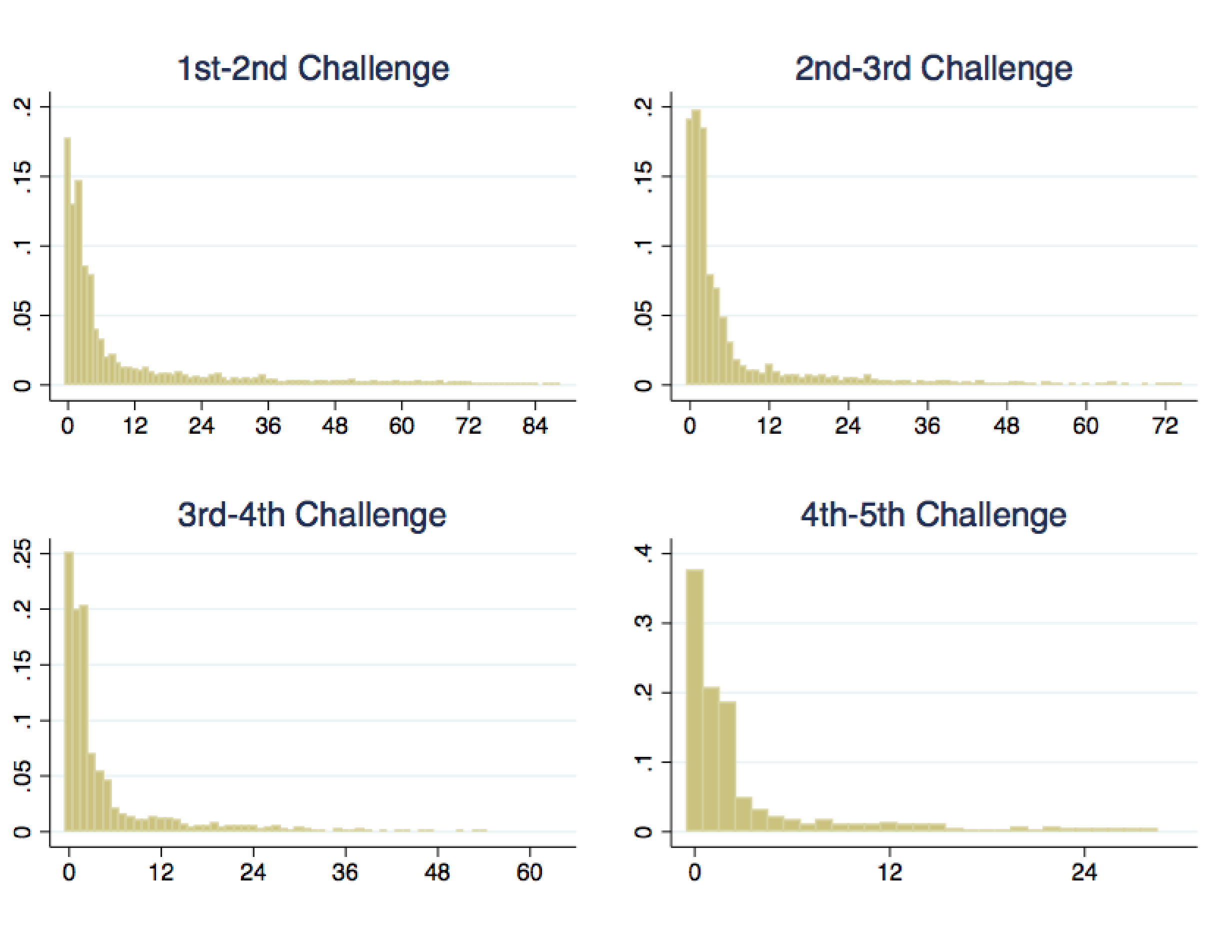
Notes: Average monthly electricity use for participant households and the full sample of non-participant households.

Table 1.2: **Probability of Challenge Outcomes**

Challenge	Households Undertaking Challenge	**Probability Of Re-Enrolling If**			**Probability Of Passing If**		
		All	Failed Challenge	Passed Challenge	All	Failed Prev. Challenge	Passed Prev. Challenge
1	9,817	0.57	0.50	0.71	0.34		
2	5,638	0.59	0.55	0.70	0.31	0.33	0.28
3	3,346	0.60	0.56	0.71	0.28	0.30	0.24
4	2,014	0.54	0.51	0.64	0.26	0.28	0.23
5	1,091	0.46	0.41	0.60	0.24	0.27	0.17
6	498	0.38	0.36	0.44	0.24	0.25	0.21
7	188	0.27	0.26	0.28	0.29	0.28	0.31
8	50	0.12	0.07	0.33	0.18	0.20	0.13
9	6	0.00	0.00	0.00	0.33	0.33	0.33

Notes: Probability of Re-Enrolling is the probability of re-enrolling in a subsequent conservation challenge, conditional on being in the current challenge. Probability of Re-Enrolling if Failed [Passed] Challenge is the probability of re-enrolling conditional on failing [passing] the current challenge. The Probability of Passing is for a household's current challenge, while the Probability of Passing if Failed [Passed] Prev. Challenge is the probability of passing the current challenge conditional on the Fail or Pass status of the previous challenge.

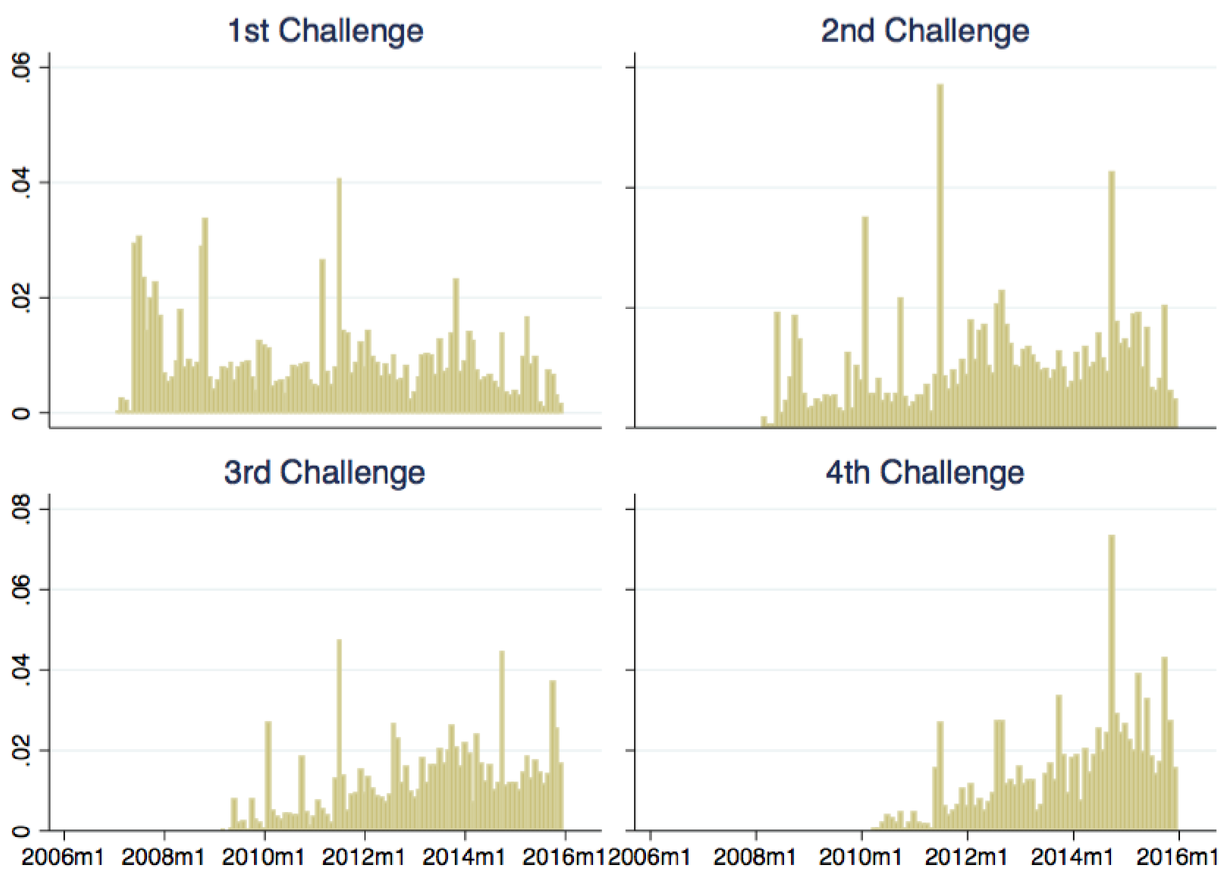
Figure 1.4: **Time Delay Between Conservation Challenges**



Notes: These histograms show the number of months households wait between conservation challenges. The majority of households which continue to additional challenges do so shortly after completing their prior challenge. The median wait after the 1st challenge is 3 months, 2 months after the 2nd challenge, 2 months after the 3rd challenge, and 1 month after the 4th challenge.

as time proceeds households that complete challenges continue to subsequent conservation challenges. Several dates show large increases in sign-ups; these are likely due to periods of significant promotion of the TPS program by BC Hydro as they do not coincide with previous months of unusually large or small electricity use or unusual changes in weather.

Figure 1.5: **Distribution of Challenge Start Dates**



Notes: These histograms show the start date for conservation challenges one through four. Several dates show large increases in the number of households starting a challenge. Periods of increased sign up do not coincide with unusual weather, seasons, or consumption, and are likely due to promotion of Team Power Smart by BC Hydro.

1.3 Empirical Strategies

There are three principal challenges to estimating the causal effect of repeated conservation challenges. First, as in all voluntary programs, households may self-select into Team Power Smart based on observable and unobservable time invariant characteristics. This self-selection could make non-participant households an unsuitable counterfactual for electricity use among participant households, had they not participated in the program. Secondly, households may start their first conservation challenge based on shocks to their past electricity consumption or expectations of their future consumption. For example, households may select into the program in response to a particularly cold winter which caused a large electricity bill. Households may also take advantage of anticipated reductions in their electricity use such as the purchase of an efficient dryer or leaving on holiday. By signing up in advance or in conjunction with their anticipated reduction in electricity use, a household could receive credit for reductions in electricity use that were not caused by the conservation program. Lastly, all households have the option of continuing to additional conservation challenges. This makes the persistence of energy savings and the causal effect of subsequent conservation challenges dependent on the households' decisions to select into additional challenges.

To address these challenges I employ two complementary empirical strategies. In Chapter 1 I use an event study research design to estimate the monthly average changes in electricity use associated with participation in Team Power Smart. By estimating changes over time within households, this strategy identifies program effects independent of self-selection into the program on observable and unobservable time invariant characteristics. For clarity, I use "program effects" to refer to changes in electricity use that are associated with participation in conservation challenges but which may or may not be causally due to the program. I restrict "treatment effect" to refer to standard causal effects. Plotting these estimates provides a visual time trend of the changes in electricity use leading up to the initial conservation challenge, during the months of each challenge, and over the months after a household leaves the program. These trends provide insight into potential self-selection into the program based on households' expectations of future electricity use and past shocks to their consumption, along with the persistence of reductions in electricity use and the effect of subsequent conservation challenges. This strategy provides detailed information on households' electricity conservation decisions and self-selection into the program, and finds no evidence that reductions in electricity associated with participation are not causally due to the program. However, because self-selection cannot be ruled out it cannot strictly identify causal treatment effects. To identify causal treatment effects, Chapter 2 exploits a discontinuity in the probability that households continue to a second conservation challenge. I use this discontinuous probability change in a fuzzy Regression Discontinuity design (fuzzy-RDD) to identify the Local Average Treatment Effect (LATE) of a second conservation challenge.

1.3.1 Event Study Empirical Strategy

An event study model identifies the monthly changes in electricity use associated with a conservation challenge by comparing participant households to households that have not participated in a challenge (Angrist and Pischke, 2008). With a control group of non-participant households an event study model is a difference-in-difference model generalized to multiple time periods and where the timing of treatment may vary across households. Multiple time periods allow lag and lead program effects to be estimated and plotted to visually inspect their pattern over time. Lagged program effects — effects occurring after the start of the initial conservation challenge — show how changes in electricity use evolve over time and can help distinguish short-run and long-run program effects. The pattern of lead program effects helps test the validity of the identifying assumptions. When each unit may have a unique date of treatment it is intuitive to consider program effects in "event time" instead of calendar time. This can be thought of as centering all households in time by their date of treatment and measuring event time as the elapsed time for each household relative to its individual date of treatment.

The general event study model is of the form

$$y_{it} = \sum_{\tau=-T}^{T} \beta_\tau D_{i,t-\tau+1} + \alpha_i + d_t + \epsilon_{it} \tag{1.1}$$

where y_{it} is the log of monthly electricity use for individual i in month t, α_i is an individual fixed effect, and d_t is an indictor for date t. $D_{i,t-\tau+1}$ is a dummy variable equaling 1 if individual i in month t began treatment in month $t-\tau+1$, where τ is the measure of event time in months.[7] I define the start of treatment as the month a household undertakes its first conservation challenge, $\tau = 1$. β_τ are the non-parametric program effects τ months lag or lead of treatment and cover all periods in the panel of length T. Event study models can only identify the full set of program effects $\{\beta_\tau\}_{\tau=-T}^{T}$ up to a constant. That is, event study models can only identify changes in program effects relative to a baseline level.[8] In my preferred specification, I define the baseline as the second year before each household undertakes their initial conservation challenge. Using a baseline of the second year allows important pre-treatment trends in the 12 months preceding participation in the program can be estimated. Results are robust to excluding other time periods; see Figure A.3 in the Appendix for a baseline of the third year pre-treatment. For the baseline of the second year, I estimate equation (1.1) excluding indicators $D_{i,t-\tau+1}$, $\tau = [-23..-12]$ which is equivalent to defining $\{\beta_\tau \equiv 0\}_{\tau=-12}^{-23}$. The non-parametric estimates $\hat{\beta}_\tau$ identify the average percentage change in monthly electricity use within a household relative to their average electricity use during the baseline period.

Identification Challenges

The identifying assumption for causal treatment effects in the event-study model is that, in the absence of treatment, treated households would have the same expected outcome as non-participant households. This assumption could be violated in two ways. First, as shown in Table 1.2, participant households differ from non-participant households in observable time-invariant characteristics. The composition of participant households could potentially change over time, for example if high electricity

[7] For example, consider an observation in December 2009 for a household that began treatment (began its initial conservation challenge) in October 2009. For this household and observation, $t = \textit{December}$ 2009 and $t - \tau + 1 = \textit{October}$ 2009. This finds $D_{i,t-\tau+1} = 1$ only for $\tau = 3$.

[8] This can be seen by adding a constant to all β_τ and noting that this constant is then collinear with both the full set of individual and date fixed effects.

Table 1.3: **Pre-Program Trends in Electricity Use**

	(1)	(2)	(3)	(4)	(5)
β_1 : Date	-0.00012	-0.00024**	-0.00033***	-0.00064***	-0.00058***
	(0.00016)	(0.00011)	(0.00009)	(0.00007)	(0.00006)
β_2 : Date×Participants	0.00005	-0.00026	-0.00021	-0.00026*	0.00016
	(0.00023)	(0.00018)	(0.00016)	(0.00015)	(0.00025)
Pre-Program Years	2	3	4	5	6
Participant IDs	7182	7182	7182	7182	7182
Non-Participant IDs	5083	3782	2914	1943	467
Observations	294359	394703	484607	547498	550725

Notes: β_1 : *Date* is the pre-program time trend common to participant and non-participant households. β_2 : *Date* $\times$ *Participants* is the additional time trend specific to participant households. Pre-Program Years is the length of time for which time trends are estimated and excludes all households that start a challenge within six months after the given pre-program. The six month period is to avoid pre-treatment trends that could include anticipation effects in the final months pre-treatment. Standard errors in parentheses clustered at the household level. *** p<0.01, ** p<0.05, * p<0.1.

consumption households select into the program earlier compared to low consumption households, and bias estimates of $\hat{\beta}_\tau$. This motivates the inclusion of individual fixed effects in (1.1) to control for time-invariant characteristics that may change over the panel. With individual fixed effects the identifying assumption requires treatment to be as good as randomly assigned conditional on time-invariant characteristics. This fails if participant and non-participant households do not have parallel trends in the absence of treatment. Evidence that the parallel trends assumption holds is obtained by comparing pre-treatment linear trends between participant and non-participant households. I estimate the following specification,

$$y_{it} = \beta_0 + \beta_1 m_t + \beta_2 d_t \times TPS_i + \beta_3 TPS_i + \epsilon_{it} \tag{1.2}$$

where y_{it} is log monthly electricity use, m_t is the date, and TPS_i is an indicator equal to one if household i is a Team Power Smart participant household. The only pre-treatment period available to all households is the year 2006. To test trends over multiple years I estimate specification (1.2) for several different time periods and include only participant households that do not begin a conservation challenge until 6 months after the initial pre-treatment years indicated. Table 1.3 shows the results where β_1 is the percent change per month for non-participant and participant households and β_2 is the additional monthly percent change for participant households. Participant households do not have a significantly different, at the 5% level, pre-treatment trend from non-participants. The magnitude of diverging time trends is also small. Taking the largest point estimate of different trends, $\hat{\beta}_2 = -0.00026$, would imply an upper bound to the potential bias in estimated program effects of only 0.3% at the end of the first conservation challenge.

Households starting a conservation challenge in response to past consumption shocks or expectations of future electricity use will also violate the parallel trends assumption. Shocks to past consumption can be tested by examining pre-treatment trends. As I discuss in Section 1.4, I find no evidence households begin a conservation challenge in response to increases in their past consumption such as large electricity bills from unusually cold winters. Whether households begin a challenge based on their expected future consumption cannot be tested. For example, it is not possible to distinguish between a household participating in a conservation challenge which causally reduces their electricity use from a household purchasing an efficient dryer and the latter participating to take advantage of their upcoming decline in electricity use. In this case, reductions in electricity use that coincide with the start of participation in Team Power Smart would not be causal. Lacking a suitable instrument for initial participation or a discontinuity in eligibility for program participation such as used by Ito (2015), self-selection based on expectations of future electricity use cannot be ruled out. Instead, as I discuss below, the pattern of estimated program effects provides evidence that such self-selection is unlikely to be substantial and that the estimated reductions are primarily causally due to the Team Power Smart program.

1.4 Event Study Estimates - Initial Conservation Challenge

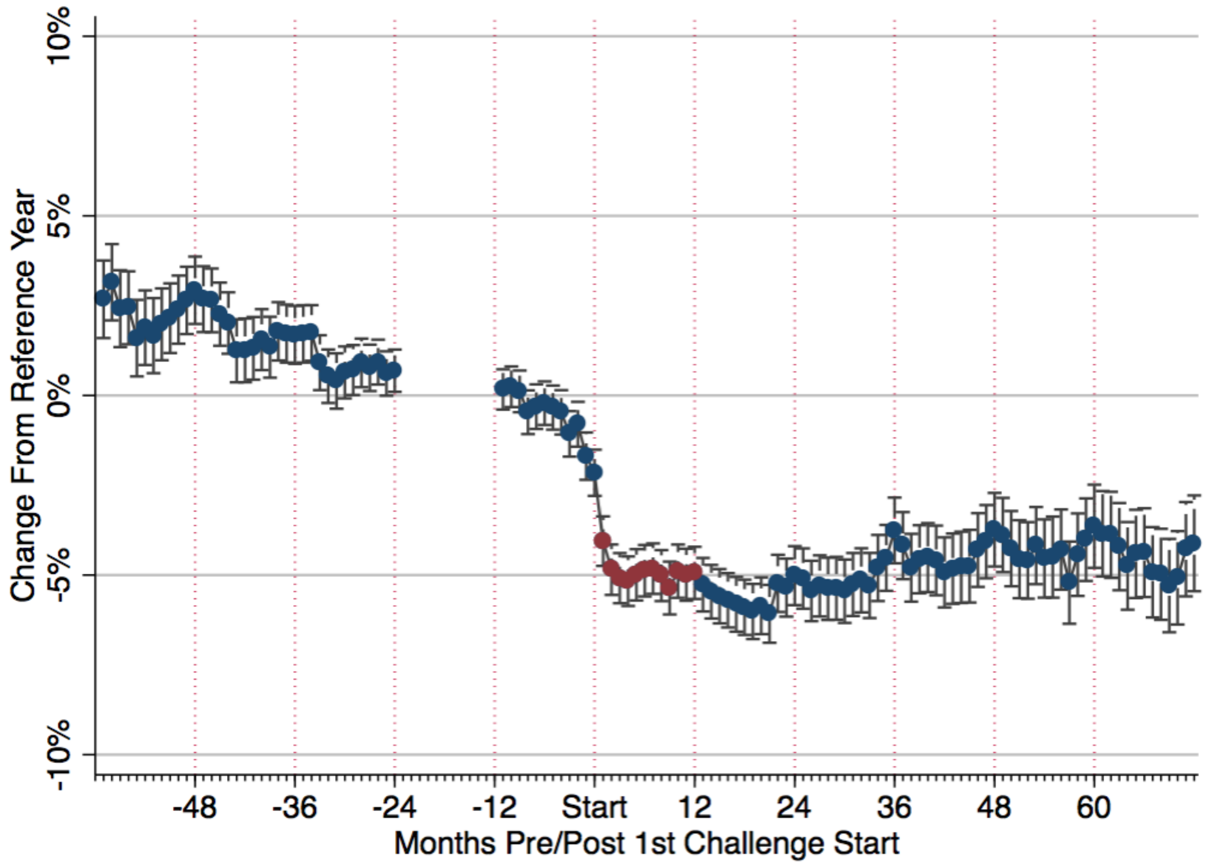
This section presents the results of estimating the event study model, equation (1.3), for all participant households and a comparison group of non-participant households.[9] Estimates are presented in Table D.1 in the Appendix. To visualize the trend of program effects over event-time τ, I plot the estimates $\hat{\beta}_\tau$ in Figure 1.6. The pattern of monthly estimates provides substantial insight into the short run and long run reductions in electricity use and whether reductions are causally due to the Team Power Smart program. The estimates in Figure 1.6 can be separated into three time intervals: the pre-treatment period leading up to the initial conservation challenge (months -60 to 0), the twelve months of the initial conservation challenge (months 1 to 12), and the months after the initial conservation challenge is completed (months 13 to 72).

In the second time interval, the twelve months of the initial conservation challenge show a substantial average reduction in electricity use of 5.3% relative to the baseline of the second year pre-treatment. BC Hydro measures households' reductions relative to the twelve months preceding a challenge; households achieve a 4.3% reduction relative to this year and their 10% conservation target. With the exception of the first month, the reductions over the twelve months of the initial challenge are not statistically significantly different from each other at a 5% level. This shows that, on average, households are not strongly increasing or decreasing their electricity conservation during the challenge. This stable conservation is consistent with households making either an initial investment that causes persistent reductions throughout the challenge, or maintain a constant conservation effort throughout the challenge. The consistent program effects also indicate that self-selection to take advantage of upcoming short-term changes in electricity use, such as a holiday, is not a significant cause of participation; if they were, reductions in energy use would be expected to spike in the initial months of the challenge before partially rebounding.

In the first time interval, the months leading up to the initial conservation challenge show a gradual

[9] Due to the small sample size for months far from the initial conservation challenge I pool monthly indicators before and after the ± 5 year estimation window around the initial conservation challenge into separate indicators. Results are robust to this simplification. To ensure a closer comparison group of non-participants to participant households, I also use a random sub-sample of non-participant households with the same composition of building type and heating characteristics as participant households. Results are robust to including all non-participant households; see Figure A.2 in the Appendix.

Figure 1.6: **Estimated Program Effects For All Households**



Notes: This figure plots estimated program effects $\hat{\beta}_\tau$ with 95% confidence intervals from equation (1.3) ordered by event-time τ. Estimates are in Appendix Table D.2. Point estimates in red denote the 12 months of the initial conservation challenge ($\tau = [1..12]$). The pre-treatment period is denoted by the months prior to *Start* ($\tau \leq 0$). The period after the initial challenge concludes are the months with $\tau > 12$. The visual gap in estimates between months $\tau = -11$ and $\tau = -23$ is the excluded reference period. $\hat{\beta}_\tau$ identify the percent change in electricity use relative to the average electricity use within a household during this excluded reference year. Estimates include individual and date fixed effects and I cluster standard errors at the household level.

decline in estimated program effects. This pre-treatment trend could arise for two reasons. First, the pre-treatment trend could indicate a violation of the parallel trends assumption — differences in exogenous trends between the treated and control households — despite the lack of statistically significant different pre-trends estimated previously. To test this, I estimate specification (1.1) without non-participant control households. This identifies the treatment effect by comparing currently treated households to a control group composed of households treated at a later date in the panel (Borusyak and Jaravel, 2016). I plot estimates in Appendix Figure A.1. Estimates lose precision due to the smaller set of households but continue to show the same pre-treatment trend, showing the pre-treatment decline is not a due to a violation of the parallel trends assumption. Second, the pre-treatment decline could reflect different time trends among participant households that are not fully controlled for by common date fixed effects. Figures 1.7 through 1.9 plot the program effects for households undertaking different numbers of challenges. This shows that the declining pre-treatment trend is limited to households undertaking only one or two challenges and, importantly, does not bias the short or long-run program effect estimates.

A threat to estimated program effects not being causal is households beginning a conservation challenge in response to a high electricity bill, such as after a cold winter. In this case, reversion to the mean would result in reductions in electricity use, relative to the previous year, being credited to the program. If this self-selection occurs it will manifest itself as positive pre-treatment effects in the months immediately prior to the initial conservation challenge. Instead, the opposite effect is found; the last two months (months $\tau = [-1..0]$) in Figure 1.6 show a decline in electricity use relative to the previous months. These reductions in electricity use before the program starts do not count towards the conservation credited to households and, in addition, make it harder to achieve the financial reward as they lower the baseline from which the 10% conservation target is measured from. This lack of a positive pre-treatment effect in the months prior to the challenge suggests that households do not self-select into the program based on past consumption shocks.

The pre-program decline of months $\tau = [-1..0]$ may indicate that households are undertaking electricity conservation prior to the program start. This could indicate self-selection into Team Power Smart as a result of making an energy efficiency investments, which would bias the estimated program conservation upwards. There are two reasons such self-selection and bias to the estimated program effects is unlikely to be large. First, participation in Team Power Smart as a result of making an investment would result in persistent reductions in electricity use; in contrast, Subsection 1.4.1 shows that electricity use rebounds as households leave the program. Second, the pre-program roll-off can result mechanically from the billing process. During the period of the panel studied in this thesis BC Hydro does not record electricity bills on a fixed monthly basis.[10] Instead, BC Hydro uses a rolling billing period where different houses are billed on different days of not more than 62 days and the use is calendarized to monthly consumption. As a result, reductions that occur after the start of a

[10]BC Hydro began a rolling installation of Smart meters in 2011 which allowed collection of hourly electricity use. The event study estimates of this section includes households beginning their initial challenge prior to February 2013 and so will include some households with electricity use recorded at a higher frequency. However, all data provided to the researcher was harmonized and calendarized by BC Hydro to a monthly level regardless of the recording frequency. As a result, challenges over the period studied are recorded as beginning on the first day of a month regardless of when during the month a household signed up for a challenge, or when the electricity meter was read.

conservation challenge cannot be separated within a billing cycle from electricity use that occurred prior to the challenge start. This can result in reductions due to a conservation challenge being partially credited to up to the last two months before a household begins its challenge.[11]

The third time interval includes all months after the initial challenge completes, months $\tau = [13..72]$. Pooling all participant households in estimating (1.1) includes those ending their program participation after their first challenge, those immediately continuing to additional challenges, and those waiting several months to years before starting a subsequent conservation challenge. Estimates $\hat{\beta}_\tau$, $\tau > 12$ are the average change in electricity use across these households, including any rebound in electricity use, and additional treatment effects from subsequent challenges undertaken by households. While Figure 1.6 shows that participation in Team Power Smart is associated with long-run average reductions in electricity use, this does not distinguish persistent energy savings due to a challenge, from the effects of subsequent challenges. To distinguish these effects I modify the event-study model (1.1) and estimate it separately for households ending their participation and those re-enrolling in subsequent conservation challenges.

1.4.1 Event Study Estimates - Additional Conservation Challenges

Comparing the estimated program effects across households that undertake different numbers of conservation challenges sheds light on both the process of self-selection into additional challenges and the persistence of energy savings. To do this, I pool households by the number of conservation challenges they undertake and estimate separate event study models for each group of households, equation (1.3)

$$y_{it} \quad = \quad \alpha_i + d_t + \sum_{\tau=-59}^{72} \beta_\tau D_{i,t-\tau+1} + \sum_{g=1}^{8} \theta_g G_{itg} + Pre_{it} + Post_{it} + \epsilon_{it} \tag{1.3}$$

where y_{it}, α_i, d_t, and $D_{i,t-\tau+1}$ are defined as in the event-study model of equation (1.1). I pool monthly indicators before and after the ± 5 year estimation window around the initial conservation challenge into, respectively, indicators Pre_{it} and $Post_{it}$. As shown in Figure 1.4, some households undertaking multiple conservation challenges have a gap in time between when their previous challenge completes and they re-enroll in a subsequent challenge. To account for this variable gap I include the indicator G_{itg} in equation (1.3). G_{itg} is 1 if household i in month t has completed challenge g but has not yet re-enrolled in challenge $g+1$. This indicator is not necessary for estimating the event study model; instead, it simplifies the comparison of program effects across households with different gap lengths. For households undertaking two or more challenges, including G_{itg} defines the coefficients $\beta_\tau, \tau = [12..23]$ as the program effects of the twelve months of the second conservation challenge regardless of the length of gap between conservation challenges. For households ending their participation after a single challenge, $\beta_\tau, \tau = [12..23]$ are the post-program effects for the first 12 months immediately following the initial challenge. Similarly, including G_{itg} defines $\beta_\tau, \tau = [24..35]$ as the program effects during a third conservation challenge for households that re-enroll, and as the second year post-program effects for households that end their participation. In my preferred specifications I pool all households that

[11] Households could also potentially begin reducing their electricity use in response to a conservation challenge but not complete their online registration until some weeks later.

re-enroll within 12 months of completing their previous challenge and exclude households with longer gaps between challenges; results are robust to alternate gap lengths between challenges.

Estimates for Additional Conservation Challenges

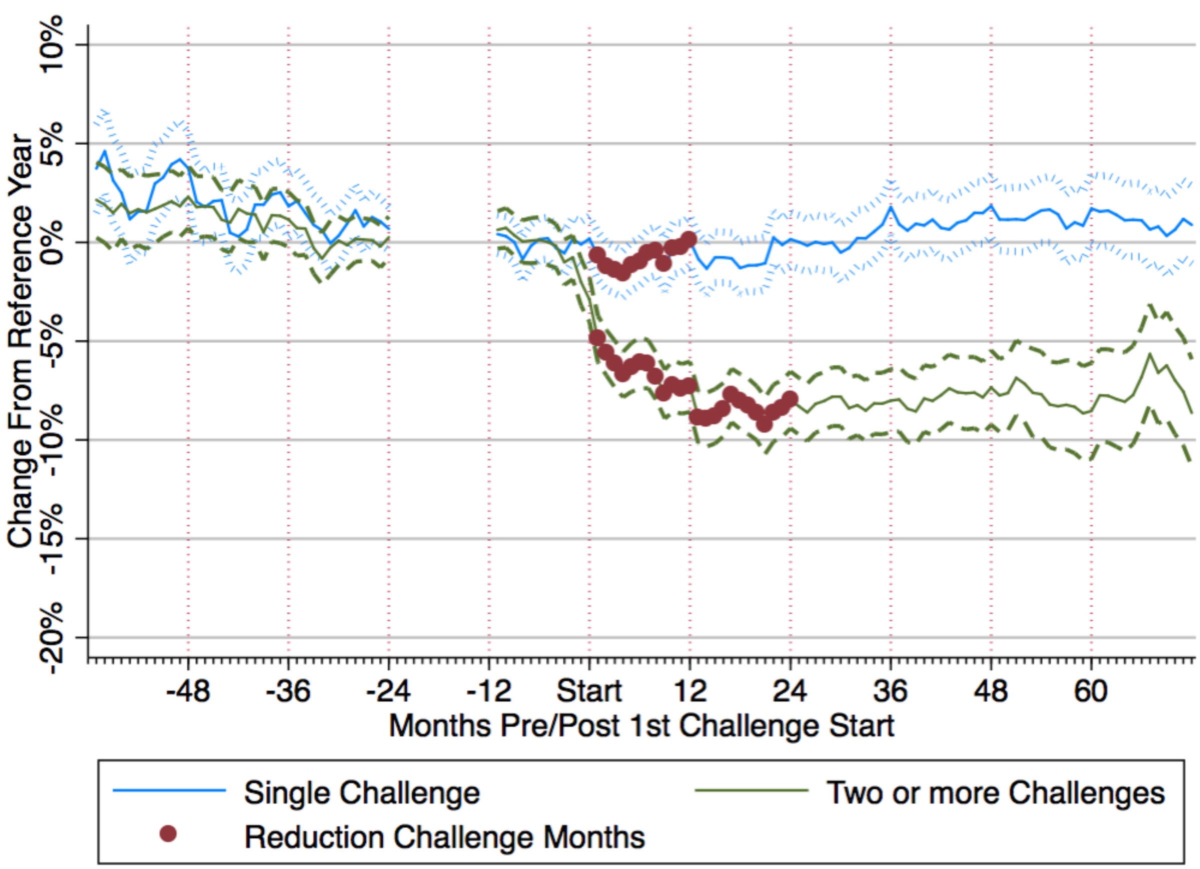
I separately estimate equation (1.3) for six subsets of households depending on how many conservation challenges they undertake; these subsets are not mutually exclusive. Estimates are presented in Table D.2 in the Appendix. In Figure 1.7 I plot estimates $\hat{\beta}_\tau$ for households that end their program participation at a single challenge, and separately for households that re-enroll in a second conservation challenge. These two household groups show significant differences. Households that end their program participation after the initial challenge have average reductions during the challenge of 0.5%. Households that re-enroll in a second conservation challenge have average reductions during their first challenge of 6.5%. This suggests self-selection into additional challenges based on the reductions in electricity use achieved during the initial conservation challenge. Alternatively, it could be that households which are ex-ante likely to continue to additional conservation challenges are also those households that achieve large reductions in energy use.

Estimates $\hat{\beta}_\tau$ in Figure 1.7 for households that re-enroll are the average program effect including both households ending their participation after their second conservation challenge, and those that re-enroll in a third challenge. Figure 1.8 separates these households and plots estimates $\hat{\beta}_\tau$ for those undertaking only two conservation challenges against those that re-enroll in a third challenge. Both these household groups have similar conservation during their first challenge. However, during their second conservation challenge those households that end their participation show a rebound in their use over the months of the challenge. Their electricity use returns close to the pre-program levels as the challenge ends. In comparison, households that continue to a third conservation challenge continue to decrease their electricity use during their second challenge.

Figure 1.9 shows this pattern repeats again between households ending program participation after the third challenge and those re-enrolling in a fourth challenge. Electricity use remains similar across both groups of households until the last challenge, it continues to decrease among those who re-enroll and rebounds among those who leave the program. This rebound does not return to the pre-program use; electricity use remains persistently lower than pre-program levels by approximately 4%. In addition, the rebound occurs during the months leading up to the end of their final conservation challenge instead of after the challenge has completed. This is consistent with households stoping their conservation effort, or "giving up," prior to the end of the challenge.

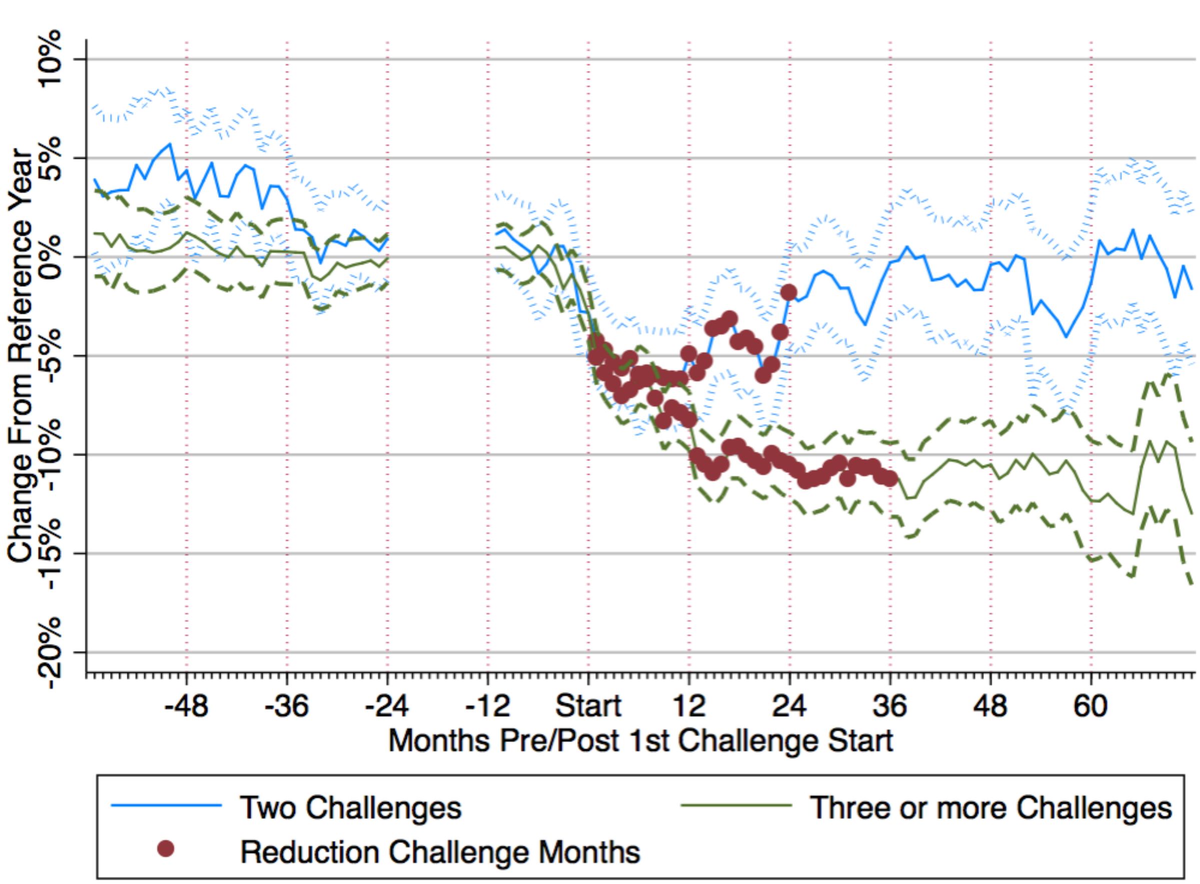
The reductions in electricity use are not driven by a particular heating type. Figure 1.10 plots estimated changes for primarily Electric and primarily non-Electric heating households. These have similar reductions despite the higher average use among electric heating households. As discussed further below, this similarity suggests that reductions are not due primarily to heating-related changes such as a lower thermostat set point or improved insolation.

Figure 1.7: **Single Challenge vs. Two Or More Challenges**



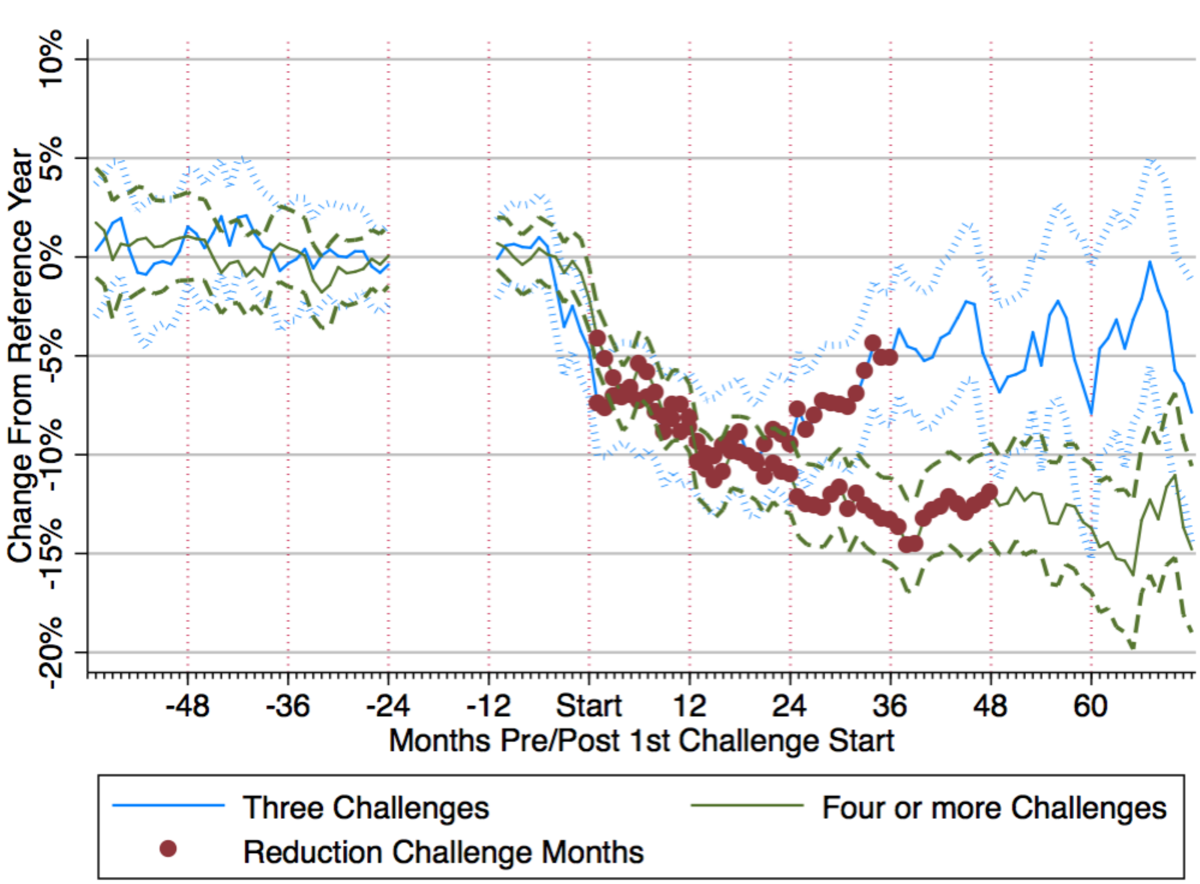
Notes: This figure plots estimated program effects $\hat{\beta}_\tau$ and 95% confidence intervals from equation (1.3) estimated for two mutually exclusive groups of households. Individual monthly estimates are presented in Table D.2 in the Appendix and plotted here as a line for clarity. Estimates in blue are households that undertake a single challenge and then end their program participation. Event study estimates in green are for households that undertake at least two conservation challenges and continue to their second challenge within 12 months of completing their initial challenge. Not shown are estimates θ_g for electricity use during the gap between the first and second challenges. Months 13-24 are estimates of the average change in electricity use among households in their second conservation challenge independent of any gap between challenges. Estimates include individual and date fixed effects and I cluster standard errors at the household level.

Figure 1.8: **Two Challenges vs. Three Or More Challenges**



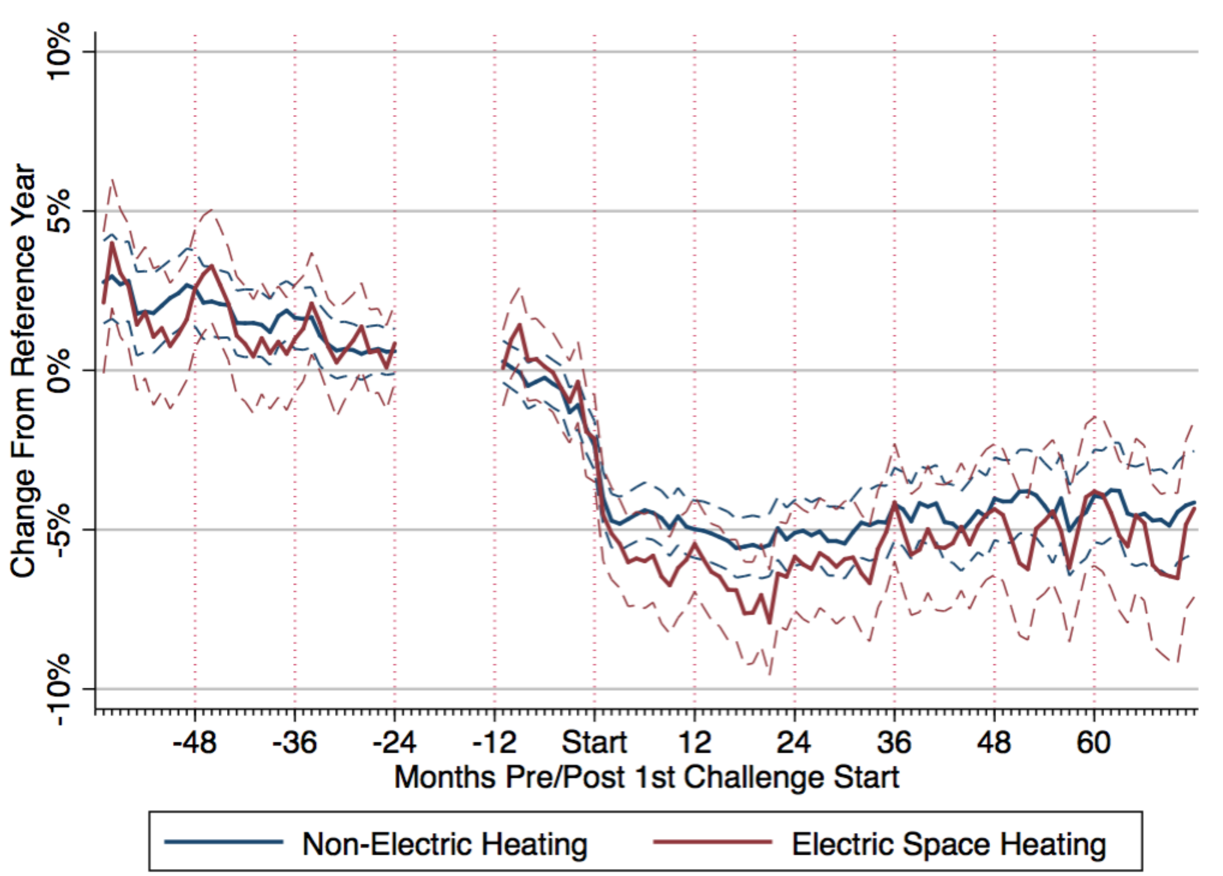
Notes: This figure plots estimated program effects $\hat{\beta}_\tau$ and 95% confidence intervals from equation (1.3) estimated for two mutually exclusive groups of households. Estimates are presented in Table D.2 in the Appendix. Estimates in blue are households that end their participation after a second conservation challenge. Estimates in green are households that continue to a third conservation challenge. Estimation sample restricted to households that continue to subsequent challenges within 12 months. Not shown are estimates θ_g for electricity use during the gap between challenges. Estimates include individual and date fixed effects and I cluster standard errors at the household level.

Figure 1.9: **Three Challenges vs. Four Or More Challenges**



Notes: This figure plots estimated program effects $\hat{\beta}_\tau$ and 95% confidence intervals from equation (1.3) estimated for two mutually exclusive groups of households. Estimates are presented in Table D.2 in the Appendix. Estimates in blue are households that end their participation after a third conservation challenge. Estimates in green are households that continue to a fourth conservation challenge. Estimation sample restricted to households that continue to subsequent challenges within 12 months. Not shown are estimates θ_g for electricity use during the gap between challenges. Estimates include individual and date fixed effects and I cluster standard errors at the household level.

Figure 1.10: **Estimated Treatment Effects By Heating Type**



Notes: This figure plots estimates of $\hat{\beta}_\tau$ and 95% confidence intervals from specification (1.3) for all participant and non-participant households by heating type. Non-Electric Heating households (blue line) are those that do not primarily heat with electricity while Electric Space Heating households use electricity as the primary heat source. Estimates $\hat{\beta}_\tau$ are ordered by event-time τ. Point estimates in red denote the 12 months of the initial conservation challenge ($\tau = [1..12]$). The pre-treatment period is denoted by the months prior to *Start* ($\tau \leq 0$). The visual gap in estimates between months $\tau = -24$ and $\tau = -36$ is the excluded reference period. $\hat{\beta}_\tau$ identify the percent change in electricity use relative to the average electricity use within a household during this excluded reference year.

1.4.2 Seasonal Treatment Effects

Electricity conservation may vary by season depending on how households reduce their electricity use. Improved insulation, smart thermostats, or reductions in the household temperature will produce larger energy savings in the winter and among households that heat primarily with electricity. More efficient dryers, lightbulbs, or other changes that affect primarily non-seasonal electricity use will generate energy savings year round for both household heating types. Comparing when during the year that reductions in electricity use occur, and between household heating types, sheds light on how households have responded to the conservation challenge.

I estimate the event study model (1.1) separately for each calendar month and household heating type, and use the year immediately before the initial conservation challenge as the reference baseline to estimate the reductions in electricity BC Hydro credits to households. In Figure 1.11, I plot the estimated program effects for each month of year during the initial challenge. Both household heating types have similar reductions in the summer months, while Electric Space Heating households have larger reductions in the winter months. To estimate the fraction of reductions due to heating, I use the reductions in electricity use over the four warmest summer months as a measure of changes in non-heating use. Comparing this to the reductions over the remainder of the year finds that 13% of electricity conservation among non-electric heating households is related to heating, while 48% of the conservation among electric heating households is due to changes that affect heating. That non-electric heating households have reductions due to heating is not unexpected; non-electric heating households may still use after-market baseboard heaters in addition to their non-electric primary heat source.

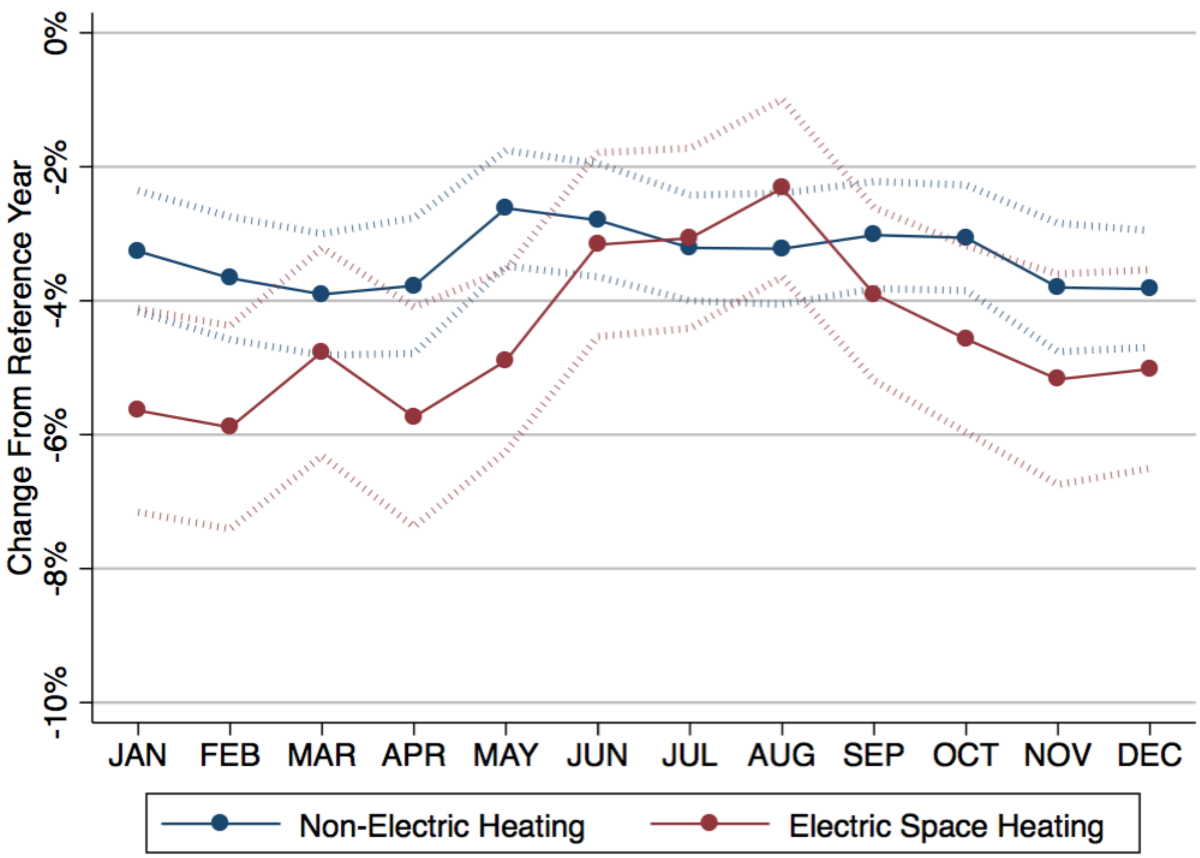
1.4.3 Program Effects By Household Characteristics

In Table 1.4, I estimate the program effects over the initial challenge for three household characteristics as well as pre-determined electricity use. To estimate the average change during a conservation challenge, I use an event study model with annual event-time indicators,

$$y_{it} = \sum_{\Upsilon=-4..-1,1..5} \theta_\Upsilon D_{i,t,\Upsilon} + \alpha_i + d_t + Pre_{it} + Post_{it} + \epsilon_{it} \tag{1.4}$$

where y_{it} is log monthly electricity use for household i at date t, and $D_{i,t,\Upsilon}$ is an indicator for if household i in (monthly) date t is in year Υ pre or post the challenge start date. α_i and d_t are indicator and date fixed effects. Pre_{it} and $Post_{it}$ are indicators for households outside the ± 5 year window. θ_1 is the average change during the initial conservation challenge, relative to the average during the reference year. I use the last pre-program year as the reference year, $\Upsilon = 0$. In panel A, I find that treatment effects in percentage terms are not statistically different across quartiles of pre-program use. This shows that reductions in absolute electricity use are larger for high-consumption household, but are not statistically different in percentage terms. In panels B, C, and D I find no statistically significant differences in treatment effects across building types, quartiles of assessed value, or quartiles of floor area. If higher assessed property values and larger floorspace is taken as a proxy for income, the results imply that households respond similarly across quartiles of wealth.

Figure 1.11: **Seasonal Treatment Effects**



Notes: This figure plots estimated changes in electricity use during the initial conservation challenge relative to the same month in the year prior. Estimates are from equation (1.1) estimated separately for each month of the year. 95% confidence intervals shown by the dashed lines. *Non-Electric Heating* are households expected by BC Hydro to heat primarily from sources other than electricity. *Electric Space Heating* are households whose primary source of heating is expected to be from electricity.

Table 1.4: **Treatment Effects by Pre-Determined Variables**

Panel A: Quartiles of Pre-Program Electricity Use				
	1st	2nd	3rd	4th
$\theta_1 : Initial\, Challenge$	-0.0440***	-0.0488***	-0.0425***	-0.0557***
	(0.00749)	(0.00546)	(0.00595)	(0.00699)
Avg. Use in 2006 (kWh)	393	701	1007	1642
Panel B: Building Type				
	1 Sty SFD	2 Sty SFD	Apartment	Townhouse
$\theta_1 : Initial\, Challenge$	-0.0359***	-0.0299***	-0.0525***	-0.0312***
	(0.00335)	(0.00381)	(0.00737)	(0.00581)
Panel C: Quartiles of Assessed Value				
	1st	2nd	3rd	4th
$\theta_1 : Initial\, Challenge$	-0.0388***	-0.0410***	-0.0327***	-0.0303***
	(0.00454)	(0.00420)	(0.00414)	(0.00448)
Avg. Assessed Value ($1,000)	$289	$450	$684	$1,208
Panel D: Quartiles of Floor Area				
	1st	2nd	3rd	4th
$\theta_1 : Initial\, Challenge$	-0.0463***	-0.0402***	-0.0281***	-0.0278***
	(0.00517)	(0.00398)	(0.00404)	(0.00452)
Avg. Floor Area (sq. ft.)	979	1642	2194	3254

Notes: Estimated average change in electricity use from the year pre-program to year of the initial conservation challenge. Panel A: Quartiles of pre-program electricity use determined from households' average electricity use in the pre-program year, 2006. Quartiles are defined separately for the balanced set of participant and non-participant households. Estimates exclude households starting their initial challenge before 2009 to avoid biasing estimates with a reversion to the mean. Panel B: Building type includes the four principal housing types of single story single family dwellings, two story single family dwellings, apartments, and town homes. Panel C: quartiles of assessed value are from the 2010 BC Assessment for individual units and include both structure and land value. Panel D: quartiles of floor area are in square feet. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

1.5 Conclusion

This chapter investigates how households respond to financial rewards for achieving electricity conservation targets. I work with a large electrical utility company to study a program that allows households to participate in successive annual electricity conservation challenges. Using a panel of monthly electricity use, I track households' decisions whether to re-enroll in the program along with changes in their electricity use. The ten years of the panel allow me to estimate the long-run persistence of electricity conservation as households leave the program, as well as the long-run changes among households that repeatedly re-enroll in the program. By comparing the electricity conservation across households I provide insights into what changes within the home households made to conserve electricity, and when conservation occurs. Three findings distinguish this work. First, this chapter undertakes one of the longest studies to date of the persistence of electricity conservation. I do not find that financial rewards cause persistent changes; instead, electricity use rebounds close to pre-program levels as households leave the program, and continues to decline among households that re-enroll. This shows that the ongoing incentive of successive financial rewards is necessary to cause long-run lower electricity use. A potential implication of this is that programs in fields outside energy conservation, such as education, that aim to cause persistent effects should consider incentivizing specific changes known to be permanent rather than reward a general outcome. Second, despite the programs voluntary nature I find no evidence that households strategically manipulate their participation to receive credit for conservation that is not due to the program. This supports that voluntary electricity conservation reward programs can be implemented and with a large majority of resulting conservation being causally due to the program. The lack of strategic manipulation is supported by the third finding of few differences in electricity conservation across household characteristics. This suggests that segmenting consumers by type is a less effective margin for improving similar conservation programs compared to increasing re-enrollment.

Chapter 2

The Extensive Margin of Electricity Conservation

2.1 Introduction

In Chapter 1, I documented how households may participate in multiple Team Power Smart conservation challenges, and that electricity use repeatedly diverges between those who leave the program and those that re-enroll in additional conservation challenges. This suggests that energy use may rebound in the absence of program participation and that voluntary re-enrollment in additional conservation challenges could cause additional reductions in electricity use. This chapter uses the same Team Power Smart program to study two aspects of these households' extensive margin re-enrollment decisions. First, I study the decision to re-enroll. I find that re-enrollment decisions differ little across household characteristics or their level of electricity use. Instead, I find that households' decisions on whether to remain in the program are sensitive to their success or failure in a conservation challenge but are insensitive to their actual effort and reductions in electricity use. As a result, households make the opposite decision from what the program design incentivizes; households facing an increased likelihood of achieving their next conservation target are less likely to re-enroll and attempt another conservation challenge. Importantly for designing general incentive programs this suggests that households use simple heuristics in making decisions rather than incorporate the detailed information that is readily available to them.

Second, comparisons of electricity conservation across multiple challenges in Chapter 1 compared households making different extensive margin re-enrollment decisions. A concern with voluntary re-enrollment is that a selection bias may arise if households differ in a way that determines post-challenge electricity conservation and is correlated with their decision to remain in the program. For example, households that make a capital investment in energy efficient appliances may, conditional on the reductions in energy use achieved in the first challenge, be more likely to continue to a subsequent challenge than a household which is making only short-run adjustments to behaviour. This potential self-selection bias prevents the event study estimates of Chapter 1 being interpreted as full program treatment effects. I address this self-selection by undertaking an instrumental variables estimate of the causal reductions in electricity use from continuing to a second conservation challenge. Households' success or failure in a conservation challenges creates a discontinuity in the probability they continue to an additional challenge. Using this in a fuzzy regression discontinuity empirical strategy I find that an additional conservation challenge causes lower electricity use. This is consistent with the event study results of Chapter 1 and shows that voluntary financial reward programs can cause reductions in electricity use, but that electricity use tends to rebound in the absence of the financial incentive.

Information provision is a ubiquitous feature of energy conservation and other incentive programs, both as a stand alone intervention and bundled with incentives such as price changes or financial rewards. I find that households' re-enrollment decisions are responsive to passing or failing their annual conservation challenge, but not to the detailed information on their actual conservation effort that is provided to them.[12] This suggests that while households are responding to information, they either do not take the full set of information into account or do not respond to it as a standard neoclassical model of consumer decisions would predict.[13] This is important for programs based on providing consumers with detailed information. For example, Schleich et al. (2017) undertakes a RCT where households can access detailed information on their consumption through an online portal or through mailed reports and finds this is effective in persistently reducing electricity use. Jessoe and Rapson (2014) use a RCT to find in-home displays on the price and quantity of electricity used increases consumers price-elasticity, while Martin and Rivers (2016) show how similar in-home displays can directly cause energy conservation.[14] This chapter's findings are also consistent with previous work that suggests the arrival of information, not only the information content itself, can serve as a nudge or reminder on electricity use. Gilbert and Graff Zivin (2014) show that simply receiving an electricity bill can reduce electricity use in the short term by 1%, and Sexton (2015) finds that increased billing frequency lowers electricity use. The responsiveness of households to their passing or failing a challenge but not the incentive structure is also suggestive that intrinsic motivations for conserving energy may be important (Gneezy et al., 2011). Previous work considering intrinsic motivations include the home energy reports studied in Allcott (2011) and Allcott and Rogers (2014), and Ito et al. (2015) who directly compare moral suasion against the incentive of short-term higher prices through two RCTs. They find that while both interventions cause lower short-run electricity use, only the short-term price increase generated persistent effects.

Information also plays an important role in programs that change a price schedule or offer a financial reward. These programs assume consumers are sufficiently informed of their own energy use and the program design to respond to the incentive as intended. Programs can fail to deliver anticipated outcomes if consumers under-respond to the incentive or do not use information available to them. Ito (2014) exploits a spatial discontinuity among households experiencing different block pricing schedules and finds that consumers respond to average prices instead of marginal prices. He concludes that the nonlinear price schedule is unsuccessful in achieving energy conservation due to this suboptimal response. The use of heuristics and inattention to information found in this chapter, whether rational or a sub-optimal response to the incentive, may be a pervasive feature of consumers' responses and is

[12]This information is provided to participants both through a letter at the end of their conservation challenge and through an online portal. The majority of households regularly log into their online portal (Kassirer et al., 2014) so this their lack of response is not not due to a lack of awareness of their progress or success.

[13]An important model of decision making in Psychology is self-efficacy, which refers to an individuals beliefs about their ability to affect outcomes through their actions (Bandura, 1977). Households could observe their success, interpret this as evidence they can affect the outcome of electricity conservation, and decide to re-enroll due to their updated beliefs on their self-efficacy. However, the sharp discontinuity in this setting - and the information provided to households on their actual conservation achieved - precludes the concept of self-efficacy from alone explaining households' behaviour.

[14]There is no a-priori reason that additional information would lead to decreases rather than increases in consumption; Wichman (2017) find that switching from bi-monthly to monthly billing for water increased its use.

an important consideration in the design of incentive programs.

The remainder of Chapter 2 is organized as follows. In Section 2.2 I discuss the weather adjustment and how changes in electricity use are measured. Section 2.3 analyzes what determines the decision to re-enroll including the discontinuous effect of success. Section 2.4 introduces the fuzzy regression discontinuity empirical strategy and provides evidence the identifying assumptions hold. I present the estimates in 2.5, discuss the program cost-effectiveness in Section 2.6, and conclude in Section 2.7.

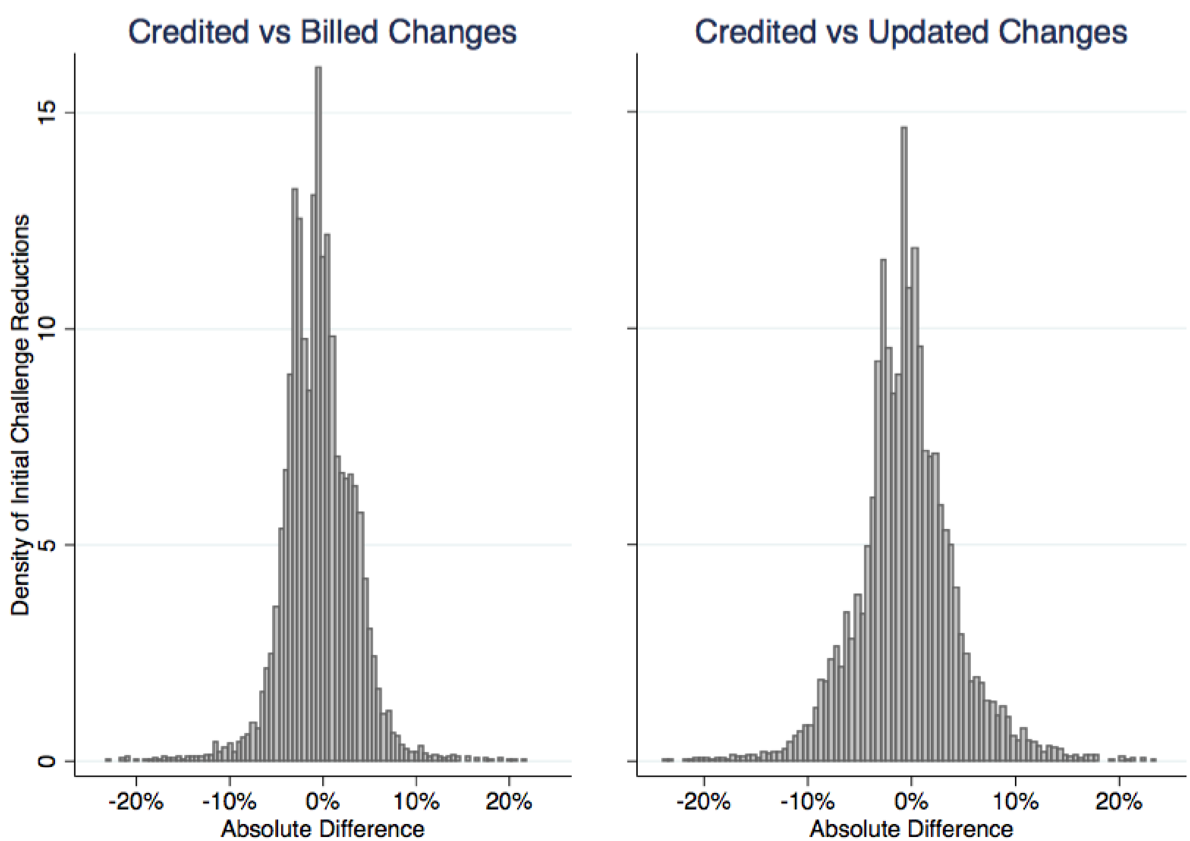
2.2 The Weather Adjustment

Weather changes affect households' electricity use. BC Hydro applies a weather-adjustment algorithm to avoid penalizing or unnecessarily rewarding households for changes in electricity use during a conservation challenge that are due to idiosyncratic changes in weather, and not their conservation effort. This algorithm is applied to the changes in actual electricity use that customers are billed for; I refer to changes in billed electricity use, prior to their weather-adjustment, as billed changes. The weather-adjustment algorithm adjusts billed changes for year-to-year changes in heating degree days, and results in a second measure which is the electricity conservation houses receive credit for; I refer to these as credited changes. Importantly, the weather-adjustment algorithm used by BC Hydro resulted in large adjustments to households' electricity conservation beyond those necessary for correcting for weather changes.[15] Figure 2.1, Credited vs. Billed changes, shows a histogram of the difference in absolute percentage points between credited and billed changes. These differences are not small; they have a mean of -0.43% and a standard deviation of 4.5%. In the second panel, Credited vs. Updated Changes, I show the difference between credited changes used to evaluate a household's success and changes using the updated algorithm where the effect of weather on electricity use has been removed as recommended by BC Hydro. These have a mean of -0.51% and standard deviation of 5.4%. This shows the adjustment caused many households to receive random shocks to their electricity conservation comparable in magnitude to half of their 10% conservation goal. When households view their online progress towards their 10% reduction goal, or the reductions in electricity use they are credited with achieving during a challenge, they are shown the credited — not billed — changes. As a result, households were not aware of the random shock applied to their conservation. As I discuss later, this random conservation shock significantly strengthens the fuzzy regression discontinuity identifying assumptions by mechanically randomizing households into and out of success in their conservation challenge.[16]

[15] Adjusting for weather is not an exact science. Some households heat with electricity more than others, some households that do not principally heat with electricity - and so are defined as non-electric heat households - still make significant use of electric heat via baseboard heaters, and household-specific characteristics like insulation or number of residents will drive large differences in the use electricity in response to weather changes. The weather-adjustment algorithm used to calculate credited changes was improved and updated in 2014; I exclude households that use the updated weather-adjustment algorithm. The weather-adjustment algorithm, formally available on the BC Hydro Team Power Smart website, is available from the author.

[16] In theory a household could calculate their own reductions from their billed consumption, and from this, determine the weather correction applied. This is unlikely to have occurred for many households.

Figure 2.1: **Weather Adjustment Discrepancies**



Notes: The left panel, Credited vs. Billed changes, is a histogram of the absolute differences between the changes in electricity use credited to a household after applying the first weather-adjustment algorithm, and the changes in their billed electricity use. The right panel, Credited vs. Updated Changes, is the histogram of differences between credited changes and changes in billed electricity use where the effect of weather has been removed as recommended by BC Hydro and used in the updated algorithm. Differences are in absolute percentage points such that a 10% Absolute Difference is equivalent in magnitude to the 10% conservation target.

2.3 Re-Enrollment Decisions

All households that participate in Team Power Smart have the option of re-enrolling in additional conservation challenges. Do households differ in their likelihood of re-enrolling? To explore what correlates with their decision I estimate a Probit model for the probability of re-enrolling. Table 2.1 shows the marginal effects. Specification (1) includes households' electric heating category and building type. I use the most common household type, Single Story Single Family Dwellings that heat primarily without electricity, as the reference category; marginal effects show the change in probability of re-enrolling relative to this household type. Specification (1) shows Townhouses are the only household type with a statistically significant difference (4.75%) in the probability of re-enrolling. Specification (2) shows that the probability of re-enrolling does not materially differ across the number of bedrooms, household value, or size of the house. Specifications (3) and (4) control for, respectively, households credited changes in electricity use and an indicator *Success* for whether they achieve their conservation target. These show households with larger electricity conservation are more likely to re-enroll as are households that pass their conservation challenge. This highlights the importance of success in the challenge relative to differences across household types. Taking the largest difference in point estimates across household types in Specification (4) finds Townhouses are 8.6% more likely to re-enroll than homes classified Other. In comparison, households that pass their conservation Challenge are 19.8% more likely to re-enroll. This supports the findings from Subsection 1.4.3 that electricity conservation does not differ substantially across household types. Specification (5) includes a household's Pre-Program use measured in standard deviations from the mean of households' 2006 electricity use within heating and building type categories. This shows that households with higher Pre-Program electricity use are more likely to re-enroll; households three standard deviations above the mean 2006 electricity use are 5.9% more likely to re-enroll. However, this magnitude is not large compared to the effect of Success — Specification (6) — or differences between Townhomes and Other. Specification (7) shows the re-enrollment probability differs little between Winter and other seasons. Taken together these results show that a households pre-determined characteristics have little direct, or in-direct through the electricity conservation, effect on the probability of re-enrolling. This is particularly the case in comparison to the importance of passing the conservation challenge, which I explore in detail below.

2.3.1 By Level of Reductions

Figure 2.2 plots the probability of re-enrolling in a second challenge against the credited changes in electricity use from a household's initial conservation challenge. This provides several insights into households' decisions whether to re-enroll. Among households that fail their initial challenge, their probability of continuing to a second challenge is largely independent of their credited changes in electricity use. Figure 2.2 shows households with large increases in electricity use of around 20% have a similar probability of re-enrolling as households that had no change, and these households are approximately 7% less likely to re-enroll as those that nearly achieved their 10% target. A similar pattern repeats amongst households that passed their challenge; those that barely pass with reductions ∼10% are equally likely to continue as households that achieved reductions of ∼20% or more. In

Table 2.1: **Probit Model: Re-Enrolling in a Second Challenge**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable: Re-enrollment in a second conservation challenge							
Non-Electric Heat	-	-	-	-	-	-	
Electric Heat	0.00368	0.00112	0.00273	-0.00646	0.00335	-0.00682	
	(0.0148)	(0.0152)	(0.0149)	(0.0150)	(0.0149)	(0.0151)	
Heating Unknown	-0.0141	-0.00335	-0.0122	-0.00740	-0.0134	-0.00655	
	(0.0178)	(0.0185)	(0.0179)	(0.0178)	(0.0178)	(0.0178)	
1 Story Sfd	-	-	-	-	-	-	
2 Story Sfd	-0.00895	0.00879	-0.00873	-0.00661	-0.00915	-0.00684	
	(0.0147)	(0.0160)	(0.0147)	(0.0149)	(0.0147)	(0.0149)	
1.5 Story Sfd	-0.0161	0.000633	-0.0157	-0.0190	-0.0162	-0.0194	
	(0.0294)	(0.0301)	(0.0293)	(0.0297)	(0.0294)	(0.0298)	
Apartment	0.00710	-0.0438*	0.00725	0.00821	0.00796	0.00897	
	(0.0199)	(0.0263)	(0.0201)	(0.0201)	(0.0200)	(0.0202)	
Townhouse	0.0475**	0.0215	0.0510***	0.0569***	0.0476**	0.0572***	
	(0.0187)	(0.0206)	(0.0187)	(0.0187)	(0.0187)	(0.0187)	
Other (home type)	-0.0361	-0.0501	-0.0334	-0.0294	-0.0362	-0.0295	
	(0.0350)	(0.0360)	(0.0350)	(0.0350)	(0.0351)	(0.0350)	
Bedrooms		-0.0117					
		(0.00739)					
Value		-0.0195					
		(0.0132)					
Floor Area		-0.0260					
		(0.0239)					
Credited Changes			-0.259***				
			(0.0470)				
Success				0.198***		0.200***	
				(0.0114)		(0.0114)	
Pre-Program Use					-0.0195***	-0.0236***	
					(0.00567)	(0.00570)	
Winter							-
Spring							0.0357*
							(0.0184)
Summer							0.0301*
							(0.0165)
Fall							-0.0207
							(0.0169)
Households	7182	6880	7182	7182	7181	7181	7182
Pseudo R^2	0.001	0.003	0.010	0.030	0.003	0.031	0.002
χ^2	12.892	23.482	44.942	273.666	24.662	288.620	16.978

Notes: This table shows how differences in household characteristics affect the probability of re-enrolling in a second conservation challenge. Dependent variable: indicator $C_i = 1$ if household i re-enrolls, 0 otherwise. All households in this sample begin an initial conservation challenge between February 2007 and February 2013. Estimates for specifications (1) - (6) are relative to the reference category of One Story Single Family Dwellings that are primarily non-electric heating. Specification (7) is relative to Winter. Value and Floor area are natural logs, Credited Changes is the percent change in Challenge 1 electricity conservation credited to households, and Success an indicator equal to 1 if a household achieves their Challenge 1 conservation target. Pre-program use is the number of standard deviations between a household's electricity use in 2006 and the average electricity use among households within the same building and heating type category. All coefficients are marginal effects at the covariate means. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

contrast, there is a sharp discontinuous jump in the probability of continuing to a subsequent challenge at the 9.5% threshold for success.[17] This pattern shows that in deciding whether to continue in the program, households are responsive to their success or failure in a conservation challenge but are largely insensitive to the level of reductions in electricity use they are credited with or achieve.[18]

This insensitivity of re-enrollment to changes in electricity use is not consistent with the incentive structure of the program. Each conservation challenge is an additional 10% reduction compared to the year prior; this creates a greater incentive for households that had increases, or smaller reductions, in electricity use to re-enroll compared to if they had larger reductions. Households with large increases during a challenge establish a new higher baseline for their subsequent 10% conservation challenge, while those that decreased their electricity use establish a new lower baseline. Under the reasonable assumption that the marginal cost of electricity conservation is increasing, the greater the electricity conservation achieved in a challenge the more costly, in terms of effort or financial investment, the subsequent challenge will be.[19][20]

2.4 The Fuzzy Regression Discontinuity Empirical Strategy

The simplest way to identify the causal treatment effect of a second challenge would be to randomly assign households that complete the first challenge into and out of the second challenge. Random assignment would permit causal identification by making households ending their participation a valid counterfactual for those which were randomly re-enrolled. In the absence of random assignment, a discontinuity in the probability of treatment can, under several identifying assumptions, produce random assignment into treatment for a subset of households. This permits causal identification for that subset despite potential self-selection, and is the basis of the fuzzy regression discontinuity design (fuzzy-RDD) employed here (Angrist and Pischke, 2008).

A fuzzy-RDD requires a threshold in a running variable that, as the running variable increases from just below the threshold to just above the threshold, causes a discontinuous change in the probability of treatment. If units cannot precisely manipulate the running variable then, at the threshold, they will be as good as randomly assigned to either side of it (Lee and Lemieux, 2010). As a result, some units are randomly assigned into and out of treatment by their random assignment to either side of the discontinuity. Those units that 'comply' with the running variable by changing their treatment status,

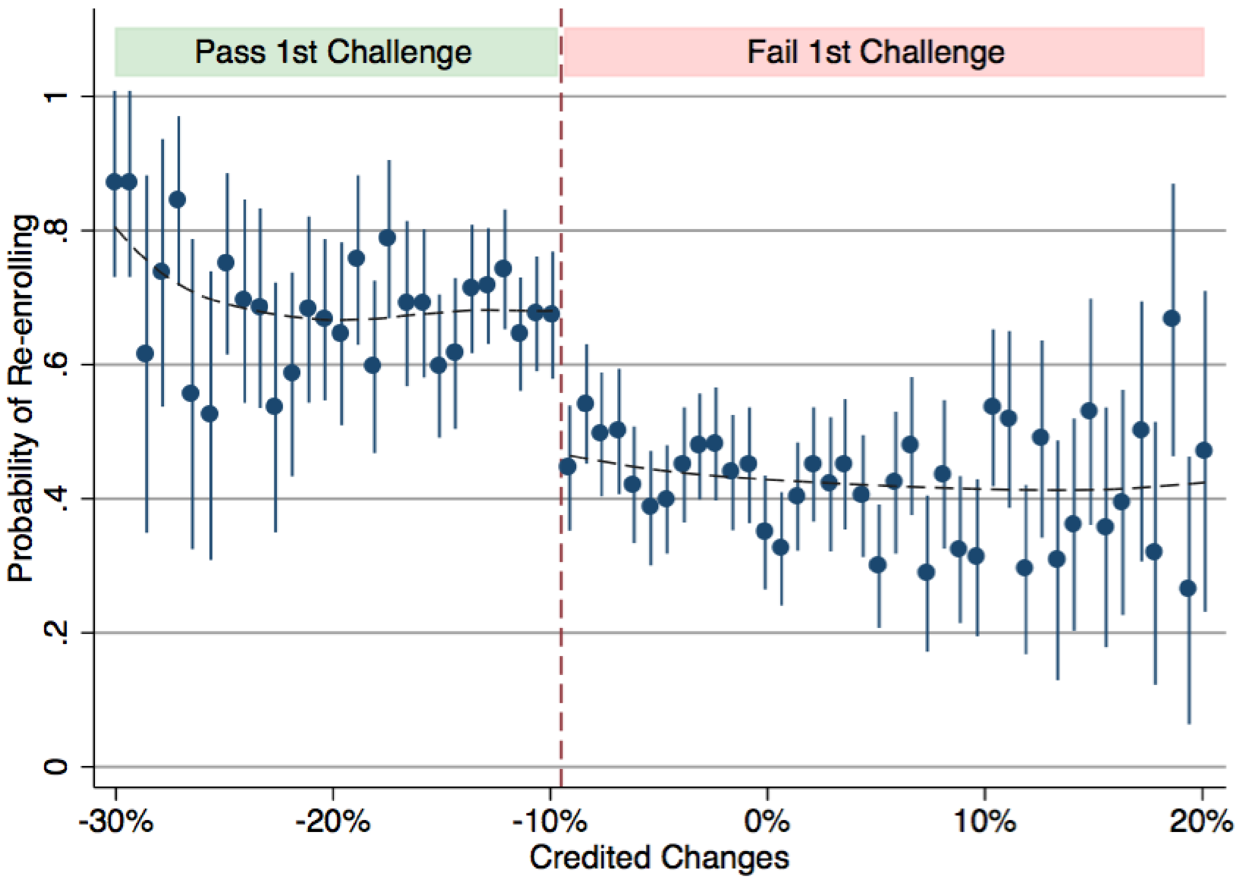
[17] While similar discontinuities exist after the second and subsequent conservation challenges, the sample of these households is too small for instrumental variable estimates.

[18] See Figure B.1 in the Appendix for the comparison to billed reductions.

[19] This is supported by Figure B.1 in the Appendix, which shows the fraction of households that, conditional on re-enrolling, pass their second conservation challenge. Consistent with increasing marginal cost of electricity conservation described above, approximately half of households with billed increases in electricity use of 10-20% pass their second challenge; these households are twice as likely to pass as households that had reductions during the first challenge of 10% or larger. This finding also holds when using credited changes.

[20] A potential concern is that households may differ substantially depending on their credited changes in a way that affects their probability of re-enrolling, and this could offset the incentive to re-enroll caused by smaller reductions. The large weather adjustment makes such an offset unlikely to produce the observed insensitivity of re-enrollment to conservation. The weather adjustment significantly randomizes households achieving the same actual reduction in electricity use across the credited changes, increasing the comparability of households with different credited changes.

Figure 2.2: **Probability of Re-Enrolling**



Notes: Credited changes are the annual changes in electricity consumption between the pre-program baseline year and the year of the first conservation challenge after processing by the BC Hydro weather adjustment algorithm. Credited changes are those changes shown to households during their conservation challenges and after their completion. Point estimates and 95% confidence intervals are the mean probability of re-enrolling among households within 0.75% width binds in credited changes. The dashed line is a first order local polynomial fit. The vertical dashed line indicates the 9.5% threshold. By definition, households to the left of the dashed line pass their conservation challenge while those to the right fail.

based on their position above or below the threshold, are called 'compliers.' This difference in compliers changing their treatment status across the threshold changes, through the compliers treatment effect, the outcome variable across the threshold. The fuzzy-RD recovers the Local Average Treatment Effect (LATE) for these compliers by dividing the average change in outcome variable by the average change in the number of compliers. This is a *Local* Average Treatment Effect because it is estimated only for those units with the running variable in the vicinity of the threshold.[21] The fuzzy-RD estimate of the LATE can be found by dividing two regression discontinuity estimates; the numerator is the RD estimate of the change in outcome across the threshold, and the denominator is the RD estimate of the change in probability of treatment across the threshold. As Angrist and Pischke (2008) discuss, this can be interpreted as a two-stage least squares instrumental variables estimation where the instrument for treatment is a binary indicator for a unit being above or below the threshold.

I use the discontinuity in the probability of continuing to a second challenge at the 9.5% conservation threshold for success, shown in Figure 2.2, as the instrument for treatment in a second conservation challenge. The instrumental variable is a binary indicator for success in the initial conservation challenge. The first stage relationship is

$$C_i = \gamma_0 + \gamma_1 1\{R_i \leq \bar{R}\} + \gamma_2 R_i + \gamma_3 1\{R_i \leq \bar{R}\} \times R_i + \gamma_4 B_i + \gamma_5 X_i + \eta_i \tag{2.1}$$

where C_i is a binary indicator for whether a household continues to a second challenge, R_i are households' credited changes in electricity use from the first challenge, $\bar{R}$ is the threshold for success in the challenge and is -9.5%, $1\{R_i \leq \bar{R}\}$ is the dummy variable for success in the initial challenge, B_i are the billed changes from the initial challenge, and X_i is a vector of other controls. In my main specification I control for a linear trend in credited reductions and allow this trend to have different slopes on either side of the discontinuity. The instrument excluded from the second stage is $1\{R_i \leq \bar{R}\}$.

The second-stage relationship is

$$y_i = \beta_0 + \beta_1 C_i + \beta_2 R_i + \beta_3 1\{R_i \leq \bar{R}\} \times R_i + \beta_4 B_i + \beta_5 X_i + \epsilon_i \tag{2.2}$$

where y_i is the post-challenge percent change in electricity use.[22] y_i is defined

$$y_i \equiv \frac{(u_{i,\tau=2} - u_{i,\tau=1})}{u_{i,\tau=0}} \tag{2.3}$$

where $u_{i,\tau}$ is household i's aggregate electricity use during the year indexed by event-time τ. For households that do not undertake a second challenge $u_{i,\tau=2}$ is the total electricity use in the 12 months immediately following the completion of their initial challenge, $u_{i,\tau=1}$ is the total electricity use during their initial challenge, and $u_{i,\tau=0}$ is the use during the pre-challenge year. For households that immediately undertake a second conservation challenge with no gap between challenges $u_{i,\tau=2}$ is the total electricity use during the second challenge and $u_{i,\tau=1}$ and $u_{i,\tau=0}$ are as before. For households that

[21] In contrast to an Average Treatment Effect which would be the average treatment effect across all units, regardless of their value of the running variable.

[22] In the Appendix Section B.3 I present an alternative fuzzy-RD specification that uses the log of monthly electricity use. This finds broadly similar results to using post-challenge percent changes in electricity use. My preferred results use post-challenge changes due to its parsimonious specification and easy interpretation of estimated coefficients.

wait before beginning a second conservation challenge I define $u_{i,\tau=2}$ as the 12 months of electricity use during their second challenge and $u_{i,\tau=1}$ as the 12 months of electricity use immediately preceding that second challenge. This makes y_i a consistent measure of the reductions in electricity use a household is trying to achieve in its second challenge regardless of whether that household waited before undertaking a challenge or began it immediately. I center the billed and credited changes at the 9.5% threshold. β_0 is the post-challenge change in billed electricity use at this threshold for households that do not continue in the program. β_1 is the additional effect on post-challenge billed changes in electricity use relative to households that left the program.

2.4.1 First Stage and Reduced Form

Figure 2.3 plots the probability that households re-enroll in a second conservation challenge by their credited changes in the first challenge.[23] The solid vertical line shows the 10% target, and the dashed vertical shows the 9.5% threshold for success or failure. Importantly, the discontinuity occurs at the 9.5% threshold for determining success or failure, and not at the 10% target that households are trying to achieve. Figure 2.4 plots y_i from equation 2.3 against the same bins of credited changes during a households' first conservation challenge. The discontinuity occurs again at the 9.5% threshold, not 10% conservation target.

2.4.2 Identifying Assumptions

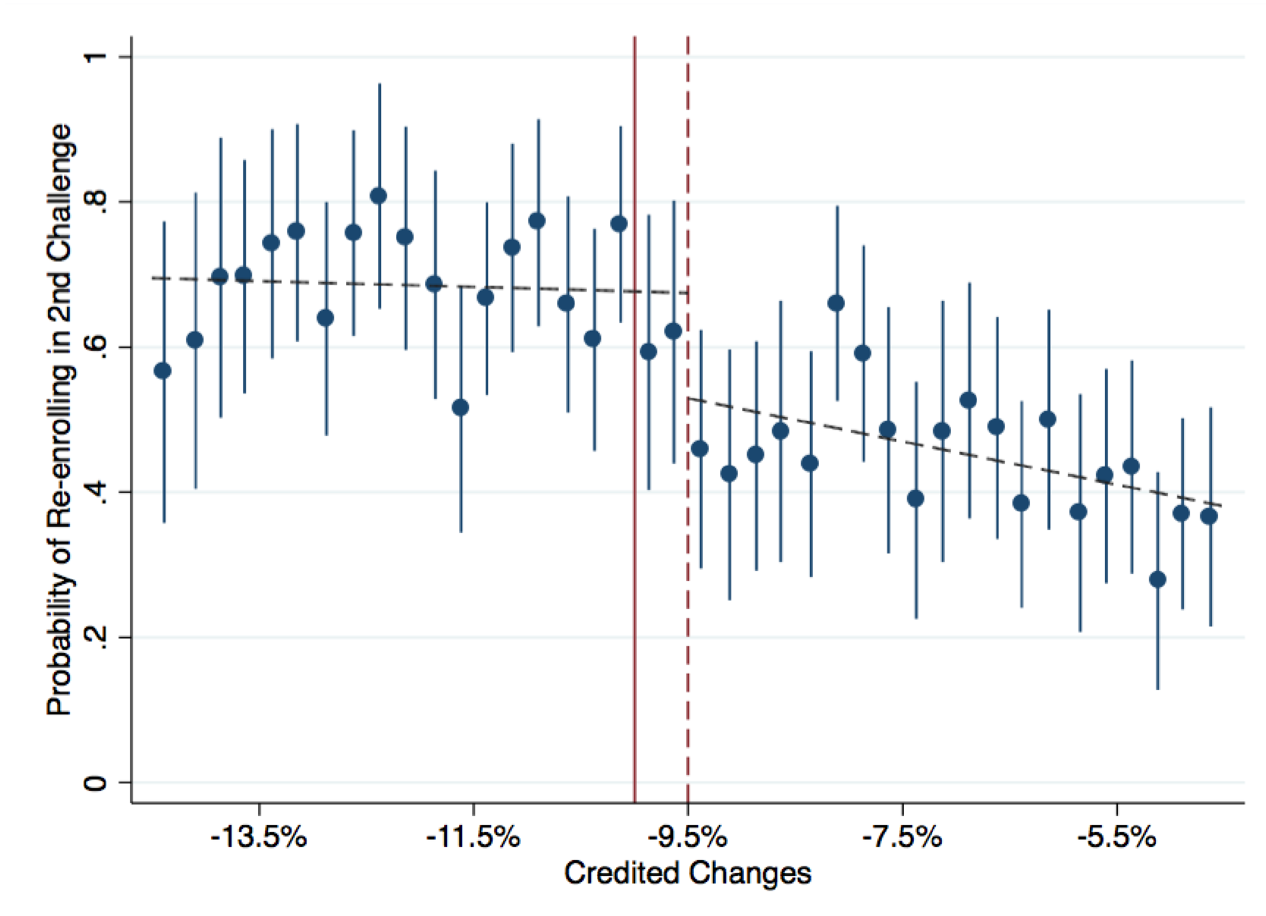
The fuzzy RD estimation strategy requires that assignment into and out of treatment in the vicinity of the threshold is as good as randomly assigned, such that households on one side of the threshold are a suitable counterfactual for households on the other. This assumption is violated if households can precisely manipulate their assignment into treatment. Such manipulation is a particular concern in this setup as households are explicitly trying to achieve a 10% conservation target. Sorting at the discontinuity could occur if, for example, households are heterogeneous in their attention to their progress and high-information type households exert additional effort in the last months of a conservation challenge and self-select into passing their challenge.

Sorting discontinuously at the threshold is unlikely for several reasons. Because the weather adjustment is applied each month and used to update households' cumulative progress towards their 10% conservation target, households that are attentive to their progress could in theory take the weather adjustment partially into account in updating their effort over the first 11 months.[24] However, households do not know their last month's conservation until they have completed the challenge and BC Hydro applies the final weather adjustment. This final weather adjustment mechanically randomizes households near the threshold into and out of treatment. I estimate that, within a $\pm 5\%$ window around the 10% conservation target, 17.3% of households that are on track to succeed at month 11 ultimately fail, while 14.7% of households that are on track to fail ultimately pass their challenge. The

[23] This is similar to Figure 2.2 but plotted on a narrower bandwidth around the 9.5% threshold.

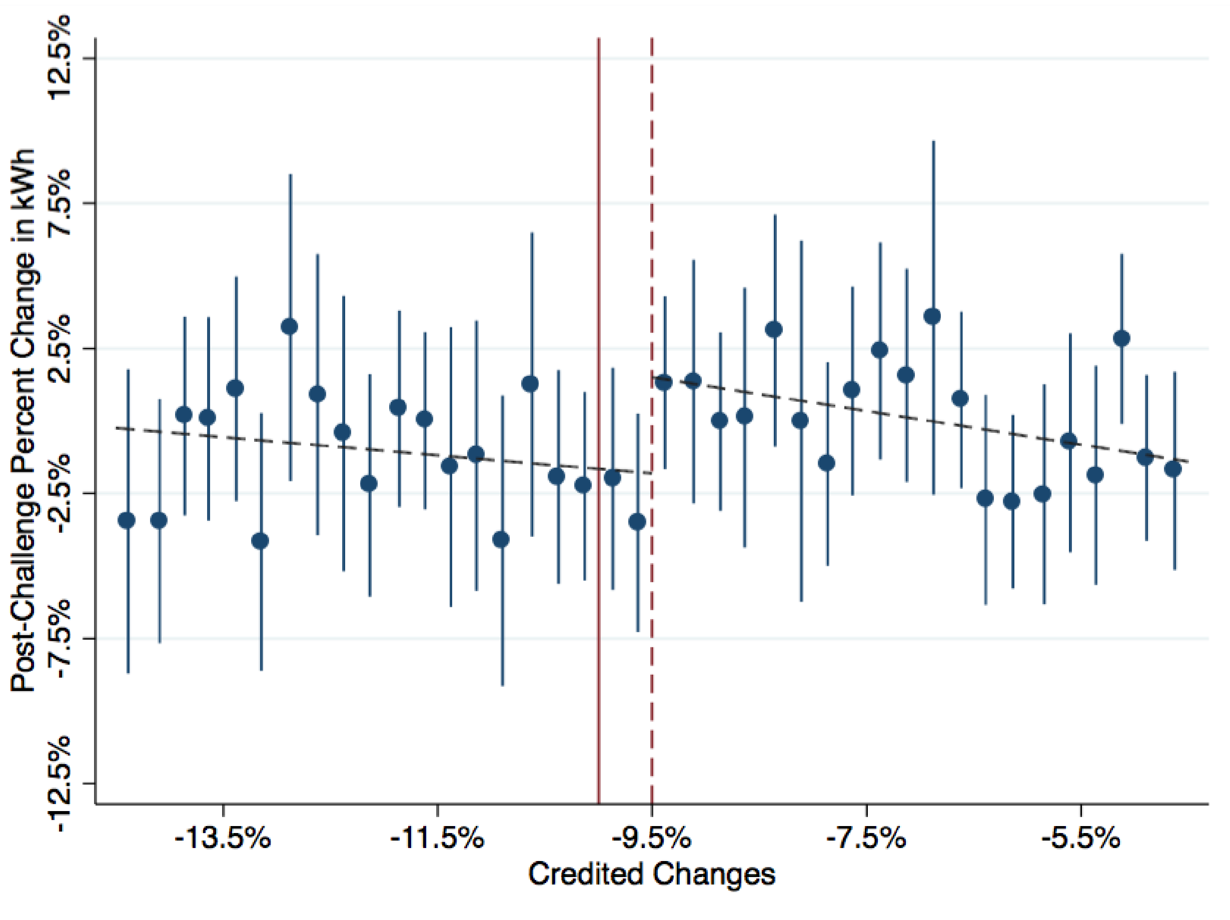
[24] In separate ongoing work I find no evidence that households just succeeding in their challenge update their next months conservation effort differently than those households which are just failing their challenge. This suggests that households are not precisely targeting the 10% conservation target.

Figure 2.3: **First Stage - Probability of Re-Enrolling**



Notes: Credited changes and the Probability of Re-enrolling are as defined in Figure 2.2. The sold vertical line is at the 10% conservation target and the dashed vertical line denotes the 9.5% threshold for success in a Challenge.

Figure 2.4: **Reduced Form - Post-Challenge kWh Changes**



Notes: Post-Challenge Percent Change in kWh denotes the percentage annual change in electricity use from the year of the conservation to the post-challenge year. The sold vertical line is at the 10% conservation target and the dashed vertical line denotes the 9.5% threshold for success in a Challenge.

cumulative nature of the challenge also makes precise manipulation difficult. A household at a 9% cumulative reduction entering the last month of their challenge would have to double their previous monthly reductions and reduce their use in the last month by 21% to achieve their 10% target.

Most importantly, households were not aware that their success or failure would be evaluated against a 9.5% threshold instead of the advertised 10% target. This does not remove the potential problem of sorting; households sorting around the 10% target in a way that changed potential outcomes in the absence of treatment would invalidate the causal interpretation of Regression Discontinuity estimates at the 9.5% threshold.[25] Instead, this feature of a separate threshold and target provides evidence that households are not sorting at the threshold on either observables or unobservables, despite bunching below the 10% target.

Figure 2.5 shows a histogram of credited changes during households' first conservation challenge. The mass of observations just below the 10% threshold suggests that households may be bunching around the 10% target. A density test by McCrary (2008), Figure 2.6, rejects the null hypothesis that there is no discontinuity in the density of the running variable at the 10% threshold, supporting that households are bunching at the 10% target.[26]

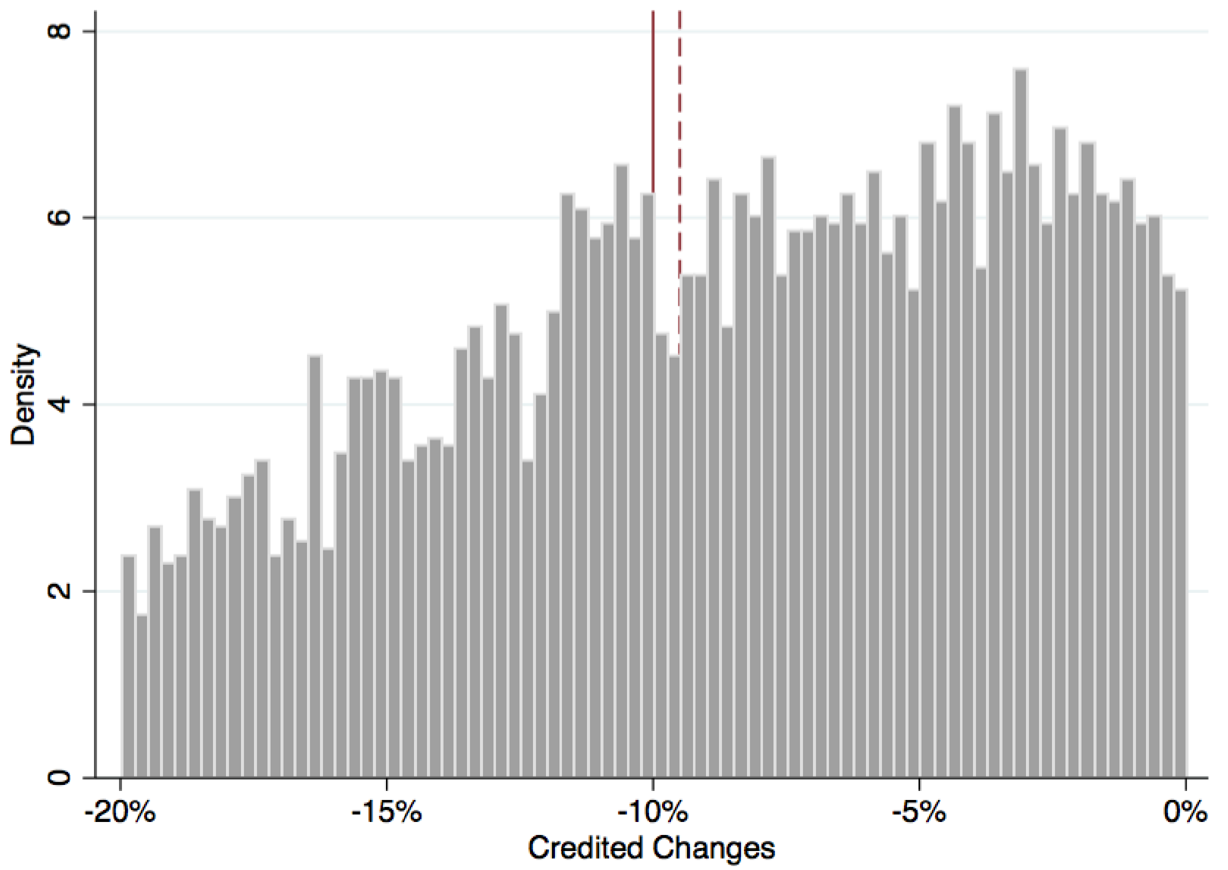
Evidence on whether there is sorting on observables can be gained by testing the continuity of covariates and pre-determined variables across the 9.5% threshold and 10% target. I find no statistically significant changes at either threshold which supports that households are not sorting at the discontinuity or 10% threshold in a manner correlated with observables (Appendix Table B.1 and Figure B.3). BC Hydro's separation of the 10% target from the 9.5% threshold for success also provides evidence that households are not sorting on unobservables. If households were sorting around the 10% target, their decision to re-enroll in a second challenge or the post-program outcome would be expected to be discontinuous at the 10% target. Figures 2.3 and 2.4 suggest that these outcomes change discontinuously only at the 9.5% threshold and not the 10% target.

The exclusion restriction requires that the instrument only affect the outcome, y_i, through the decision to continue to a second challenge (Angrist and Pischke, 2008). Conditional on credited and billed reductions, success in a challenge can have no direct effect on post-program changes in electricity use and therefore can be excluded from the second stage. This assumption could be violated if households receive a warm-glow effect from succeeding that affects their subsequent effort at reductions independent of continuing, or if the \$75 rebate causes an income effect and alters post-program conservation. As \$75 is small relative to household's incomes I assume there is no income effect that influences electricity use. Any warm-glow effect on subsequent conservation effort is likely to be short lived compared to the effect of the financial incentive, which remains throughout the twelve months of the challenge. In addition, if a warm glow effect was substantial it is likely to be particularly strong during the initial months of the next challenge while a household's success is still fresh in their minds. The event study estimates of Figure 1.7 indicate that additional program effects during the second challenge are consistent throughout the twelve months of challenge. This suggests that there is little

[25] In theory, a sufficiently large number of observations would allow estimation limited to bandwidths of $\pm 0.5\%$ around the 9.5% threshold, thus avoiding the problem of sorting around the 10% threshold.

[26] A density test by McCrary (2008), Appendix Figure B.2, fails to reject the null hypothesis at the 5% level (one sided p-value 0.093) that there is no discontinuity in the density of the running variable at the 9.5% threshold.

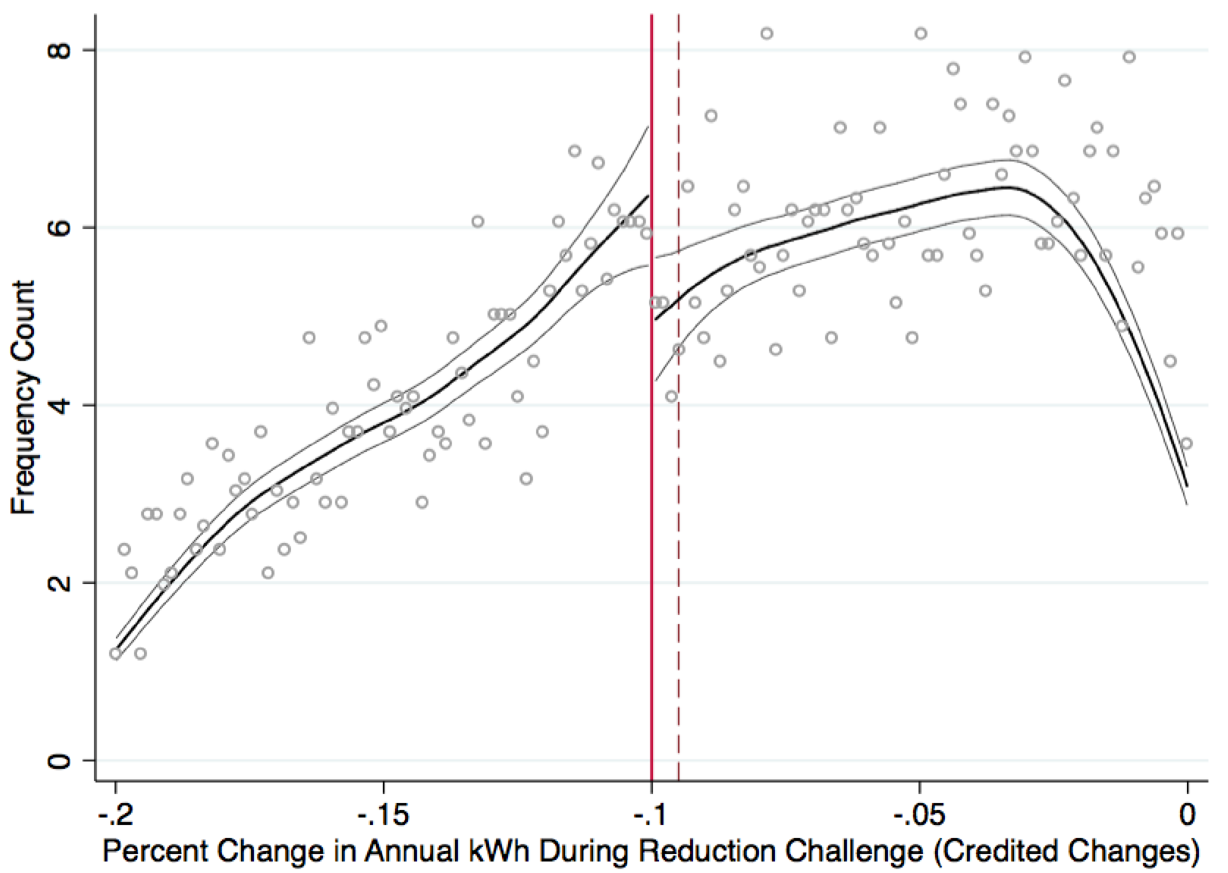
Figure 2.5: **Histogram of Credited Changes**



Notes: Histogram of households' credited changes during their initial conservation challenge. The sold red line is the 10% conservation target and the dashed line the 9.5% threshold for success. The increase in mass to the left of the vertical line demonstrates the potential for bunching at the 10% target.

warm-glow effect in the initial months, unless they are cancelled out by an equal and opposite increase in program effects from the financial incentive.

Figure 2.6: **Density Test of the Running Variable - 10% Target**



Notes: McCrary (2008) density test of the percent change in electricity use from a household's initial conservation challenge. The dark line is a smoothed local linear fit to the density of changes in electricity use, with 95% confidence intervals indicated by the light grey line. Point estimates of the density are grey circles. The dashed red line is the 9.5% reduction threshold, and the sold line is the 10% reduction target.

2.5 Fuzzy Regression Discontinuity Estimates

This section presents the fuzzy-RD estimates of the treatment effect of a second conservation challenge. Across a wide variety of specifications and robustness checks I find a consistent pattern where re-enrolling in a second conservation challenge causes a large additional reduction in electricity use. These results are consistent with the event-study results and support that additional conservation challenges cause additional reductions in electricity use, and that electricity use rebounds as households leave the program.

Table 2.2 presents my preferred specification. Columns (3) through (7) show results estimated for different bandwidths from ±7% to ±3% around the threshold of a -9.5% change in credited electricity use.[27,28] I restrict the estimation sample to households starting a second challenge within 12 months of completing their prior challenge.[29] Panel (A) shows the first-stage results for the probability of continuing to a second challenge, estimated from equation (2.1). For my preferred bandwidth of 5% I find that, conditional on failing the challenge, 53% of households continue to a second conservation challenge. At the 9.5% threshold for success, households that just succeed in their initial conservation challenge are 14.5% more likely to continue to an additional challenge than those which just failed. This pattern repeats across various estimation window widths. Approximately 50% of households re-enroll in a second conservation challenge if they fail their initial challenge, while success in a challenge causes an additional 14% to 20% of households to re-enroll.[30] Across bandwidths from ±7% to ±3% I find the F-statistic on the instrument decreases from 22 to 6.5. This indicates that the first-stage is reasonably strong for larger bandwidths, but the small sample size becomes relevant as the bandwidth narrows. The large and significant estimate on the discontinuity, $\hat{\gamma}_1$, and the generally small and insignificant estimate on billed changes, $\hat{\gamma}_4$, shows that the significant correlation between billed changes in electricity use and the probability of continuing, visible in Appendix Figure B.1, is a composition effect from the weather adjustment changing the probability of success near the threshold.

Table 2.2 panel (B) reports the OLS and second-stage instrumental variable estimates of equation (2.2). Specification (1) is the OLS result for all households. (1) shows that re-enrolling in an additional challenge is associated with a 1.6% decline in post-challenge electricity use, relative to households that do not re-enroll. Specification (2) is the OLS results for households within ±5% of the threshold. This finds that, for households at the 9.5% threshold, ending participation is associated with a rebound of 1.5% and re-enrolling with a reduction of 2.4%. Specifications (3) through (7) show the IV estimates for different estimation bandwidths. These estimates find a consistent pattern where, for households that comply with the instrument, continuing to a second conservation challenge causes a reduction in electricity use. For my preferred bandwidth of 5%, I estimate that continuing to a second conservation challenge causes a 23% reduction in electricity use ($\hat{\beta}_1 = -0.231$). This is a large effect. By definition,

[27] I present estimates using a range of bandwidths and a uniform weighting instead of kernel estimates. See Imbens and Lemieux (2007) for a discussion on the practical similarities of varying the bandwidth to using different kernels. I find results are robust across different bandwidths to using a triangular kernel.

[28] Plots of the First Stage and Reduced Form are presented in Table B.4 in the Appendix.

[29] Results are robust to other restrictions on the gap length between challenges. Table B.3 in the Appendix presents one such robustness check using a gap length of 6 months.

[30] A weak instruments test by Moreira (2003) rejects (p-value 0.013, 5% bandwidth) that the binary indicator for Success (γ_1) is a weak instrument.

the treatment effect is the change in electricity use for complier households relative to what they would have had, had they not re-enrolled. An instrumental variables estimation strategy cannot identify the level of electricity use, in the absence of continuing, for these compliers. As a result, the treatment effect does not separately identify a potential rebound in electricity use among those that leave the program from additional reductions in electricity use beyond those achieved in the first conservation challenge.[31] Given the average reduction of 9.1% during the first challenge for households within the ±5% bandwidth, and the target of an additional 10% conservation, the estimated treatment effect is likely comprised of both a significant rebound in electricity use among complier households that end participation and a large additional reduction among complier households that re-enroll in another challenge.[32]

Specifications (3) through (7) show that the magnitude of the treatment effect of a second challenge is sensitive to the estimation window width. Increasing the estimation bandwidth trades off decreased variance from a larger sample size against increased potential bias from misspecification (Lee and Lemieux, 2010). The concern with a large bandwidth is that households farther from the threshold may differ from those close to the threshold. This is less of a concern in this context due to the large weather adjustment. As Figure 2.1 demonstrates, households with the same billed changes in electricity use receive a large shock to their credited changes. Households receiving significantly different signals on their credited changes in use will, unknown to the households, have exerted the same effort and caused the same physical reductions in electricity use. This randomizes households by their actual effort - as measured by billed changes correctly adjusted for weather - across credited changes and around the discontinuity.

This randomization by billed changes is also important as post-challenge changes in electricity use are correlated with changes during the challenge. For example, households that experienced an unusually warm winter and decreased their use, but made no other changes, would be expected to rebound the following year to their use conditional on the expected number of heating degree days. While the credited changes, $\hat{\beta}_2$, and billed changes, $\hat{\beta}_4$, are either only marginally significant at a 10% level or not significant, the negative sign on the point estimates is consistent with the expected pattern of households that have larger increases in an initial challenge having a larger post-challenge reduction, and households with larger initial challenge reductions having larger post-challenge rebounds in use.

As discussed previously, an identifying assumption of a fuzzy RD estimation strategy is that households just on either side of the discontinuity are as good as randomly assigned. Evidence that this assumption does not hold would be if the IV estimates were sensitive to the inclusion of additional covariates. Table 2.3 shows IV estimates controlling for detailed household characteristics and changes in heating degree days. I control for the percent change in heating degree days between both the pre-program and first conservation challenge years ($HDD_{0,1}$) and between the first conservation challenge

[31] β_0 cannot be interpreted as the average electricity use, conditional on other covariates, for complier households in the absence of a conservation challenge. For example, in the simplest instrumental variable setup of a constant and a single endogenous variable, $\hat{\beta}_0 \equiv \bar{y} - \hat{\beta}_1 \cdot \bar{x}$.

[32] In comparing the OLS and fuzzy-RD estimates in Table 2.2 panel (B), it is important to note that the OLS estimates are for all households within the estimation window while the fuzzy-RDD is a LATE for compliers. As a result, the difference in estimates may be partially due to compliers being responsive to the incentive as estimated, with always-taker's and never-taker's electricity conservation remaining largely unresponsive to the financial reward incentive.

Table 2.2: **Fuzzy Regression Discontinuity Estimates of a Second Challenge**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A – First Stage				
Dependent variable: Continue to a Second Challenge C_i							
Window			±7 %	±6 %	±5 %	±4 %	±3 %
γ_1: Success Ind.			0.202***	0.190***	0.145***	0.137**	0.173**
			(0.0432)	(0.0468)	(0.0516)	(0.0573)	(0.0677)
γ_2: Cred. Reduc.			-0.532	-1.391	-2.937**	-1.290	2.090
			(0.778)	(0.971)	(1.244)	(1.760)	(2.765)
γ_3 : Success×			1.166	2.491*	2.565	-0.778	-3.973
Cred. Reduc.			(1.104)	(1.383)	(1.809)	(2.447)	(3.906)
γ_4 : Billed Reduc.			-0.300	-0.368	-0.0474	-0.282	-0.511
			(0.324)	(0.344)	(0.366)	(0.404)	(0.491)
γ_0 : Constant			0.487***	0.508***	0.530***	0.510***	0.475***
			(0.0303)	(0.0330)	(0.0364)	(0.0409)	(0.0479)
F-statistic			21.95	16.41	7.882	5.668	6.485
			Panel B – Second Stage				
Dependent variable: Percent change in post-challenge electricity use							
	OLS		Instrumental Variable Estimates				
Window		±5 %	±7 %	±6 %	±5 %	±4 %	±3 %
β_1 : Re-Enroll	-0.0160***	-0.0242***	-0.125**	-0.178**	-0.231**	-0.323**	-0.183*
	(0.00422)	(0.00683)	(0.0605)	(0.0738)	(0.116)	(0.164)	(0.111)
β_2: Cred. Reduc.		-0.171	-0.412*	-0.643*	-1.185*	-1.108	0.785
		(0.245)	(0.241)	(0.354)	(0.654)	(0.828)	(0.654)
β_3 : Success×		0.431	0.303	0.375	0.867	-0.229	-1.732
Cred. Reduc.		(0.439)	(0.310)	(0.426)	(0.624)	(0.985)	(1.229)
β_4 : Billed Reduc.		-0.0858	-0.0461	-0.113	-0.101	-0.216	-0.222
		(0.0990)	(0.0917)	(0.103)	(0.121)	(0.168)	(0.161)
β_0 : Constant	-0.00773***	0.0146*	0.0712**	0.103**	0.138**	0.184*	0.0944
	(0.00289)	(0.00773)	(0.0356)	(0.0447)	(0.0698)	(0.0950)	(0.0625)
N	5432	1475	2050	1763	1475	1196	888

Notes: This table reports fuzzy-RD estimates corresponding to equations (2.1) and (2.2). Estimation sample is restricted to households that either start their next challenge within 12 months or do not undertake an additional challenge. Estimation window is restricted to ± the listed percent around the 9.5% threshold in credited changes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

and post-challenge year ($HDD_{1,2}$). Increases in heating degree days during the post-challenge period are positively correlated with post-challenge changes in electricity use. This is consistent with colder weather increasing the demand for electricity. The inclusion of these additional covariates has only a small effect on the estimated effect of a second conservation challenge and supports the identifying assumption that households are as good as randomly assigned at the discontinuity.

A potential concern with the weather adjustment is if households with the same credited changes differ substantially in billed changes in the vicinity of the discontinuity. If billed reductions affect the post-program outcomes, for example if households were to exhibit a strong reversion to the mean, then outliers in the weather adjustment could cause a violation of the good-as-randomly assigned assumption. Evidence that this is not a problem is gained by further restricting the estimation sample to households that had billed changes within ±5% of 9.5% in billed reductions. This excludes those households receiving large weather adjustments to their billed electricity use, in addition to the estimation bandwidth in credited reductions. Estimates, Table 2.4, are robust to this restriction. A potential concern with RD estimates is that the using observations away from the threshold increases the risk of biased estimates (Calonico et al., 2014). In Table 2.5 and Appendix Table B.2 I present bias-corrected estimates using 1st and 2nd order polynomial fits and the method of Calonico et al. (2014). Specification (8) presents the optimal bandwidth, 9%, determined using the variance-bias tradeoff method of Calonico et al. (2014). Bias-corrected estimates lose significance for small bandwidths in the 1st order estimates, and in all bandwidths narrower than 9% for 2nd order estimates. However, in all cases the point estimates stay a consistent sign and large magnitude. This suggests that while the magnitude of point estimates vary, the causal effect of an additional conservation challenge is a large additional reduction in electricity use for complier households — those households whose decision to re-enroll in Team Power Smart is affected by their success or failure in their prior conservation challenge.

Table 2.3: **Fuzzy Regression Discontinuity Estimates: Additional Covariates**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A – First Stage				
Dependent variable: Continue to a Second Challenge C_i							
Window			±7 %	±6 %	±5 %	±4 %	±3 %
γ_1: Success Ind.			0.202***	0.188***	0.143***	0.137**	0.169**
			(0.0432)	(0.0467)	(0.0515)	(0.0572)	(0.0675)
γ_2: Cred. Reduc.			-0.484	-1.335	-2.834**	-0.981	2.186
			(0.800)	(0.995)	(1.268)	(1.781)	(2.774)
γ_3 : Success×			1.130	2.467*	2.524	-0.979	-3.635
Cred. Reduc.			(1.108)	(1.385)	(1.807)	(2.437)	(3.892)
γ_4 : Billed Reduc.			-0.329	-0.417	-0.169	-0.475	-0.870
			(0.358)	(0.381)	(0.403)	(0.439)	(0.539)
$\mathrm{HD}D_{1,2}$			0.0867	-0.0662	0.142	-0.128	-0.145
			(0.360)	(0.386)	(0.398)	(0.429)	(0.482)
$\mathrm{HD}D_{0,1}$			-0.0805	0.0703	-0.00803	0.301	0.591
			(0.351)	(0.375)	(0.384)	(0.410)	(0.463)
γ_0 : Constant			0.462***	0.476***	0.505***	0.483***	0.468***
			(0.0342)	(0.0372)	(0.0407)	(0.0456)	(0.0531)
F-statistic			21.82	16.27	7.736	5.757	6.266
			Panel B – Second Stage				
Dependent variable: Percent change in post-challenge electricity use							
	OLS		Instrumental Variable Estimates				
Window		±5 %	±7 %	±6 %	±5 %	±4 %	±3 %
β_1 : Re-Enroll	-0.0159***	-0.0241***	-0.119**	-0.168**	-0.214*	-0.306*	-0.176
	(0.00421)	(0.00683)	(0.0600)	(0.0726)	(0.113)	(0.157)	(0.112)
β_2: Cred. Reduc.	0.352***	0.348***	0.310***	0.297**	0.376***	0.255	0.299*
	(0.0748)	(0.111)	(0.103)	(0.120)	(0.137)	(0.177)	(0.166)
β_3 : Success×	-0.105	-0.113	-0.142	-0.113	-0.122	0.0475	0.0459
Cred. Reduc.	(0.0698)	(0.102)	(0.0964)	(0.114)	(0.130)	(0.173)	(0.167)
β_4 : Billed Reduc.		-0.0507	-0.323	-0.510	-0.962	-0.853	0.906
		(0.250)	(0.245)	(0.354)	(0.630)	(0.772)	(0.646)
$\mathrm{HD}D_{1,2}$		0.432	0.320	0.368	0.826	-0.239	-1.543
		(0.438)	(0.308)	(0.421)	(0.601)	(0.956)	(1.207)
$\mathrm{HD}D_{0,1}$		-0.216**	-0.142	-0.222*	-0.249*	-0.418**	-0.442**
		(0.108)	(0.103)	(0.114)	(0.130)	(0.185)	(0.183)
β_0 : Constant	0.00460	0.0302***	0.0782**	0.105**	0.138**	0.185**	0.109*
	(0.00447)	(0.00977)	(0.0342)	(0.0421)	(0.0656)	(0.0872)	(0.0626)
N	5432	1475	2050	1763	1475	1196	888

Notes: This table reports fuzzy-RD estimates corresponding to equations (2.1) and (2.2). All specifications include building type and heating category fixed effects. $HDD_{0,1}$ and $HDD_{1,2}$ are, respectively, the percent change in heating degree days from the pre-program year to the initial challenge, and initial challenge to the post-program year. Estimation sample restricted to households that either start their next challenge within 12 months or do not undertake an additional challenge. Estimation window is restricted to ± the listed percent around the 9.5% threshold in credited changes. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table 2.4: **Fuzzy Regression Discontinuity Estimates: Restricted Billing**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A – First Stage				
Dependent variable: Continue to a Second Challenge C_i							
Window			±7 %	±6 %	±5 %	±4 %	±3 %
γ_1: Success Ind.			0.185***	0.181***	0.121**	0.121**	0.172**
			(0.0485)	(0.0516)	(0.0561)	(0.0614)	(0.0721)
γ_2: Cred. Reduc.			0.124	-0.0825	-2.105	-1.072	3.004
			(1.016)	(1.205)	(1.476)	(2.003)	(3.073)
γ_3 : Success× Cred. Reduc.			-0.697	-0.292	0.0330	-1.981	-5.156
			(1.412)	(1.679)	(2.104)	(2.740)	(4.291)
γ_4 : Billed Reduc.			0.177	0.0615	0.215	0.283	-0.163
			(0.612)	(0.625)	(0.643)	(0.691)	(0.762)
γ_0 : Constant			0.475***	0.482***	0.515***	0.500***	0.458***
			(0.0349)	(0.0371)	(0.0401)	(0.0444)	(0.0511)
F-stat			14.57	12.34	4.649	3.889	5.665
			Panel B – Second Stage				
Dependent variable: Percent change in post-challenge electricity use							
	OLS		Instrumental Variable Estimates				
Window		±5 %	±7 %	±6 %	±5 %	±4 %	±3 %
β_1 : Re-Enroll	-0.0160***	-0.0239***	-0.171**	-0.191**	-0.269*	-0.307	-0.118
	(0.00422)	(0.00769)	(0.0782)	(0.0848)	(0.159)	(0.188)	(0.103)
β_2: Cred. Reduc.		-0.282	-0.510*	-0.638*	-1.240*	-1.118	0.953
		(0.282)	(0.308)	(0.355)	(0.733)	(0.876)	(0.631)
β_3 : Success× Cred. Reduc.		0.385	-0.153	-0.00562	0.319	-0.386	-1.527
		(0.505)	(0.441)	(0.487)	(0.719)	(1.132)	(1.204)
β_4 : Billed Reduc.		0.192	0.283	0.203	0.238	0.252	0.0407
		(0.172)	(0.190)	(0.194)	(0.238)	(0.279)	(0.212)
β_0 : Constant	-0.00773***	0.00903	0.0886**	0.102**	0.150	0.166	0.0524
	(0.00289)	(0.00817)	(0.0448)	(0.0490)	(0.0916)	(0.106)	(0.0564)
N	5432	1147	1394	1291	1147	982	763

Notes: This table reports fuzzy-RD estimates corresponding to equations (2.1) and (2.2). Sample restricted to households with billed changes within ±5% of the 9.5% conservation target along with restricting households to those within the listed estimation window around the 9.5% threshold in credited reductions. Estimation sample restricted to households that either start their next challenge within 12 months or do not undertake an additional challenge. Estimation window is restricted to ± the listed percent around the 9.5% threshold in credited changes.

Table 2.5: **1st Order Bias-Corrected Fuzzy Regression Discontinuity Estimates**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Conventional	-0.187*	-0.236**	-0.261**	-0.239**	-0.201**	-0.173**	-0.164**	-0.164**
	(0.104)	(0.118)	(0.128)	(0.108)	(0.0858)	(0.0727)	(0.0668)	(0.0669)
Bias-corrected	-0.281***	-0.176	-0.194	-0.265**	-0.299***	-0.293***	-0.257***	-0.189***
	(0.104)	(0.118)	(0.128)	(0.108)	(0.0858)	(0.0727)	(0.0668)	(0.0669)
Robust	-0.281*	-0.176	-0.194	-0.265*	-0.299**	-0.293***	-0.257***	-0.189**
	(0.157)	(0.177)	(0.188)	(0.156)	(0.122)	(0.103)	(0.0936)	(0.0763)
Observations	888	1196	1475	1763	2050	2296	2543	2538
Order Poly. (p)	1	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2	2
BW Poly. (h)	3%	4%	5%	6%	7%	8%	9%	9%
BW Bias (b)	3%	4%	5%	6%	7%	8%	9%	17%
F-Conv.	6.8	7	7.1	9.2	12.9	16.5	19.4	19.3
F-Bias	6.8	11.2	12.9	10.4	8.9	9.4	12.1	17.7
F-Robust	3.1	4.9	5.5	4.5	3.9	4.1	5.3	13.5

Notes: This table reports fuzzy-RD estimates using the method of (Calonico et al., 2014). All specifications use 1st order local polynomial regressions using a triangular kernel and restricted to households that either start their next challenge within 12 months or do not undertake an additional challenge. Specifications (1) through (7) are for bandwidths BW Poly. (h) around the threshold. Specification (8) determines the optimal polynomial and bias-correction bandwidths to be 9% and 17%, respectively. Conventional, bias-corrected, and bias-corrected and robust F-stats on the 1st stage instrument respectively denoted by F-Conv., F-Bias, and F-Robust. The bias correction is 2nd order, local polynomial. Standard errors in parentheses. Conventional and Bias-corrected have conventional standard errors, Robust estimates use robust standard errors. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

2.6 Cost Effectiveness

The Team Power Smart program is designed to produce electricity generation capacity savings and reduce the expected future increase in demand for electricity. A full cost-benefit analysis of the TPS program is beyond the scope of this work. Instead, I provide a lower bound on the cost of avoided electricity generation caused by this financial reward program. I estimate a lower bound for two reasons. First, because the costs of administering and advertising the Team Power Smart reward program are confidential to BC Hydro, I consider only the cost of the $75 rebates rewarded to households, and leave aside the costs of administering the program.[33] Second, I make the assumption that the estimated electricity conservation from the event-study model, Figure 1.6, are treatment effects; any overestimate of the true treatment effect will bias upward the cost of avoided generation. This is a reasonable assumption given the findings of Sections 1.3.1 and 2.1. From the estimated electricity conservation and the average electricity use among participants, Table 1.1, I find that the average aggregate reduction in electricity use over the first six years after an initial challenge is 2.7 MWh per household. This is the average across all households and accounts for their decisions whether to re-enroll after each challenge. Taking into account households' average success in their conservation challenges, and the number continuing to additional challenges, the average aggregate rebate payment over the six years

[33]Because the program is administered online, variable costs excluding the rebate are likely to be negligible. Program fixed costs may not be insignificant relative to the cost of the rewarded rebates. One full time equivalent employee compensated at $70,000 per year for managing the program would add approximately $6 per challenge to the program. This is $20 per awarded rebate using the 30% success rate over the initial five challenges households undertake.

is \$53 per household. This finds an average cost of avoided generation of \$20/MWh. In comparison, participants paid an average retail price of \$96/MWh in 2016, while a large hydroelectric dam under construction in the province is estimated to have a levelized cost of electricity of \$34-\$83/MWh (British Columbia Utilities Commission, 2017). This makes the Team Power Smart program a cost-effective way to reduce the demand for electricity in comparison to the cost of new generation.

What is the cost of avoided carbon emissions due to this energy conservation program? It is important to note that this energy conservation program was not designed to principally reduce carbon emissions and the cost of avoided emissions is not particularly relevant to BC Hydro. British Columbia generates over 90% of its electricity from hydroelectric dams and has a low emissions intensity of 9kgCO2e/MWh (BCH, 2016); \$20 per avoided MWh is a cost of avoided greenhouse gas emissions of \$2,222/tCO2eq. However, BC Hydro engages in large cross-border trade in electricity with the United States, primarily California. Lower electricity use in British Columbia allows BC Hydro to sell relatively low cost and low emissions power to California. Assuming all reductions in B.C. electricity use reduces generation in California finds, using the 2017 California average emissions intensity (EPA, 2017), a cost of emissions abatement of \$71/tCO2. At the 2017 U.S. average emission intensity, this falls to \$45/tCO2 (**?**). These abatement costs are within the range of commonly discussed estimates of the SCC (**?**), and indicate that in some jurisdictions, repeated financial reward programs similar to the one studied in this work may be cost effective from a social perspective. The results of Chapter 1 and Chapter 2 show that the continued incentive of repeated financial reward is important for maintaining and causing additional reductions in electricity use. The continued incentive improves the program's cost-effectiveness, compared to a program offering a single annual conservation challenge. This improved cost-effectiveness occurs for two reasons. First, the program administration fixed costs are spread across additional conservation challenges. Second, the repeated incentive causes additional reductions, and keeps use from rebounding back close to pre-program levels.

2.7 Conclusion

In this chapter, I analyze the re-enrollment of households in a program offering financial rewards for energy conservation. I first study households' extensive margin decisions to re-enroll. I find that in deciding whether to re-enroll, households are responsive to their success or failure in achieving their conservation target, and yet, conditional on this success or failure, are notably unresponsive to their degree of electricity conservation. The decision to re-enroll also differs little across observable household characteristics. These patterns suggest that households use heuristics in making participation decisions, rather than accounting for the incentive structure of the program or information on their degree of conservation effort. I then exploit the discontinuous change in the probability of re-enrollment at the threshold for success in a fuzzy regression discontinuity strategy; I estimate that, despite self-selection into continuing, re-enrollment causes additional reductions in electricity use. This chapters findings imply it is important for repeated incentive programs to explicitly consider the process of self-selection into or out of re-enrollment. This is especially important when treatment effects fully dissipate after the incentive ends. Programs may be able to improve their cost effectiveness by encouraging re-enrollment, particularly when consumers fail to achieve their goal and self-select out of the program as a result. One way to address this could be to offer tiered incentives, such as offering an alternative challenge only to consumers that missed their target and are likely to self-select out of the program. Such an alternative incentive could be either transparently offered at the outset of participation, or revealed only to consumers that have not re-enrolled after some time period to target those likely to be ending participation. The responsiveness of households to their success or failure also opens up the possibility of modifying the reward structure to exploit this sensitivity; for example, by using unexpected consolation prizes instead of an all-or-nothing reward.

The use of information is central to how we model the decisions of agents across many questions within economics, from how tax rates affect labour market responses to voting behavior. The use of heuristics may be ubiquitous in these decisions. Heuristics could result from rational inattention and allow households to avoid the cognitive and time costs of considering complex incentives, or alternatively they could result from households responding to incentives and information based on models of behaviour beyond standard neoclassical models of decisions. Disentangling the ways in which consumers use information is an important avenue for future research. It is particularly important to the design of complex price schedules that are increasingly facilitated by the spread of smart meters, distributed electricity generation, and electricity storage.

A further direction for inquiry are the consequences of repeating interventions. Combined together, chapter 1 and chapter 2 showed how interventions that do not cause persistent effects can still produce cost-effective long-run changes by being repeatedly offered. These findings also suggest that consumers respond to each discrete reward in isolation, and do not consider the dynamic consequences of their effort on the future incentive structure. This raises the questions of whether this is a feature of the annual time scales of this program, or is a common feature of consumer responses to discrete rewards, and how consumers would respond as the time between potential rewards was reduced towards the limit of a continuous incentive.

Chapter 3

The Choice of Transportation Mode In International Trade

3.1 Introduction

All international trade of physical products requires a method of transportation. Transportation modes differ in many ways. Air freight is substantially faster but more expensive than sea freight. Fast delivery times are particularly important for perishable products, to industries utilizing just-in-time manufacturing, and in assisting firms meet unexpected demand shocks or recover from supply chain interruptions. Geographic constraints affect access to markets, with planes able to deliver directly to the interior of continents while ships are limited to navigable waters and accessible ports. Infrastructure improvements and technological changes vary between transport modes, causing different impacts across countries and industries depending on the modes they rely upon. In addition, transportation modes have very different environmental impacts. For example, airplanes generate on the order of 100 times the carbon emissions of shipping the same good the same distance by sea, while the high sulfur bunker fuels used by ships are responsible for 15%–30% of global NOx emissions (Corbett et al., 2007; Cristea et al., 2012). Understanding how these factors affect trade, evaluating policies like Pigovian taxes or infrastructure investments, and exploiting variation in mode-specific factors like distance all require a model of trade differentiated by the mode of transport. Critical to this is the degree of substitution between, and choice of, transport modes.

In this chapter, I develop a general model of the choice of transport mode in international trade. I use this model to show how several extant trade models incorporating mode choice start from different theoretical motivations, yet produce closely similar reduced form equations for bilateral trade differentiated by transport mode. I then show that this reduced form equation—which is common across several trade models—imposes strong and possibly unrealistic restrictions on substitution patterns across modes and countries. Importantly, all these models share a common assumption of a positive cross-elasticity for the share of imports carried by separate transport modes with respect to their trade costs. This positive cross-elasticity has been used to argue that transport modes are substitutes, but—as I will demonstrate—this is in fact insufficient to define transport modes to be substitutes. Instead, I show that the condition under which separate transport modes within bilateral trade are complements or substitutes depends on the magnitudes of this mode-share elasticity and a conventional trade elasticity. Prior modeling approaches that have estimated or assumed positive cross-elasticities—and have concluded that modes are substitutes—are in fact imposing modes to be complements, while other models are imposing modes to be substitutes. This assumption can substantially affect counterfactual trade patterns. As a result, these models are particularly sensitive to the choice of substitution elasticities and can poorly approximate the policies and trade patterns they are designed to analyze: those involving mode-specific changes in freight costs.

Improving on these limitations motivates the empirical contributions of this chapter. I first use detailed data on U.S. imports arriving by air and ocean transport to undertake estimates of the substitutability of transport modes within international trade. By exploiting idiosyncratic changes in freight rates to estimate reduced form own- and cross-price elasticities for mode-specific imports, I find that while higher freight costs reduce imports, little of this reduction substitutes to the alternative transport mode within bilateral trade. Instead, the evidence found implies a small or zero cross elasticity of demand for imports typically carried by air, and a larger and negative cross elasticity—complements—for imports typically carried by sea. These estimates demonstrate that the way multiple transportation modes have been modeled to date does not accurately approximate real substitution patterns. It thus suggests future work should return to the drawing board of modeling multiple modes of transportation within international trade, and must further empirically establish the substitution patterns across modes and countries that need to be approximated. I then reconsider the previously made observation that a substantial share of imports arrives by both transport modes. I find evidence that this overlap in products arriving by both modes reflects unobserved heterogeneity in product quality and therefore may not—as has been argued previously—be evidence of a large potential for substitution between transport modes. In addition, I find that unobserved quality is a particularly important, and largely unrecognized, determinant of the choice of transport mode across countries.

A variety of work has considered international trade differentiated by the method of transport; of particular importance are models of Lux (2011), Hummels and Schaur (2013), and Shapiro (2016). These models start from different theoretical motivations, and their models have been used to consider different policies and aspects of mode-specific trade. As a result, the close similarities between these models are not immediately obvious. The first contribution of this chapter is to develop a generalized model of mode choice in international trade—generalized in the sense that the reduced form equations for bilateral trade derived and estimated by Lux (2011), Hummels and Schaur (2013), and Shapiro (2016) can all all be viewed as special cases of this generalized model. This demonstrates how these models are closely connected and that they share common restrictions on trade patterns. In particular, whether they implicitly model transport modes as substitutes or complements depends on the values of two key elasticities—the elasticity of trade shares and mode shares.[34] This "generalized" model is not an attempt to build a new microfounded model of endogenous mode choice or demonstrate alternative reduced form equations for bilateral trade. In particular, the goal of this model is not to accommodate different theories of how importers choose a transport mode—such as the theoretical origin of transport modes being complements vs. substitutes—or accommodate some important features of the data, such as unobserved product heterogeneity. Instead, the assumptions underlying this generalized model are intentionally chosen to show the similarity between the prior models of Lux (2011), Hummels and Schaur (2013), and Shapiro (2016), their shared strengths and weaknesses in modeling substitution patterns across countries and methods of transport, and demonstrate how the elasticities of mode

[34]Trade shares refers to standard bilateral trade shares and mode shares are the value share within bilateral trade carried by a specific transport mode. Both are defined in Section 3.2.

shares and trade shares jointly determine whether modes are complements or substitutes.

Lux (2011) extends the work of Eaton and Kortum (2002) to incorporate multiple transportation modes and endogenous mode choice in a general equilibrium trade model. Lux (2011) uses this model to estimate a mode share elasticity that implicitly treats transport modes as highly substitutable. As a result of this high substitutability between modes, Lux (2011) concludes that reductions in mode-specific trade have relatively little impact on welfare. In this chapter, I contribute contrary empirical estimates that find modes are not highly substitutable. In addition, I show how a low substitutability between transport modes is inconsistent with restrictions on across-country substitution patterns that are implicit in Lux (2011) and similar models. Closely related to the work by Lux (2011) and this chapter is a paper by Shapiro (2016), who builds an Armington trade model differentiated by transport mode to consider welfare impacts of regulating carbon emissions from international transportation. As discussed in Shapiro (2016), shipping a product by air releases approximately 103 times the emissions per tonne-km as shipping the same product the same distance by sea.[35] While not considered by Shapiro (2016), this substantially larger emissions intensity for air is further exacerbated by the additional radiative forcing of contrails and ozone formation—a tonne of CO2 released by an airplane contributes 1.2–2.7 times the warming impact of a tonne of CO2 released at sea level (Grassl and Brockhagen, 2007). These differences in radiative forcing and carbon emission intensities make the choice of transport mode particularly critical to the direct climate change impacts of international transportation. I show that the model of Shapiro (2016) has close analytic similarities to that of Lux (2011), however Shapiro (2016) does not estimate the substitutability across modes. Instead, Shapiro (2016) assumes parameter values for his model that, as I will show, implicitly impose air and ocean to be complements; this assumption has a substantial effect on estimated changes in trade flows, carbon emissions, and policy findings.

How substitutable are air and ocean transport? The magnitude of substitutability and how to model it depends on what determines mode choice. Several important and inter-related determinants of mode choice have been considered previously in the literature. A particularly important one is differences in delivery time between air and sea transport. Without a value on time, importers would almost always choose the typically cheaper option of ocean freight. Hummels and Schaur (2013) model this trade off between time and cost in mode choice to estimate a value of time: they find "*each day in transit is equivalent to an ad-valorem tariff of 0.6 to 2.1 percent.*" In this work, I show that the equation estimated by Hummels and Schaur (2013) is equivalent to that of Lux (2011). Hummels and Schaur (2013) exploit cross-sectional variation in freight rates across U.S. coasts while using the same data and estimating equation, and they estimate an elasticity of substitution that implies that air and ocean are complements. This directly, though not obviously, contradicts the estimates of Lux (2011) and supports the assumption of complements made implicitly by Shapiro (2016). While I note that this paper does not attempt to model a microfounded reason why transport modes may be complements or substitutes, one potential source of complementarity could be fixed costs in supply chains. For example, consider firms that pay a fixed cost per country to source imports. A shock to a single mode's transport costs may, on the margin, induce a firm to pay the fixed cost for an alternative country—from which it then sources imports by both air and sea transport. Similarly, firms may wish to source intermediate

[35] This point was also previously discussed by Cristea et al. (2012).

inputs used in Leontief type production from the same or similar suppliers to ensure compatibility. If they rely on different transport modes for different products, a single mode cost shock may induce them to import products carried by multiple modes from a different exporting country.

It has been previously observed that many products within detailed product category-trade route-years arrive by both air and ocean transport (Hummels and Schaur, 2010; Lux, 2011; Hummels and Schaur, 2013). As these authors note, the same products arriving by both air and ocean transport imply a greater potential substitution across modes compared to if different modes were used exclusively for different products. In this paper, I document a substantial heterogeneity in unit values which implies that, even within detailed product categories, air and ocean transport modes may be specializing in different varieties of a product. This reduces the potential for substitution between transport modes. In addition, as freight charges are predominantly determined on a per weight basis, iceberg trade costs fall as the unit value of a product rises (Hummels and Skiba, 2004). This makes the ad-valorem price premium paid for air transport fall as unit values rise. The importance of unit values to mode choice has been previously considered by Harrigan (2010), who builds a model of comparative advantage between transport modes based on product weight and shipping distance. He tests and confirms two model predictions: the U.S. imports heavier products from nearby countries and light products from distant countries, and products imported from distant locations have higher unit values. The findings of this chapter support the observation by Harrigan (2010) that the importance of distance to mode choice may be highly non-linear, and that a comparative advantage in high quality products by countries with high productivity may be particularly important to the choice of transport mode. As a result, these comparisons across modes within a country suggest a potentially new way to study unobserved quality differences within products across countries (Schott, 2004; Hallak and Schott, 2011; Hummels and Klenow, 2005). In a separate paper to their work on time as a trade barrier in Hummels and Schaur (2013), Hummels and Schaur (2010) show how fast delivery by air, compared to slow but cheap ocean shipping, can be used to hedge uncertainty in demand and prices.[36] I add to both of these papers by contributing evidence that differences in distance between countries, and therefore in delivery times, are a relatively unimportant determinant of the choice of mode. In contrast, I show that unobserved product quality—as proxied by unit values within the most detailed product category available, HS10—is highly correlated with the choice of transport mode across countries. This largely unrecognized factor of unobserved quality may be more important in determining the choice of mode than the variation in distance between countries and across U.S. coasts that has been previously exploited for identification, for example by Harrigan (2010) and Hummels and Schaur (2013).

The remainder of this chapter is organized as follows. Section 3.2 derives a model of mode-specific bilateral trade and discusses implications of the model. Section 3.3 shows the similarities between this model and existing models that incorporate mode choice. I present reduced form estimates of the elasticity of substitution across modes in Section 3.4. In Section 3.5 I discuss how unobserved product quality can explain much of the variation in mode choice across countries and within detailed product

[36] An additional literature considers the value of time in delivery without explicitly differentiating across transport modes. In particular, Harrigan and Venables (2006) consider time as a factor in agglomeration of stages of production and Evans and Harrigan (2016) consider how different valuations for fast delivery determine the distance from which goods are imported.

categories. I conclude in Section 3.6.

3.2 A Model of Endogenous Transportation Mode Choice

In this section I develop a simple trade model incorporating endogenous mode choice, which I refer to as the nested CES model.[37] After deriving the nested CES model, Subsection 3.2.1 shows how this model predicts trade flows to change in response to a shock to trade costs. Subsection 3.2.2 explores the implications of this model for substitution patterns across countries and transport modes, including the definition and conditions under which modes are modeled as complements or substitutes.

The primary advantage of this nested CES model is that it produces an intuitive and general model of mode-specific trade. This model does not, and is not intended to, produce any new insights into theoretical mechanisms for substitution across transport modes, nor a reason for why transport modes may be complements or substitutes. As I show in Section 3.3, this general nested CES model yields the mode-specific trade patterns derived in Lux (2011), Hummels and Schaur (2013), and Shapiro (2016), and the estimating equations of Lux (2011) and Hummels and Schaur (2013). By transparently clarifying the relationships between these models, the nested CES model demonstrates that, although they have been derived from different theoretical assumptions and to answer different questions, these models are closely related and share strengths and weaknesses in approximating mode-specific trade flows.

Consider a representative consumer with CES preferences over a composite bundle Q_w of product variety w delivered by mode m, and with elasticity of substitution $\eta > 1$ across varieties,[38]

$$U = \left[\sum_{w} (Q_w)^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}} \tag{3.1}$$

Within the bundle Q_w the consumer has CES preferences over physical goods q_w^m delivered by mode m with elasticity of substitution $\rho > 1$ across modes,

$$Q_w = \left[\sum_{m=1}^{2} \lambda_w^m (q_w^m)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} \tag{3.2}$$

λ_w^m is a price-equivalent preference shifter that captures preferences over different modes, in particular shipping time, and models delivery time as a form of perceived quality in the same manner as Hallak (2006) and Hummels and Schaur (2013). The representative consumer of importing country n maximizes utility subject to their budget constraint of total expenditure X_n[39],

[37] CES refers here to constant elasticity of substitution.

[38] The bundle Q_w includes both imports and domestically sourced varieties.

[39] In this model I abstract away from a production function and factor markets and consider only the consumer demand. I hold expenditure X_n fixed. However, this is not only a partial equilibrium analysis; the model includes the indirect effect of a change in trade costs through its impact on all other modes and countries via the price indexes. As Section 3.3 will show, this yields the same equation for mode-specific trade as the general equilibrium models of Lux (2011) and Shapiro (2016).

$$\sum_w \sum_m p_w^m q_w^m \leq X_n$$

With the Armington assumption that each country produces a differentiated variety w, the expenditure share $\frac{X_{ni}^m}{X_n}$ by country n on imports from country i by mode m can be derived,

$$\frac{X_{ni}^m}{X_n} = \underbrace{\frac{(\lambda_{ni}^m)^\rho (p_{ni}^m)^{1-\rho}}{P_{ni}^{1-\rho}}}_{\gamma_{ni}^m} \underbrace{\left(\frac{P_{ni}}{P_n}\right)^{1-\eta}}_{\pi_{ni}} \tag{3.3}$$

where price indexes for bilateral trade P_{ni} from i and P_n for country n are given by

$$P_{ni}^{1-\rho} = \sum_m (\lambda_{ni}^m)^\rho (p_{ni}^m)^{1-\rho}$$

$$P_n^{1-\eta} = \sum_i {P_{ni}}^{1-\eta}$$

The derivation is discussed in Appendix C.1.1 and follows Sillard and Wilner (2015).[40] λ_{ni}^m is perceived quality across modes and which may differ across origin-destination, and p_{ni}^m is the Cost-Insurance-Freight (CIF) unit price of quantity q_{ni}^m.[41] I assume that the Free-On-Board (FOB) price c_i is independent of mode and there are iceberg trade costs τ_{ni}^m such that $p_{ni}^m = c_i \tau_{ni}^m$.[42]

Equation (3.3) gives the mode-specific bilateral trade shares conditional on unobserved preferences λ_{ni}^m, origin prices c_i, and iceberg trade costs τ_{ni}^m.[43] Mode-specific bilateral expenditure X_{ni}^m can be decomposed into a conventional bilateral trade share π_{ni} and, within that bilateral trade, a mode share γ_{ni}^m, such that $X_{ni}^m = X_n \pi_{ni} \gamma_{ni}^m$. The trade share π_{ni} is the share of value imported from country i to country n by both modes, and for this model is derived as[44]

$$\pi_{ni} \equiv \frac{X_{ni}}{X_n} = \left(\frac{P_{ni}}{P_n}\right)^{1-\eta} \tag{3.5}$$

Mode share γ_{ni}^m is defined as the share of value within bilateral imports X_{ni} from country i to n that

[40] An interpretation of this derivation is the consumer uses two-stage budgeting as in Sillard and Wilner (2015). The representative consumer first chooses their optimal expenditure over variety w and then their optimal expenditure over modes within a variety. The importance of this becomes apparent when considering the income and substitution effect interpretation of the separable trade and mode shares derived here.

[41] With a single transport mode, the trade shares of equation (3.3) reduces to a standard form for bilateral trade shares common to a wide variety of structural gravity models (Head and Mayer, 2015),

$$\pi_{ni} \equiv \frac{X_{ni}}{X_n} = \left(\frac{P_{ni}}{P_n}\right)^{1-\eta} = \frac{\left({\lambda_{ni}}^{\frac{\rho}{1-\rho}} c_i \tau_{ni}\right)^{1-\eta}}{\sum_i {P_{ni}}^{1-\eta}} \tag{3.4}$$

[42] In Section 3.4 I show the assumption that c_i is independent of mode does not hold in the data. I maintain the assumption here to show the similarity to other models.

[43] It is only necessary to specify a functional form for λ_{ni}^m for some estimates of ρ. This is discussed in Subsection 3.2.3.

[44] This follows from Appendix C.1.2.

is shipped by mode m,

$$\gamma_{ni}^{m} \equiv \frac{X_{ni}^{m}}{X_{ni}} = \frac{\lambda_{ni}^{m\,\rho}(p_{ni}^{m})^{1-\rho}}{P_{ni}^{1-\rho}} \tag{3.6}$$

3.2.1 Transport Cost Shocks and Trade Changes

How trade changes due to trade cost shocks provides an intuitive interpretation for the elasticities of substitution ρ and η. Consider changes in imports from country i by mode m, X_{ni}^{m}, due to a change in bilateral mode-specific trade costs $\tau_{nj}^{m'}$ from exporting country j to importing country n by mode m',[45]

$$\frac{\partial ln X_{ni}^{m}}{\partial ln \tau_{nj}^{m'}} = \frac{\partial ln X_{n}}{\partial ln \tau_{nj}^{m'}} + \underbrace{\frac{\partial ln \pi_{ni}}{\partial ln \tau_{nj}^{m'}}}_{Eqn.\,(3.8)} + \underbrace{\frac{\partial ln \gamma_{ni}^{m}}{\partial ln \tau_{nj}^{m'}}}_{Eqn.\,(3.9)} \tag{3.7}$$

Imports are valued CIF and the minor difference from using FOB values is discussed in Appendix C.1.3. Equation (3.7) is general. It includes both changes in trade costs within bilateral trade $n-i$ by mode m ($m = m'$ and $j = i$) as well as changes in trade costs among alternative transport modes $m \neq m'$ and alternative origin countries $j \neq i$. I assume that total expenditure X_n, which includes expenditure on domestic sourcing, is constant and therefore the first term of equation (3.7) is zero. Changes in imports X_{ni}^{m} due to changes in a trade cost has two components. The second term of equation (3.7) is the change in trade shares,

$$\frac{\partial ln \pi_{ni}}{\partial ln \tau_{nj}^{m'}} = \begin{cases} (1-\eta)(1-\pi_{ni})\gamma_{ni}^{m'} & < 0 \; if \; i \; = \; j \; \forall \, m, m' \\ (1-\eta)(-\pi_{nj})\gamma_{nj}^{m'} & > 0 \; if \; i \; \neq \; j \; \forall \, m, m' \end{cases} \tag{3.8}$$

The third term of equation (3.7) is the change in mode shares,

$$\frac{\partial ln \gamma_{ni}^{m}}{\partial ln \tau_{nj}^{m'}} = \begin{cases} (1-\rho)(1-\gamma_{ni}^{m'}) & < 0 \; if \; i = j \; and \; m = m' \quad (3.9a) \\ (1-\rho)(-\gamma_{nj}^{m'}) & > 0 \; if \; i = j \; and \; m \neq m' \quad (3.9b) \\ 0 & if \; i \neq j \; \forall \, m, m' \quad (3.9c) \end{cases}$$

Combining, the total changes in expenditure from country i by mode m is

$$\frac{\partial ln X_{ni}^{m}}{\partial ln \tau_{nj}^{m'}} = \begin{cases} (1-\eta)(1-\pi_{ni})\gamma_{ni}^{m} + (1-\rho)(1-\gamma_{ni}^{m}) & if \; i = j \; and \; m = m' \quad (3.10a) \\ (1-\eta)(1-\pi_{ni})\gamma_{ni}^{m'} + (1-\rho)(-\gamma_{ni}^{m'}) & if \; i = j \; and \; m \neq m' \quad (3.10b) \\ (1-\eta)(-\pi_{nj})\gamma_{nj}^{m'} & if \; i \neq j \; \forall \, m, m' \quad (3.10c) \end{cases}$$

Equation (3.10) shows how imports X_{ni}^{m} involve a change in trade shares π_{ni} as trade redistributes to or from other origin countries than i, and within this bilateral trade a change in the share transported by

[45]Derivation in Appendix C.1.2. Partial derivatives indicate that while all indirect effects of $\partial ln \tau_{nj}^{m'}$ through other countries and modes are accounted for, other trade costs are held fixed.

mode m, mode share γ_{ni}^m. Equations (3.8) and (3.9) provide a simple interpretation of the substitution elasticities ρ and η. $1-\eta \equiv \theta$ is the conventional trade elasticity common to a variety of gravity models (Head and Mayer, 2015).[46] $1-\rho$ is the elasticity of mode shares γ_{ni}^m to changes in transport costs $\tau_{nj}^{m'}$. Note that other factors that affect trade, like perceived quality for fast delivery, distance, or differences in origin prices, are determinants of the observed trade and mode shares through equations (3.5) and (3.6).

3.2.2 The Substitutability of Transport Modes

The previous section showed how changes in mode-specific bilateral trade shares, equation (3.3), can be decomposed into changes in conventional trade shares and mode shares. While the analytic form of this decomposition is convenient, in this subsection I show this analytic form comes at a cost of imposing potentially strong restrictions in how trade redistributes across countries and modes in response to a trade cost shock. Because equation (3.3) incorporates the mode-specific trade shares of Lux (2011), Hummels and Schaur (2013), and Shapiro (2016), the restrictions discussed here also apply to their models.[47] In other words, the restrictions on substitution patterns discussed here are not only characteristic of the nested CES model but are a general characteristic of the class of models that yield mode-specific trade flows in the form of equation (3.3). In addition, these restrictions are a direct consequence of the assumptions underlying these trade models. Alternative trade models, for example a model with an alternative ordering of preferences over varieties and modes, produce a different decomposition of mode-specific bilateral trade and alternative sets of restrictions over counterfactual trade patterns.

Complements or Substitutes

There are two ways to define the substitution across transportation modes. The models of Lux (2011) and Shapiro (2016) use one potential definition: the elasticity of mode *shares* to a change in trade costs,

$$\frac{\partial ln\gamma_{ni}^m}{\partial ln\tau_{ni}^{m'}} \equiv \begin{cases} Own - price\, Mode\, Share\, Elasticity < 0 & for\ m = m' \\ Cross - price\, Mode\, Share\, Elasticity > 0 & for\ m \neq m' \end{cases}$$

This form is intuitive. An increase in trade costs for mode m decreases the mode share for m, while an increase in trade costs for an alternative mode m' causes a rise in the mode share for m.[48] However, defining transportation modes as substitutes if they have a positive cross elasticity in mode *shares* may be misleading.

In contrast, consider the own and cross-price elasticities for the *level* of imports, X_{ni}^m.

[46] The additional terms multiplying $1-\eta$ in equation (3.8) reflect the fact the change in trade shares is not isoelastic and capture the general equilibrium effect of freight cost changes on the price index. Similarly for $1-\rho$ in equation (3.9).

[47] The restrictions are equivalent to the Independence of Irrelevant Alternatives (I.I.A.) property common, and well known, to nested discrete choice models. However, it may not be apparent that this is also an implicit property of many trade models incorporating transportation mode choice.

[48] This is the general form of the own and cross-price elasticities of mode shares to bilateral trade costs in the nested CES model, equations (3.9a) and (3.9b).

$$\frac{\partial ln X_{ni}^{m}}{\partial ln \tau_{ni}^{m'}} = \begin{cases} Own - price\, Import\, Elasticity & if\, m = m' \\ Cross - price\, Import\, Elasticity & if\, m \neq m' \end{cases} \tag{3.11}$$

The specific analytic form of these will depend on the particular model of mode-specific trade. For the nested CES model and equivalent models these have the form of equations (3.10a) and (3.10b). The own-price elasticity for changes in expenditure X_{ni}^{m} to a trade cost change to the same mode m within that bilateral trade route is unambiguously negative,[49]

$$\frac{\partial ln X_{ni}^{m}}{\partial ln \tau_{ni}^{m}} = \underbrace{(1-\eta)(1-\pi_{ni})\gamma_{ni}^{m}}_{(-)} + \underbrace{(1-\rho)(1-\gamma_{ni}^{m})}_{(-)}$$

This is intuitive: increases in freight costs for a transport mode decrease imports carried by that same mode. The sign of the cross-price elasticity for the level of imports arriving by mode m depends on two terms,

$$\frac{\partial ln X_{ni}^{m}}{\partial ln \tau_{ni}^{m'}} = \underbrace{(1-\eta)(1-\pi_{ni})\gamma_{ni}^{m'}}_{(-)} + \underbrace{(1-\rho)(-\gamma_{ni}^{m'})}_{(+)} \tag{3.12}$$

Imports carried by each mode are gross substitutes if this cross-price elasticity of equation (3.12) is positive—imports by mode m rise in response to a rise in trade costs for the alternative mode m'—and gross complements if the cross-price elasticity is negative. The first term is the change in trade shares and is strictly negative while the second term is the cross-elasticity of mode *shares* and is strictly positive. The cross-price elasticity for this level of imports is positive, and modes are substitutes, if

$$(1-\pi_{ni}) \quad < \quad \frac{(1-\rho)}{(1-\eta)} \tag{3.13}$$

Note that, while expenditure here is the CIF value, the cross-price elasticity is the same when expressed in FOB values.[50] Intuitively, complementarity arises when the decrease (increase) in total imports through trade share changes is larger than the increase (decrease) in imports substituting from mode m to m'.[51]

This highlights the important distinction between the elasticity of mode *shares* and the elasticity of substitution for the *level* of imports. An increase or decrease in the mode share of imports does not identify the changes in levels of constituent imports carried by each mode. A rise in the share of imports carried by air can occur along with a decline in the level of imports arriving by air, as long as imports arriving by the alternative mode ocean fall faster. In this way, a positive elasticity of mode

[49] I assume that $\eta > 1$ and $\rho > 1$.

[50] A change in freight costs will change the value of imports that include those freight costs; products redistributing to other modes and countries will face those other country and mode freight costs and modes, which are unaffected in this model, and therefore the percentage change in value will be the same for a cross-price elasticity whether measured in CIF or FOB values. This is shown formally in Appendix (C.1.3).

[51] This can be considered as an income and substitution effect. Complementarity arrises when the income effect of reduced expenditure on imports within a bilateral trade route is larger than the substitution effect between modes within that bilateral trade.

shares is consistent with the level of imports carried by air and ocean transport being either substitutes or complements.

I stress that the potential for both substitutes and complements is not in itself a model downside. Does complementarity occur only for unreasonable values of the elasticities of substitution ρ and η, or only within the nested CES model of mode choice? In Section 3.3 I show that complementarity between modes can occur in several additional models of mode choice, and prior work has estimated and assumed values for ρ and η such that air and ocean transport are treated as complements and substitutes.

Independence of Trade and Mode Elasticities

Equation (3.8) shows changes in trade shares are weighted by the share of bilateral trade subject to the cost increase, $\gamma_{nj}^{m'}$. Notably, these changes in trade shares are independent of the mode share elasticity. An advantage of this independence—an assumption of these models—is it allows the trade elasticity to be identified from mode-specific changes in freight costs, for example as in Shapiro (2016), without considering how trade redistributes across modes within bilateral trade. Is this independence assumption realistic? Perhaps not. It is reasonable to expect that a shock to a single transport mode's trade costs would induce a smaller change to trade shares if imports are highly substitutable across modes, compared to if modes have a low elasticity of substitution. For example, consider this model framework and a region with a 50% air shares. A 1% increase in both this region's air and ocean trade costs would then cause the same reduction in trade shares as a 2% increase to a single mode's trade cost. This rules out the situation where the 2% single-mode cost shock, compared to the 1%, primarily induces a substitution to the alternative mode with little reduction in trade shares.[52]

Independence of Mode Shares to Costs in Alternative Trade Routes

A closely related restriction is that mode shares within a bilateral trade route are independent of the trade costs in alternative routes. This is shown by the elasticity of 0 in equation (3.9c).[53] The importance of this restriction can be seen by considering hypothetical imports arriving from China to the U.S. carried via air freight. What happens if there is a decline in air freight costs for imports to the U.S. from a different country, Japan? On the margin it is reasonable to expect this negative cost shock to induce products to switch from China to Japan to take advantage of the now-lower air freight cost. It is also reasonable for this reduction in trade from China be largest for products that are already transported by air from China, compared to products that are already shipped by ocean from China. As a result, we would expect the share of products imported by air from China to fall due to the negative shock to Japan's air freight costs. This cannot happen in these models. Instead, in response to trade cost shocks, imports redistribute across or from alternate trade routes proportional to those routes' initial mode and trade shares; whether these imports were previously transported by air or ocean does not matter. A further and important implication of (3.9c) is that these models are particularly

[52] This independence assumption may be particularly important to estimates of sectoral trade elasticities where both the elasticity of substitution between products and level of mode shares are likely to vary across sectors.

[53] This is the Independence of Irrelevant Alternatives (I.I.A.) property common to nested discrete choice models.

poor at approximating trade changes when there is little substitutability across modes. For example, under the condition that the cross-price elasticity within bilateral trade—equation (3.12)—was 0, an increase to air freight costs from Japan would induce no substitution to imports by ocean from Japan. At the same time, the reduction in imports by air from Japan would redistribute across both air and ocean among other countries with a larger redistribution to imports via ocean—not air—among other countries.

3.2.3 Estimating Equation for the Elasticity of Mode Shares

In addition to observed trade and mode shares, simulating counterfactuals with this model requires estimates of the elasticities of substitution η and ρ. A wide variety of estimates for the trade elasticity, $1-\eta$, are available in the empirical trade literature. One approach to estimating the elasticity of mode shares ρ starts from taking the ratio of expenditures from equation (3.3) across modes within bilateral trade,

$$\frac{X^a_{ni}}{X^o_{ni}} = \frac{X_n \pi_{ni} \gamma^a_{ni}}{X_n \pi_{ni} \gamma^o_{ni}} = \frac{\gamma^a_{ni}}{\gamma^o_{ni}} = \frac{\lambda^a_{ni}{}^{\rho}(p^a_{ni})^{1-\rho}}{\lambda^o_{ni}{}^{\rho}(p^o_{ni})^{1-\rho}}$$

where λ^m_{ni} is the perceived quality over delivery times and p^m_{ni} and X^m_{ni} are measured CIF. Two further assumptions simplify the estimation. First, assume that λ^m_{ni} is constant over time and import prices are time-varying with iceberg trade costs, $p^m_{ni,t} = c^m_{i,t}\tau^m_{ni,t}$. Taking logs produces the following,

$$ln\left(\frac{X^a_{ni,t}}{X^o_{ni,t}}\right) = \rho ln\frac{\lambda^a_{ni}}{\lambda^o_{ni}} + (1-\rho)\frac{c^a_{i,t}}{c^o_{i,t}} + (1-\rho)\frac{\tau^a_{ni,t}}{\tau^o_{ni,t}} + \epsilon_{ni,t} \tag{3.14}$$

By choosing a parameterization of λ^m_{ni} and potentially additional fixed effects, equation 3.14 can be estimated. This is the general specification estimated by both Lux (2011) and Hummels and Schaur (2013), discussed below.

3.3 Comparisons to Existing Models of Mode Choice

The implications for changes in trade discussed in Section 3.2 are not unique to the nested CES model of endogenous mode choice. In this section, I show under what conditions the mode-specific bilateral trade shares derived from the nested CES model—and the resulting counterfactual trade changes and restrictions on substitution patterns—are equivalent to the models of Lux (2011), Hummels and Schaur (2013), and Shapiro (2016).

Lux (2011) is, to the best of my knowledge, the only attempt to model endogenous mode choice with an elasticity of substitution for mode shares that differs from the trade elasticity. Lux (2011) builds a general equilibrium Eaton and Kortum (2002) style model, incorporating multiple transport modes within international trade. His model incorporates transport modes by extending the Frechet distribution of productivity differences across varieties into a multi-dimensional distribution of variety-mode specific productivities. This allows Lux (2011) to derive equations for mode-specific trade shares, mode shares, and expenditure shares that are identical to those obtained from the nested CES model.

The difference, beyond the theoretical underpinnings of the model, is in the interpretation of the trade and mode share elasticities. Whereas the nested CES model interprets ρ as the elasticity of substitution over products delivered by different transport modes, Lux (2011) interprets ρ as a function of the dispersion of productivity draws from the Frechet distribution combined with the correlation of productivity draws across transport modes. As a result of the equivalence of the equations for mode-specific trade shares, the estimating equation for the elasticity of mode shares in Lux (2011) is equivalent to equation (3.14) of the nested CES model. Instead of specifying a form for λ_{ni}^{m}, Lux (2011) assumes it is time-invariant and estimates the mode share elasticity using first-differencing to control for the value of time in delivery and other time-invariant parameters. He finds a central estimate of $\rho = 7.7$.[54] Lux (2011) does not estimate a trade elasticity. Instead, he uses $\theta = 4$ from Simonovska and Waugh (2014). As a result, Lux (2011) implicitly finds that equation (3.13) is satisfied, and air and ocean transportation are substitutes.

It is important to contrast the estimate of ρ from Lux (2011) with that in Hummels and Schaur (2013). Hummels and Schaur (2013) do not consider different elasticities of substitution across modes compared to product varieties. The equivalence to the nested CES model results from imposing the restriction of a single elasticity of substitution across both varieties and modes, or $\rho \equiv \eta$.[55] Hummels and Schaur (2013) also derive an estimating equation for ρ from the ratio of imports by air and ocean within exporter-importer pairs. This is identical to that of the nested CES model, equation (3.14), when using the same parameterization for the value of delivery time as Hummels and Schaur (2013). Hummels and Schaur (2013) devote substantial time and care to estimating two key coefficients: the substitution elasticity across products and modes, and the coefficient on delivery time. The ratio of these estimated coefficients gives the value of time—the ad-valorem premium paid to reduce delivery time by one day. In their preferred specifications, Hummels and Schaur (2013) estimate ρ ranges from 1.5 to 3.3.[56] This is substantially smaller than the central estimate of 7.7 from Lux (2011). The key distinction between these estimates is that Hummels and Schaur (2013) exploit cross-sectional variation across U.S. coasts, while Lux (2011) exploits the time dimension only.

The difference in these estimates can be interpreted several ways. The first is to assume that the nested CES model (and, by extension that of Hummels and Schaur (2013)) is correctly specified such that $\rho \equiv \eta$. The estimates of Hummels and Schaur (2013) would then correspond to trade elasticities of $\theta = 0.5$ to 2.3. This is smaller than commonly estimated (Head and Mayer, 2015). The assumption that $\rho \equiv \eta$ would also imply that ocean and air freight are substitutes—albeit with a small cross-price elasticity. Alternatively, in the more flexible case of potentially different elasticities of substitution across products and modes, $\rho \neq \eta$, the estimates of Hummels and Schaur (2013) are an estimate of the mode share elasticity only. For common trade elasticities on the order of ~ 4, these estimates and equation (3.13) imply that air and ocean transport are complements. This contradicts the findings of Lux (2011). An important caveat to these estimates is in order. Identifying ρ from the ratio of air to ocean imports within a product category, as done by Lux (2011) and Hummels and Schaur (2013),

[54] Where "mode share elasticity" here refers to the nested CES definition of ρ used in this chapter. Lux (2011) uses different notation under which ρ is a rank correlation statistic.

[55] See Appendix C.2.

[56] Coast-differenced estimate from Tables 2 and 3. Note ρ in the notation of the nested CES model is called σ in the notation of Hummels and Schaur (2013).

implicitly assumes that similar products are being compared. Section 3.4 provides evidence that this ratio may compare substantially different varieties of products, calling into question what precisely is being identified through this estimation approach.

The comparison of Shapiro (2016) to the nested CES model is less straight-forward because Shapiro (2016) does not have a model of endogenous mode choice. Instead, Shapiro (2016) constructs trade shares and bilateral freight costs from mode-specific trade weighted by bilateral mode shares, then holds mode shares fixed when calculating counterfactuals. Shapiro (2016) also undertakes a robustness check for approximately endogenous mode shares by assuming a unit elasticity of mode shares to freight costs. Shapiro (2016) finds that *"[t]hese counterfactuals lead to extremely similar quantitative results to the main calculations, and provide one piece of evidence that endogenous mode share substitution for a given country pair and sector has limited impact on results."* As Shapiro (2016) notes, this robustness check is also not strictly speaking an endogenous mode choice, but instead a re-weighting of aggregate trade and freight costs by the new mode shares. This approach differs from the mode-specific trade shares of the nested CES model as the assumed mode share changes in Shapiro (2016) are not isoelastic. However, the central case considered by Shapiro (2016)—holding mode shares fixed in counterfactuals—corresponds to the nested CES model with fixed mode shares, or $\rho = 1$. Similarly, Shapiro (2016)'s assumption of a unit mode share elasticity is closely similar to the case of $\rho = 2.25$.[57] As a result, both the main policy analysis and the robustness check of a unit mode share elasticity considered by Shapiro (2016) impose that air and ocean are complements.

Lastly, I note that the structural equations for trade shares of equation (3.5) and mode shares of equation (3.6) can be derived from a discrete choice model. In particular, the equation of Hummels and Schaur (2013) is equivalent to a multinomial logit discrete choice model while the model of Lux (2011) or a nested CES model allow more general substitution patterns and are equivalent to a nested logit discrete choice model with the lower nest defined across modes.

3.3.1 The Importance of the Mode Share Elasticity

A back-of-the-envelope comparison shows that the range of substitution elasticity ρ estimated and assumed in Section 3.3 quantitatively and qualitatively matters. In Table 3.1 I show changes in counterfactual trade and emissions from three representative scenarios: 1% increases in trade costs for air imports, for sea imports, and for both together. These scenarios apply the respective trade cost changes to a region responsible for 20% of U.S. imports, which approximates the European Union trade share. Counterfactual trade is calculated using the mode-specific trade shares of equation (3.3).[58] I assume that any trade redistributing to domestic sourcing generates no direct transportation emissions, and that every dollar of trade that substitutes from ocean to air generates 50 times the emissions impact of sea. I consider three scenarios for the mode share elasticity. $\rho = 1$ corresponds to the fixed mode shares assumption of Shapiro (2016), $\rho = 2$ for the endogenous mode shares but complements assumption of Shapiro (2016), and $\rho = 7$ corresponding to the case of substitute transport modes in Lux (2011).

[57] For the case of an own-price elasticity with air shares of 20%.

[58] Using CEPII data for 2006 I find U.S. imports from home to be 75%. Including domestic expenditure makes 20% of U.S. imports a trade share of 5%. The *% Change in Imports* shows the change in total U.S. imports from all regions where total expenditure is held fixed and imports are displaced to domestic production.

Table 3.1: **Counterfactual Trade Changes - Complements and Substitutes**

	Increases in trade costs of:								
	1% Air, 1% Sea			1% Air			1% Sea		
	$\rho = 1$	$\rho = 2$	$\rho = 7$	$\rho = 1$	$\rho = 2$	$\rho = 7$	$\rho = 1$	$\rho = 2$	$\rho = 7$
Cross-price elasticity	(-)	(-)	(+)	(-)	(-)	(+)	(-)	(-)	(+)
%△ E.U. Imports	-3.80	-3.80	-3.80	-1.14	-1.14	-1.14	-2.66	-2.66	-2.66
%△ E.U. Air Imports	-3.80	-3.80	-3.80	-1.14	-1.84	-5.34	-2.66	-1.96	1.54
%△ R.O.W. Air Imports	0.20	0.20	0.20	0.06	0.06	0.06	0.14	0.14	0.14
%△ E.U. Emissions	-3.80	-3.80	-3.80	-1.14	-1.80	-5.07	-2.66	-2.00	1.27
%△ World Emissions	-0.60	-0.60	-0.60	-0.18	-0.31	-0.97	-0.42	-0.29	0.37
Emissions per $ trade	1.00	1.00	1.00	1.00	1.73	5.37	1.00	0.69	-0.87

Notes: Counterfactual percentage changes in trade due to, respectively, 1% increases in both air and ocean trade, only air trade, and only ocean trade costs. These trade costs are applied to a region approximating the E.U. trade share in U.S. imports of 5%. The Rest of the World (R.O.W.) has a trade share of 20%. Air transport is assumed to generate 50 times the emissions impact per tonne-km of ocean trade. Counterfactuals use $\eta = 5$ to match conventional trade elasticities of ~ -4. The Cross-price elasticity indicates whether air and ocean imports are complements or substitutes; (-) for complements and (+) for substitutes.

In Table 3.1 the first pattern of note is that changes in imports from the E.U. (%△ E.U. Imports) do not depend on the elasticity of mode shares under any of the trade cost scenarios. This is the consequence of the model assumption that the conventional trade elasticity is independent of the mode share elasticity. As noted previously, this is a potentially strong assumption, but the implications of this are not considered here.

The change in imports arriving by air from the E.U. is sensitive to the mode share elasticity. A 1% increase in air trade costs decreases E.U. imports arriving by air by 1.84% for complements ($\rho = 2$) and 5.34% under substitutes ($\rho = 7$). In contrast, a unilateral 1% increase in ocean trade costs decreases imports arriving by air by 1.96% under $\rho = 2$ and increases air imports by 1.54% under $\rho = 7$. Any policy analysis interested in trade volumes differentiated by mode would find not only a substantially different scale of effect but also potentially a different sign of effect, depending on the value of ρ used. In addition, differences in E.U. imports between the complements case of $\rho = 2$ and substitutes of $\rho = 7$ are larger than the differences across the scenarios of air and sea cost increases. This sensitivity to the mode share elasticity has important consequences for evaluating policies on carbon emissions. Both changes in the direct carbon emissions associated with imports from the E.U. (%△ E.U. Emissions) and in the direct emissions from all imports (%△ World Emissions) differ substantially depending on the mode share elasticity.[59] Under the complements case of fixed mode shares ($\rho = 1$), one would conclude that all three policies generate the same emissions reductions per dollar of lower trade. The policies differ only in the scale of changes. In contrast, under $\rho = 7$, regulating ocean transport increases the direct carbon emissions from trade while decreasing imports. This is likely welfare reducing.

[59]Changes in %△ World Emissions are small in levels because this policy is applied to the 20% of imports arriving from the E.U. region but are evaluated as percentage changes against all imports.

Table 3.2: **Counterfactual Trade Changes - No Substitution**

	Increases in trade costs of:					
	1% Air, 1% Sea		1% Air		1% Sea	
	$\rho = 4.8$	NS	$\rho = 4.8$	NS	$\rho = 4.8$	NS
%△ E.U. Imports	-3.80	-3.80	-1.14	-1.14	-2.66	-2.66
%△ E.U. Air Imports	-3.80	-3.80	-3.80	-3.80	0	0
%△ E.U. Sea Imports	-3.80	-3.80	0	0	-3.80	-3.80
%△ R.O.W. Air Imports	0.20	0.20	0.06	0.20	0.14	0
%△ R.O.W. Sea Imports	0.20	0.20	0.06	0	0.14	0.20
Cross-mode substitution			70	0	30	0
%△ E.U. Emissions	-3.80	-3.80	-3.63	-3.63	-0.17	-0.17
%△ World Emissions	-0.60	-0.60	-0.68	-0.57	0.08	-0.03
GHG/$ trade	1.00	1.00	3.767	3.185	-0.186	0.064

Notes: Counterfactual percentage changes in trade due to, respectively, 1% increases in both air and ocean trade, only air trade, and only ocean trade costs. These trade costs are applied to a region approximating the E.U. trade share in U.S. imports of 5%. The Rest of the World (R.O.W.) has a trade share of 20%. Air transport is assumed to generate 50 times the emissions impact per tonne-km of ocean trade. Counterfactuals use $\eta = 5$ to match conventional trade elasticities of ~ -4. The counterfactual labeled *NS* assumes there is no substitution between transport modes both within and across bilateral trade.

Independence of Mode Shares to Costs in Alternative Trade Routes

Section 3.2 discussed the independence of mode shares within a bilateral trade route from trade costs among alternative trade routes. This imposes substitution across transport modes across trade routes even in the case of no mode substitution within bilateral trade. Table 3.2 shows how this Independence of Irrelevant Alternative assumption affects counterfactual trade patterns.

The case $\rho = 4.8$ corresponds to zero substitution between transport modes within bilateral trade; this can be seen in the 0% change in imports by air [sea] from the E.U. under the 1% increase in ocean [air] trade costs. In contrast, both transport modes among the rest of the world have increases in imports (%△ R.O.W. Imports for Air and Sea.) Consider a different counterfactual that assumes all substitution is within-mode regardless of the trade route. This has no direct equivalent in the nested CES model framework, and is equivalent to modeling trade by air and ocean as completely independent trade with no general equilibrium connection between them. I denote this case in Table 3.2 by *NS (*No Substitution).

Under a 1% increase in air trade costs, the nested CES model predicts that, among air imports redistributing to alternative origin countries (R.O.W. imports), 70% will substitute to imports arriving via ocean—despite the zero substitution between air and ocean within bilateral trade. Similarly, 30% of the ocean imports substituting away from the E.U. to the R.O.W. under the 1% increase in ocean trade costs will substitute to air. This is why the nested CES model predicts an increase in global emissions from regulating ocean trade even without bilateral mode substitution, where as a true within-mode redistribution of trade would decrease world emissions. This also highlights how evaluating policies of mode-specific trade costs, particular for the direct environmental externalities of trade, can be critically dependent on the elasticity of mode shares and the restrictions over substitution patterns implicit in the model used.

3.4 Estimates of the Elasticity of Substitution Between Modes

Sections 3.2 and 3.3 demonstrated that prior models of the choice of transportation mode within international trade share close similarities in how they model counterfactual trade flows. This prior work also leaves it unclear whether air and ocean transport should be modeled as complements or substitutes, and whether import substitution between transport modes is important to consider in policy design and analysis. Are air and ocean transport complements or substitutes? There are two reasons why this key fact remains unclear, despite prior estimates of the mode share elasticity ρ. First, as Subsection 3.2.2 discussed, estimates of ρ alone do not pin down complements or substitutes. This requires an estimate of η, which to date have been estimated from different data sets and with different identification strategies than ρ. This leaves the suitability of comparing these parameters uncertain, particularly given the large differences in estimates of ρ between Hummels and Schaur (2013) and Lux (2011). Second, identifying complements or substitutes using using equation (3.13) or estimating ρ using equation (3.14) both assume that this model framework for mode choice is correctly specified. If it is not, then comparing ρ and η in equation (3.13) may not be sufficiently informative. In addition, the error term of equation (3.14) is likely to include other important explanatory factors and be subject to an omitted variable bias. This is, of course, a caveat that can be applied to any structural model. The cause for concern with this model is the potentially strong restrictions on substitution patterns discussed in Subsection 3.2.2. Because of these limitations, this section takes a different approach and I do not exploit variation in the ratio of imports carried by air and ocean transport. Instead, I directly test whether modes are complements or substitutes by estimating their reduced form own-price and cross-price elasticities. This estimation does not require the nested CES or associated models to be correctly specified, as does comparing η and ρ. The goal of this reduced form approach is to avoid estimates predicated upon a particular model being correct, and to understand the aggregate substitution patterns across countries and modes that a model of mode choice must reflect.

3.4.1 Data - U.S. Imports of Merchandise

Data suitable for studying mode choice is limited. Most publicly available customs data records import or export values, trade routes, and product categories, but not freight costs or the specific transport mode used. In addition, most industry data on freight costs is based on containers or aggregate weight, and does not observe the specific product being shipped. As a result, much work studying mode choice uses the same U.S. Imports of Merchandise dataset used in this chapter. This dataset records monthly data since 1990 on U.S. imports at an HS10-exporting country-importing district level, including the Free-on-Board value, weight, and freight costs all differentiated by air vs. ocean transport. HS10 corresponds to ten digit Schedule B product codes for U.S. imports and exports.[60] I exclude U.S. imports from Canada and Mexico, as I do not observe the method of transport and focus on imports arriving from overseas. I drop all imports of oil and natural gas and limit the panel to 1992-2007 to include complete years and exclude the Great Recession. Most results below are based on the data aggregated across months to the year level. I use monthly data to construct a measure of weight per

[60] The first 6 digits of this 10 digit code are the Harmonized System product code.

shipment, and aggregate to the six digit HS level in some estimations. Details on the data are available in Appendix C.3.

3.4.2 Reduced Form Elasticity Estimates

My preferred specification for estimating the own-price and cross-price elasticities of imports to trade costs is of the form

$$log\,X^m_{nipt} = \beta_m log(f^m_{nipt}) + \beta_{m'} log(f^{m'}_{nipt}) + \gamma_{ni} + \delta_{nt} + \alpha_{it} + \eta_p + \epsilon_{nipt} \tag{3.15}$$

where X^m_{nipt} is the FOB import value for HS6 product p by mode m exported from country i to U.S. import regions n in year t, f^m_{nipt} are per kilogram freight rates, γ_{ni}, δ_{nt}, α_{it}, and η_p are fixed effects at the origin-destination, destination-year, origin-year, and HS6 product level. Coefficients β_m and $\beta_{m'}$ are the own and cross-price elasticities of mode m and I estimate equation (3.15) separately for air and ocean imports. I use FOB import values to avoid endogeneity arising from including freight charges in the dependent variable, and use per kilogram freight rates — as opposed to ad-valorem — for the same reason.[61]

Identifying these own and cross-price elasticities has several challenges. Ideally, a natural experiment could be found that causes exogenous variation to a single mode's freight rates.[62] Lacking this, I exploit idiosyncratic origin-destination-time varying freight rates to identify the own and cross-price elasticities.

The first challenge to estimating equation (3.15) is endogeneity arising from omitted variable biases. I address most sources of omitted variable bias by using origin-year, destination-year, and origin-destination fixed effects. These control for any potential confounding variables that vary at these levels. For example, increases in region-specific energy prices that affect the demand for imports or the price of exports and are correlated with freight costs, economic booms, region-time varying labour costs, or mode-specific geography differences across routes are all controlled for through these flexible fixed effects.[63] Instead, I use any source of variation in freight costs that varies at the trade route-year-mode level. For example, changes in fuel costs that differentially affect freight rates across modes and trade routes is one source of identifying variation. This has close similarities to the identification strategies in Shapiro (2016) and, to a lesser extent, Hummels and Schaur (2013). Shapiro (2016) uses a similar specification to equation (3.15) to estimate sectoral trade elasticities. Shapiro (2016) also includes freight costs for Australian imports, which gives him two destination countries. Lacking Australian freight data, I exploit variation in destination regions within the United States; this is similar to the

[61] β_m and $\beta_{m'}$ are own and cross-price elasticities to per kilogram freight rates. These can be expressed as approximate iceberg trade cost elasticities β^τ_m, $\beta^\tau_m = \beta_m \cdot \frac{(1+s_m)}{s_m} + 1$, where s_m is the average freight charges divided by FOB value.

[62] Several possible candidates have been considered and found unsuitable. The closure of the Suez Canal studied by Feyrer (2009) would be a good candidate but that era lacks mode-specific trade data. Open-Skies Agreements considered in Micco and Serebrisky (2006) are another possible candidate. In attempting to replicate their results I find that OSA's have a smaller and largely statistically insignificant impact on freight rates. As a result, I do not find OSA's to be a suitable exogenous shock to freight rates sufficient for a first stage.

[63] Shapiro (2016) considers changes in tariffs and finds they do not affect his estimates; I do not repeat that exercise here.

country-import district variation used by Hummels and Schaur (2013).

Table 3.3 presents the results of estimating equation (3.15) separately for air imports (Panel A) and ocean imports (Panel B). Specification (1) finds own-price elasticities of the expected sign and magnitudes. Expressed in ad-valorem terms, at the median weight-to-value ratio the estimates in Specification (1) correspond to -2.9 for air and -7.1 for ocean. These are similar to trade elasticities in the literature, for example -4.25 in Melitz and Redding (2013) and in -3.7 in Shapiro (2016). To my knowledge, estimates of trade elasticities have not been undertaken for different transportation modes.

To estimate the cross-price elasticity, Specification (2) includes freight rates for imports arriving by the alternative mode within the same 6-origin-destination-year. In Panel A, for air imports, I find a small but statistically significant positive effect of ocean freight rates on air imports. This coefficient is an order of magnitude smaller than the own-price elasticity, and indicates little aggregate substitution between modes. In Panel B, for imports arriving by ocean, I find the cross-price elasticity is again smaller in magnitude than the own-price elasticity; however, I find a negative cross-elasticity. This is consistent with air and ocean imports being complements—not substitutes. Specification (3) uses an alternative set of origin country, destination region, and year fixed effects, and finds broadly similar results. This set of fixed effects is not preferred, however, as it does not allow origin and destination regions to have separate time trends. In Specification (4) I control for the value-to-weight ratio of imports, as it may be directly affect import volumes and freight rates; this finds the same pattern among the own and cross-price elasticities. Specification (5) uses import weight instead of import value as the dependent variable. This finds the opposite sign, though a still small cross-elasticity for imports arriving by air and a similar negative cross-price elasticity for imports arriving by ocean. Results are also robust to restricting the sample to different subsets of origin countries, destination regions, and time periods (Appendix C.4).

The second endogeneity challenge is simultaneity. While freight rates affect import volumes, it is also likely that import volumes affect freight rates. The sign of this bias is undetermined. Increases in imports may lead to port congestion and increase freight rates, or a higher import demand may lead to higher markups to freight rates under monopolistic competition (Hummels et al., 2009). Alternatively, rising imports may result in economies of scale that reduce freight rates. To address this source of endogeneity, I first note the unit of observation is at the HS6-year-trade route level. These observations are typically a small share of aggregate trade; as a result, they individually will have little impact on total freight volumes. I assume that, conditional on included fixed effects, including time fixed effects to control for aggregate trade volumes, the volumes in individual HS6-year-trade routes have no effect on freight rates through aggregate import volumes.

The "through aggregate import volumes" is important. A second source of simultaneity bias can arise if importers receive a discount on freight costs through bulk imports—a fact documented by Holmes and Singer (2018). This can occur independent of the contribution of an individual observations expenditure to aggregate trade. While bulk discounting may cause a negatively bias to β_m, it will not affect estimates of $\beta_{m'}$ as this is the coefficient on freight rates for the alternative transport mode. I address this potential bias in β_m by constructing a proxy for freight rates f_{ni}^m that is orthogonal to

Table 3.3: **Own and Cross-Price Elasticities**

	(1)	(2)	(3)	(4)	(5)
	Panel A : Imports by Air				
Dependent variable:	Log Air Value				Log Air Weight
Log Air Freight ($/kg)	-0.353***	-0.297***	-0.297***	-0.410***	-0.774***
	(0.00638)	(0.00816)	(0.0216)	(0.0113)	(0.0149)
Log Ocean Freight ($/kg)		0.0304***	0.0163***	0.0163***	-0.0292***
		(0.00528)	(0.00578)	(0.00508)	(0.00568)
Log Air Value/Weight				0.237***	
				(0.00893)	
Fixed Effects	A	A	B	A	A
N	4396852	1742710	1743484	1742710	1742710
r2	0.323	0.375	0.344	0.382	0.317

	Panel B : Imports by Ocean				
Dependent variable:	Log Ocean Value				Log Ocean Weight
Log Air Freight($/kg)		-0.130***	-0.125***	-0.131***	-0.136***
		(0.00692)	(0.0119)	(0.00691)	(0.00683)
Log Ocean Freight($/kg)	-0.409***	-0.420***	-0.450***	-0.501***	-0.927***
	(0.00689)	(0.00833)	(0.0167)	(0.0101)	(0.0104)
Log Ocean Value/Weight				0.159***	
Fixed Effects	A	A	B	A	A
N	4396852	1742710	1743484	1742710	1742710
r2	0.323	0.375	0.344	0.382	0.317

Notes: This table reports estimates of equation (3.15) for air and ocean imports. Fixed Effects A include HS6, Origin-Destination, Origin-Year, Destination-Year. Fixed effects B include HS6, Origin, Destination, and Year. Standard errors for Fixed Effects A are clustered at the origin-destination level and at the origin country level for Fixed Effects B. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.4: **Own and Cross-Price Elasticities - Proxy Rates**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:		Log Air Value			Log Sea Value	
Air Rate (Proxy)	-0.245***	-0.261***	-0.243***			
	(0.0168)	(0.0214)	(0.0380)			
Sea Rate	0.00517	-0.0000403	-0.00863			
	(0.00541)	(0.00784)	(0.0132)			
Air Value/Weight		0.0669***				
		(0.00973)				
Air Rate				-0.153***	-0.144***	-0.148***
				(0.00713)	(0.00992)	(0.0194)
Sea Rate (Proxy)				-0.304***	-0.315***	-0.419***
				(0.0225)	(0.0257)	(0.0690)
Sea Value/Weight					-0.0933***	
					(0.0113)	
Fixed Effects	A	A	B	A	A	B
N	1644286	163542	1644583	1656316	164758	1656622
r2	0.367	0.392	0.336	0.324	0.351	0.291

Notes: This table reports estimates of equation (3.15) for air and ocean imports. Air [Sea] Rate are log freight rates per kg. Proxy Rates are the average freight rate within the same trade route and industry excluding that observations freight rates. Value/Weights are log value to weight ratios. Fixed Effects A include HS6, Origin-Destination, Origin-Year, Destination-Year. Fixed effects B include HS6, Origin, Destination, and Year. Standard errors for Fixed Effects A are clustered at the origin-destination level and at the origin country level for Fixed Effects B. *** p<0.01, ** p<0.05, * p<0.1.

discounts for bulk shipping. For each observation I use the average of all other freight rates within the same industry (HS2 level), year, and trade route, excluding the specific HS6-trade route-year observation in calculating its own proxy freight rate. Estimates using these proxy freight rates are in Table 3.4. This finds a similar pattern of point estimates close to zero for the cross-price elasticity of air imports, and a negative and statistically significant cross-prices elasticity for ocean imports.

The cross-price elasticities estimated in Tables 3.3 and 3.4 show that air and ocean transport are not in aggregate substitutes. The negative cross-prices may indicate that imports typically carried by air have little substitution with ocean trade, where as imports typically carried by ocean are to a degree responsive to air freight rates. This is consistent with complementarity across transport modes. Such complementarity across transport modes could arise through complementarity in intermediate inputs. Based on their value-to-weight ratio and preferences over delivery time, some of these complementary inputs may typically be imported by air and others by sea. Firms facing fixed costs in sourcing complementary inputs may then source the entire bundle of goods from a new origin in response to a rise in a single modes freight costs, instead of substituting imports across modes. Similarly, fixed costs for setting up supply chains may lead firms to seek the same supplier for multiple items, or suppliers located geographically close together, even if individual items are not complementary goods. Conditional on ocean freight rates, increases in air freight rates may then lead firms to source all inputs from an alternative origin. This could cause the complementarity observed across modes. The substitutability of air and ocean transport will depend on the time frame in question. It is likely that in the short run of several months, air and ocean would exhibit little substitution due to pre-paid contracts, the relatively long time required to ship by sea, and difficulty of redirecting multi-modal freight that is already in transit. In the long run of many years to decades, supply chains and technology will adapt, leading to a likely larger substitutability. The time scale of observations used in this section - annual - results in estimates corresponding to a mid-range scale of several years where infrastructure remains mostly fixed.

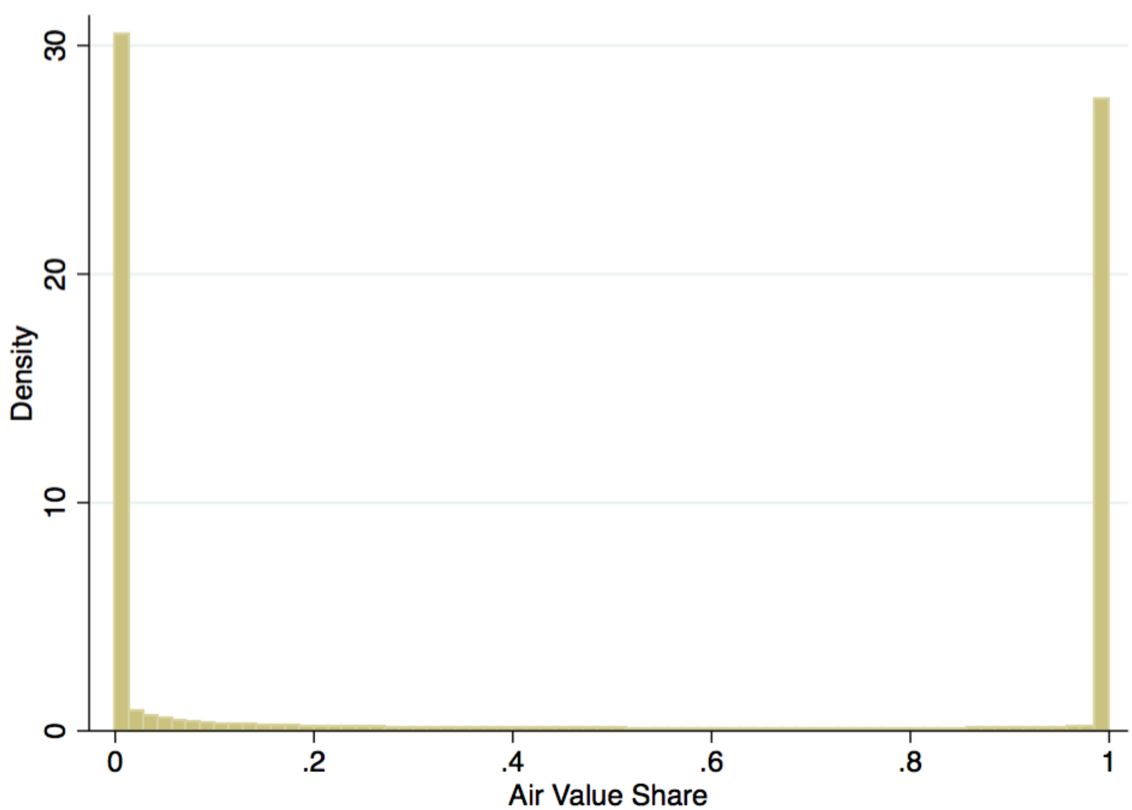
3.5 Heterogenous Unit Values and Mode Choice

The lack of strong substitution between transport modes found above–and the finding of complementarity–contradict the expectation of substitution implied by the overlap in products carried by both modes. As a result, the overlap noted by Hummels and Schaur (2010), Lux (2011), and Hummels and Schaur (2013) is important to reconsider. This overlap is also pertinent to the choice of transport mode because, as I show below, it may result from unobserved heterogeneity in product quality—which in turn is highly correlated with the across-country share of import value transported by a specific mode.

The overlap in products carried by both modes is demonstrated in Figure 3.1. This histogram shows the air value shares within detailed HS10-exporter-import district-year observations. These observations are concentrated at air shares close to 0 and 1, and the overlap in imports is relatively uniformly distributed among those products arriving by both modes. HS10-exporter-import district-year observations with both modes represent 19% of observations and 46% of total import value. However, many products may be predominantly moved by one mode while having small quantities shipped by the alternative mode, such as scientific ore specimens arriving by air when nearly all ore imports arrive by bulk ocean carriers. For HS10-exporter-import district-year observations with air value shares between 5% and 95%, the overlap between air and ocean transport falls to 13% of products representing 19% of value. 19% of value regularly carried by each mode within a trade route is not an insignificant fraction of imports. Implicit in the association between imports within HS10-year-trade routes arriving by both modes and potential substitution between modes is that, at the point of arrival to the importer or consumer, products are in fact closely similar varieties that serve as substitutes in the hands of a consumer or importing firm. However, inspection of the product categories in the U.S. Imports of Merchandise data suggests there is scope for substantial heterogeneity in varieties of products even within the most detailed HS10 product level recorded by customs.

For example, HS code 8517.12.0000 corresponds to "*Telephones for cellular networks or for other wireless networks*" (Census, 2018). This includes everything from cheap flip phones to expensive smart phones. Similarly, HS code 6110.11.0010 corresponds to *Sweaters, pullovers, sweatshirts, waistcoats (vests), and similar articles knitted or crocheted: Of wool: Men's or boys'*. While this is a detailed description, men's wool sweaters and sweatshirts vary widely in cost for a similar weight garment and may serve different markets. If expensive sweaters tend to arrive by air and low cost sweatshirts by ocean, then the arrival of imports across both modes within a HS10 code may not be indicative of a high degree of potential substitutability across modes. Instead, it reflects the aggregation over heterogeneous product varieties. Heterogeneity in quality across varieties would manifest itself as variation in unit prices within detailed product categories. Unit prices are not observed. Instead, I use the ratio of

Figure 3.1: **Histogram of Air Shares**



Notes: Density of air value shares within HS10-origin country-importing customs district-year observations.

free-on-board value to weight

$$\frac{Value}{Weight} = \frac{p * q}{w * q} = \frac{p}{w}$$

where value is a product's unit price in the origin country times quantity, and weight is a product's unit weight w times quantity.[64] This is a unit price where the unit is a kilogram of the product rather than a single item of unknown weight. Figure 3.2 plots the log of FOB value to weight ratios across HS10-year-trade route observations. Air shipping has substantially higher average FOB values per kilogram with a median of \$61/kg for air and \$7.0/kg for sea. There is also substantial heterogeneity in value to weight ratios within modes. However, this heterogeneity in part reflects product composition at the HS10 level. To control for product composition, I regress the log of value to weight on HS10 product α_p and year γ_t fixed effects,[65]

$$log\, value/weight_{nipt} = \alpha_p + \gamma_t + \epsilon_{nipt}$$

This controls for all value to weight variation that is due to observed product composition at the HS10 level or aggregate changes over time. Figure 3.3 plots the histogram of residuals from this estimation. This shows air has a higher median value/weight while there remains substantial heterogeneity within detailed product categories.

A limitation to this comparison is that value-to-weight ratios may reflect discounts from bulk purchasing even within identical product varieties. Bulk purchasing, referring to the quantity purchased by an individual firm from a particular supplier within some time period, is not observed but is likely higher among ocean imports. However, a proxy for bulk purchases can be constructed as the weight in kilograms per individual shipment listed on an individual customs declaration form. A limitation to this measure is it will not capture discounts from bulk purchases that are split across separate shipments. Data limitations require constructing the measure of bulk purchases at the most detailed observation level available, HS10-exporting country-month-importing district-unloading port. A measure of bulk purchases can be constructed for 75% of import value.[66] To control for bulk shipments in value to weight differences between modes, I estimate the following,

$$\begin{aligned} log\, value/weight_{nipt} = M_{nipt} + M_{nipt} \times WgtCard_{nipt} + (1 - M_{nipt}) \times WgtCard_{nipt} \\ + \eta_n + \delta_i + \gamma_t + \alpha_p + \epsilon_{nipt} \end{aligned} \tag{3.16}$$

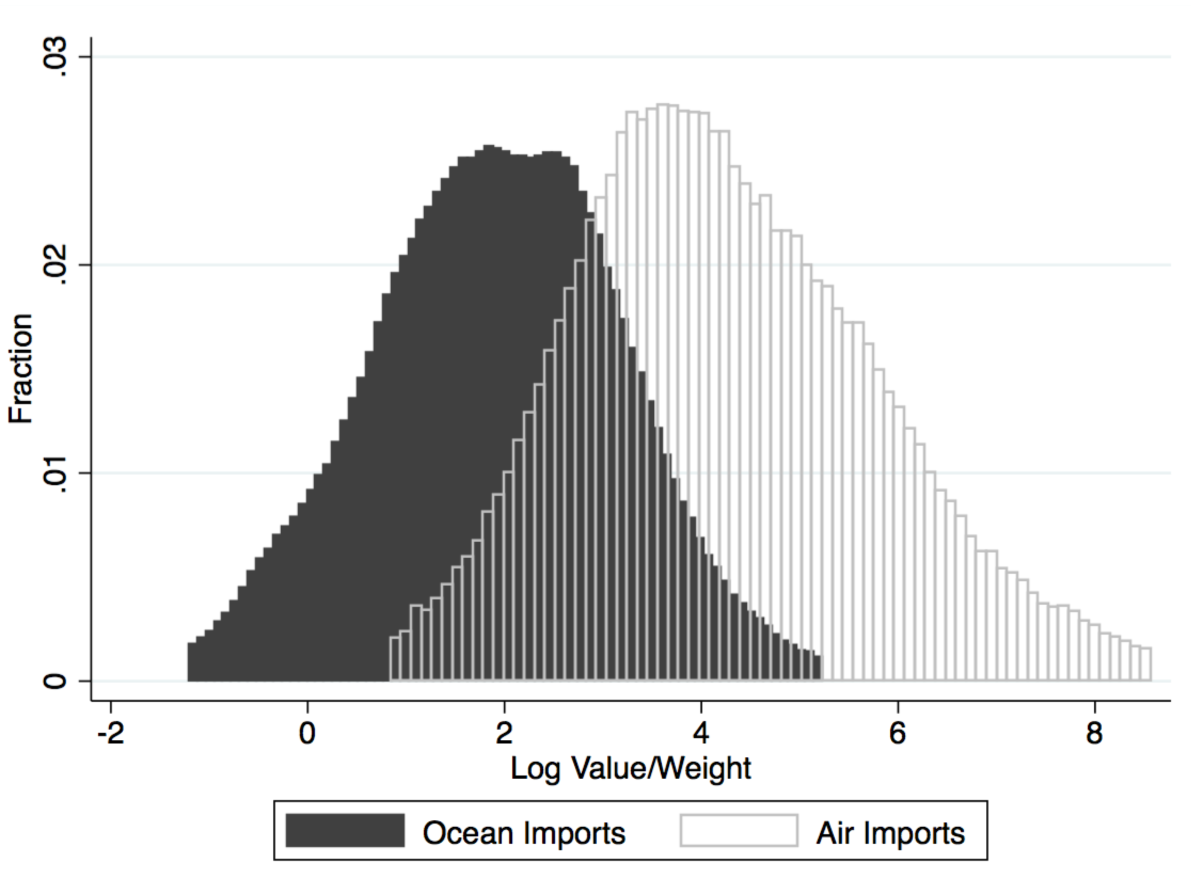
where M_{nipt} is an indicator for an observation arriving by air transport, $WgtCard_{nipt}$ is the log of weight per shipment, and remaining terms are origin country, import district, year, and HS10 product

[64] Free-on-board value excludes shipping charges. The U.S. Customs Bureau on FOB values: "*this value is generally defined as the price actually paid or payable for merchandise when sold for exportation to the United States, excluding U.S. import duties, freight, insurance and other charges incurred in bringing the merchandise to the United States.*"Census (1996)

[65] This has close similarities to an estimate by Harrigan (2010), but where I explore differences across modes while he is interested in the effect of distance.

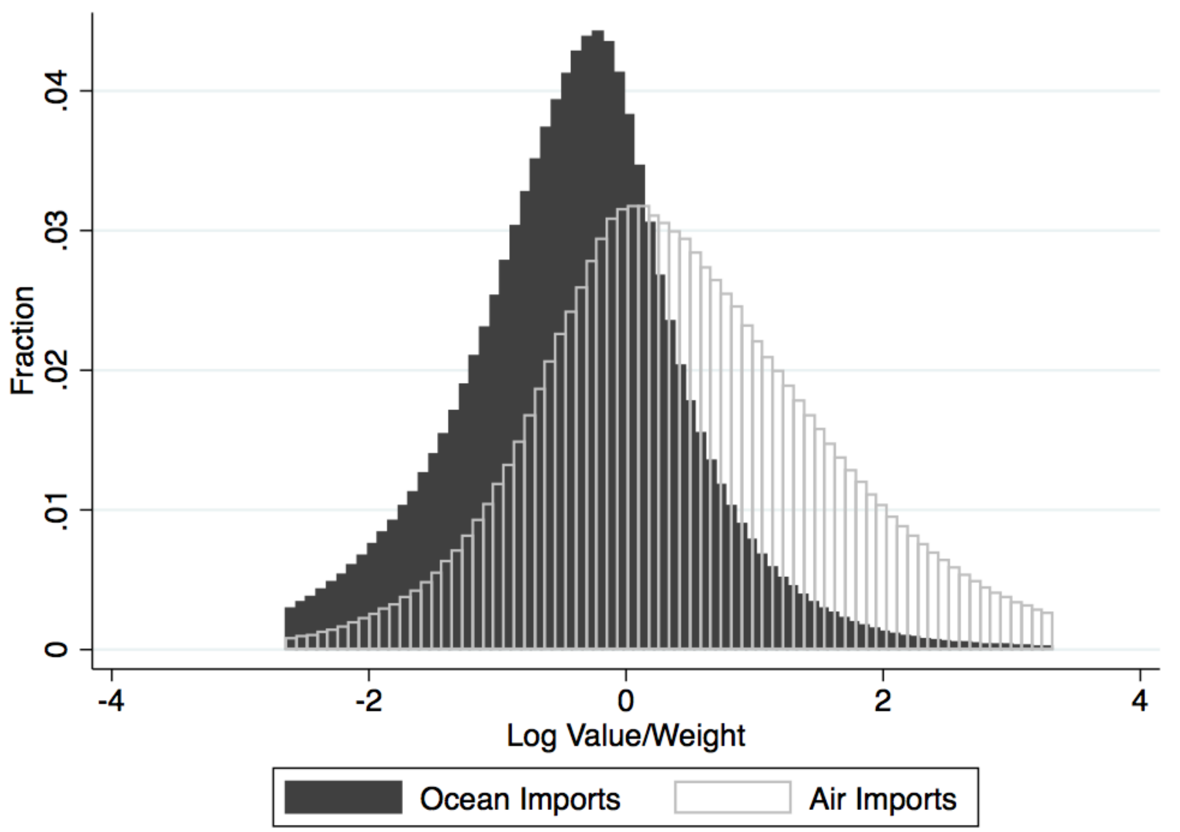
[66] As high-value imports are less likely to be assigned to individual shipments this measure of bulk purchasing has a selection bias towards those products with lower total import value. Details are available in Appendix C.3.

Figure 3.2: **Log Value/Weight Ratios**



Notes: Log of value to weight ratios within HS10-year-origin country-import district observations.

Figure 3.3: **Log Value/Weight Ratio Residuals**



Notes: Residual value to weight ratios within HS10-year-exporting country-import district observations after regressing log value to weight ratios on HS10 and year fixed effects.

Table 3.5: **Differences in Value/Weight Across Modes**

	(1)	(2)	(3)	(4)	(5)
Dependent variable:			Log Value per Weight		
Air	2.319***	2.087***	1.107***	1.078***	0.331***
	(0.0699)	(0.0685)	(0.0460)	(0.0599)	(0.0433)
Ocean X Shipment Weight				-0.361***	-0.489***
				(0.00883)	(0.0186)
Air X Shipment Weight				-0.500***	-0.482***
				(0.0146)	(0.0133)
N	4155978	4155978	4154759	3642211	1048027
r2	0.383	0.427	0.712	0.839	0.621
Year	✓	✓	✓	✓	✓
Country		✓	✓	✓	✓
District		✓	✓	✓	✓
HS10			✓	✓	✓

Notes: Estimates of equation 3.16. Fixed effects for year, origin country, import district, and HS10 product are included where indicated. Air is an indicator for the transport mode being air; ocean is the excluded reference category. Shipment weight is the log of weight attributed to that modes shipment as declared on a customs import form. This table reports *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

fixed effects. Table 3.5 presents the estimates. Specifications (1) and (2) show that without controlling for HS10 product composition air freight is on average over 209% more valuable per kilogram than sea freight. After controlling for HS10 product composition, specification (3) finds a 111% increase for air transport. Specifications (4) and (5) control for bulk purchasing through the weight per shipment. This finds that imports by both air and ocean have substantially lower unit values the larger the shipment size, consistent with discounts for bulk purchasing. In Specification (5) I limit the sample to individual shipment sizes between the 25th and 75th percentile of air shipment weight, 18 to 230 kilograms, to compare shipment sizes across ocean and air that can easily be carried within a single airplane flight or shipping container.[67] This estimate finds air shipments continue to have a 30% price premium. These differences in the value-to-weight ratio support that, while air and ocean transport specialize in different products, there remains substantial heterogeneity in product varieties within ten digit HS product categories and across transport modes.[68]

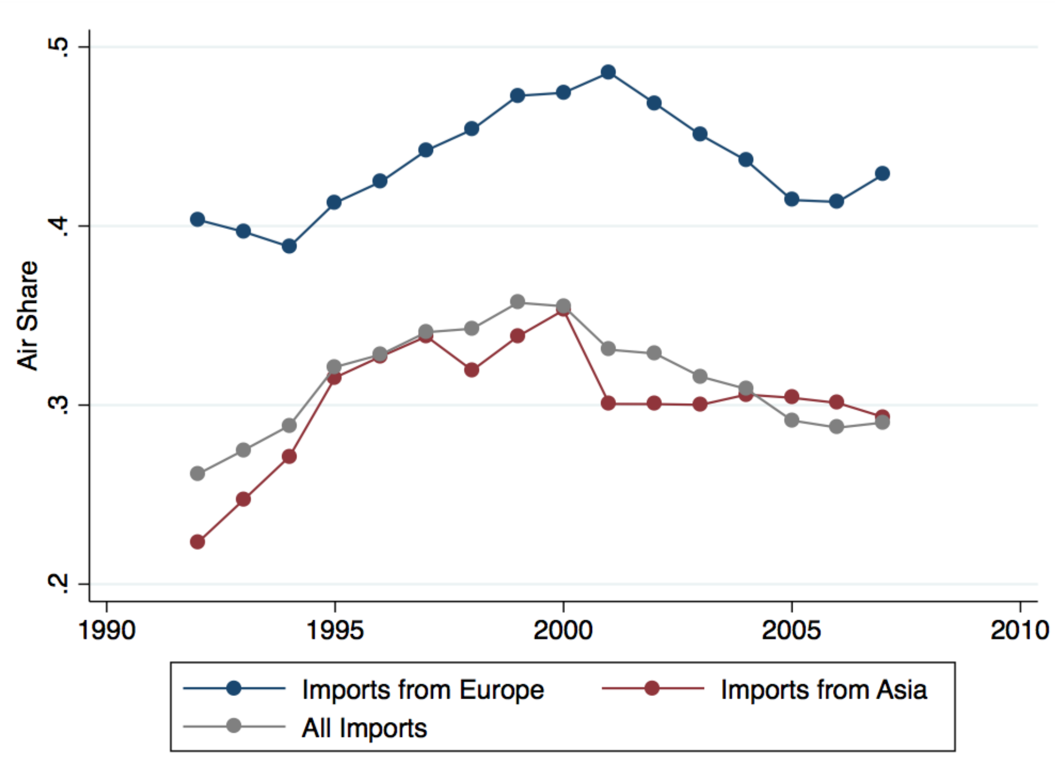
3.5.1 Differences in Air Shares Across Countries

How the share of imports carried by air within a bilateral trade route varies across countries suggests that unobserved product quality is important to the choice of mode, particularly in comparison to the importance of distance. Figure 3.4 plots the average air value shares for all U.S. Imports arriving from overseas, and separately for imports from Asia and Europe. Despite the large growth in U.S. imports over this period, the aggregate air share has not changed substantially. To explore the persistent differences in air shares and the role of product composition, I focus on the top 15 Asian and European

[67] 230 kilograms is also the maximum weight allowed for 10 checked bags on Air Canada.

[68] This point was partially noted by Harrigan (2010); "This great range of log unit values is suggestive of substantial heterogeneity even within narrowly-defined HS10 categories." However, Harrigan (2010) does not note that this heterogeneity differs systematically across transport modes.

Figure 3.4: **Trends in Air Value Shares**



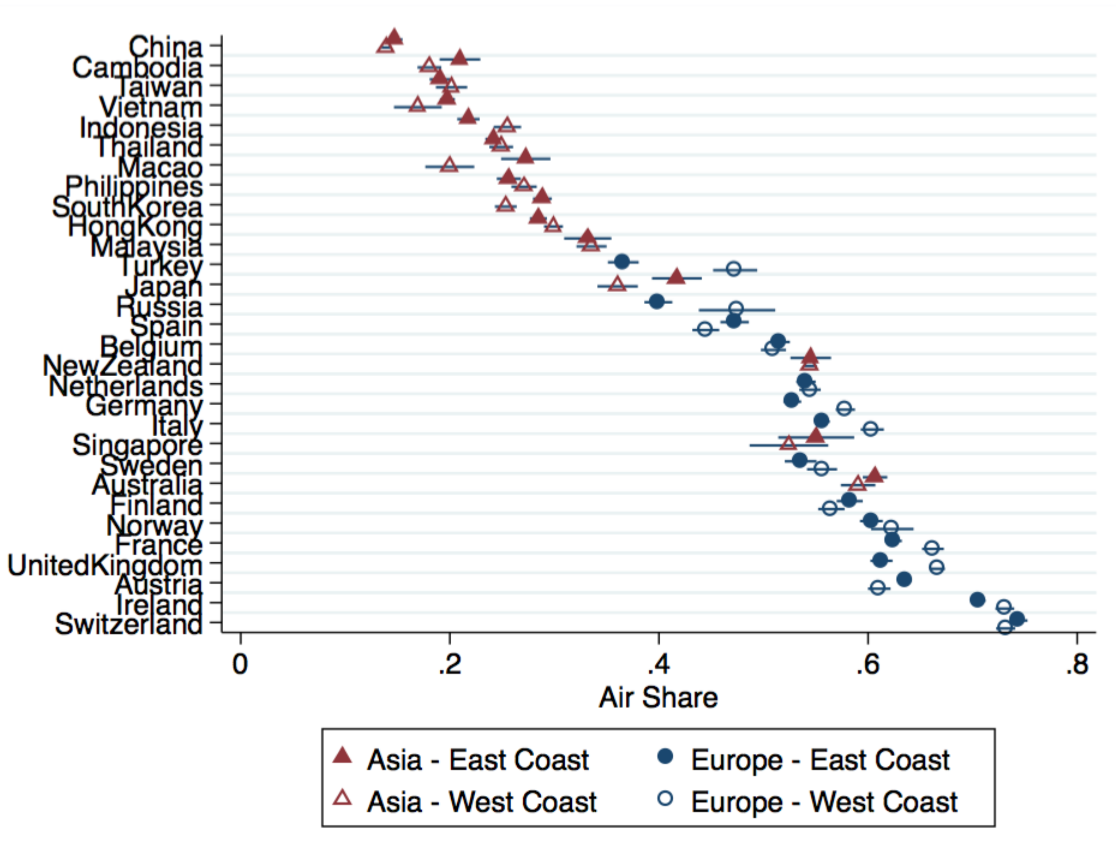
Notes: Time trends in average annual air value shares. *All Imports* includes all countries in addition to European and Asian countries.

countries by total exports to the U.S., along with U.S. import districts located on the East Coast and West Coast of the United States. These coastal districts simplify the comparison between imports by air and ocean freight as ports and airports are located close together, and because the East Coast and West Coast introduce a wedge in relative transit distance between air and ocean freight due to the necessity for ships to transit the Panama Canal. The average distance by air from countries in Asia to the East Coast of the U.S. is 20% farther than to the West Coast, but is 79% farther by sea. In contrast, for European countries the West Coast is on average 49% farther from the East Coast by air and 138% farther by sea.[69]

Figure 3.5 plots the average air value shares within country-U.S. coast trade routes and orders countries by their average air shares. Three patterns are of note. First, air shares differ greatly across countries. For example, China has an average air share of 14% compared to the United Kingdom at 62%. Second, average air shares within countries differ little between coasts of the U.S. despite the differences in distance. This suggests that differences in shipping time across coasts and countries is a relatively unimportant determinant of air shares compared to an unobserved factor that varies across countries and is correlated within countries. Third, consistent with Figure 3.4, countries with

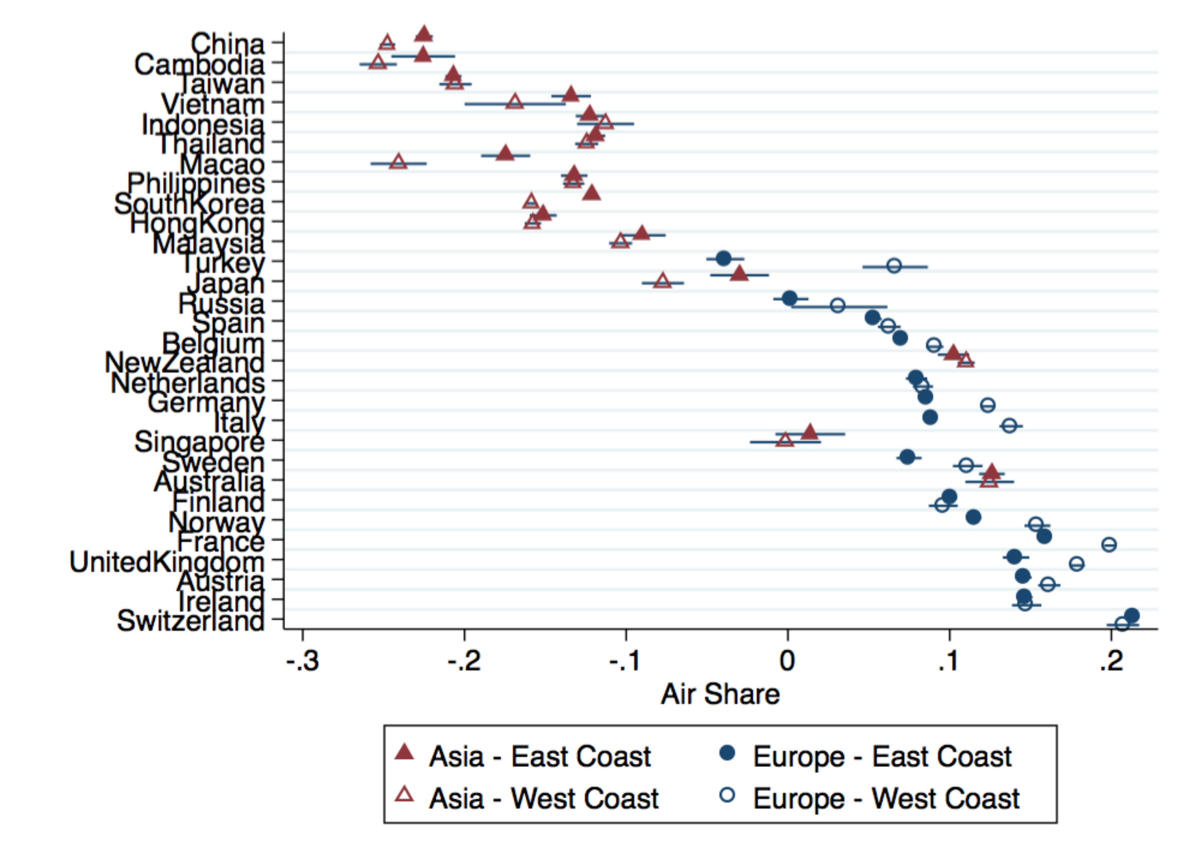
[69] Air distances are measured using Google Maps from countries' capital cities to the major port in the U.S. district of entry. Sea distances account for geography, such as the Panama canal, and are measured using Sea-distances.org from the largest port of export from a country to the largest port in the U.S. district of entry. Continent-Coast distances are the trade value weighted averages across country to import district trade routes.

Figure 3.5: **Average Air Shares By Country**



Notes: Air value shares averaged within country-coast trade routes. Countries are ordered by their average air shares, and coasts refer to the East Coast and West Coast of the United States. 95% confidence intervals on the mean air share are shown.

Figure 3.6: **Average Residual Air Shares By Country**



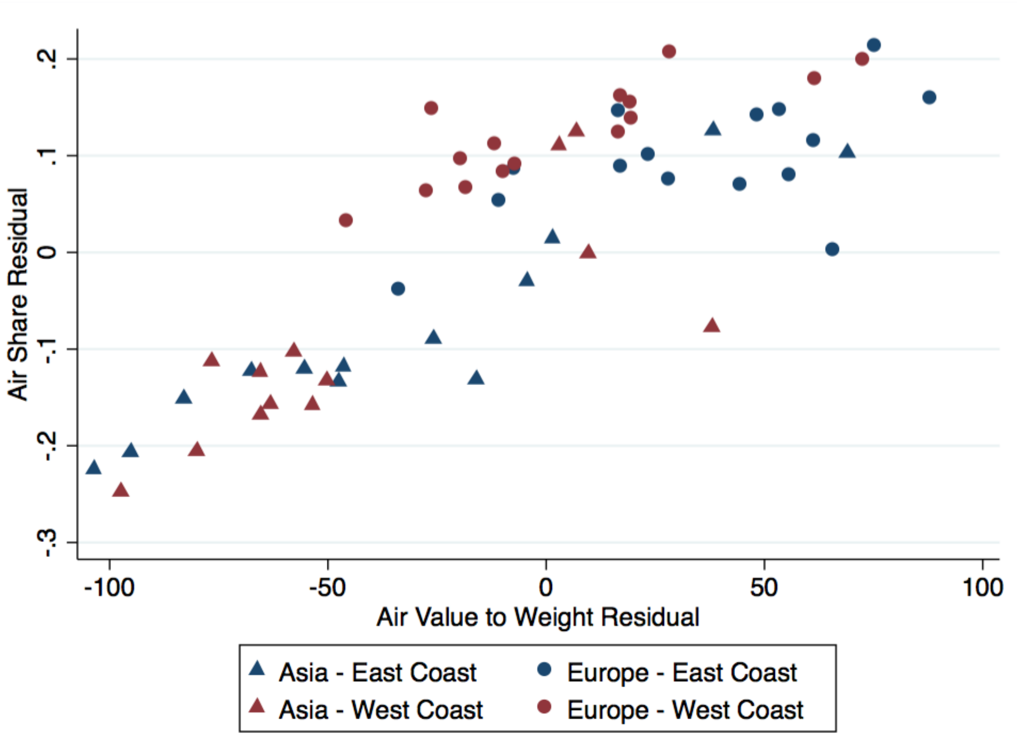
Notes: Residual air value shares averaged within country-coasts. Countries are ordered by their average air value shares as in Figure 3.5. 95% confidence intervals on the mean of the residual air share are shown.

high air shares are concentrated in Europe. Notable exceptions in Asia are Australia, Singapore, and New Zealand. This is consistent with the correlation of per capita GDP with air shares noted by Lux (2011).

However, this comparison of air shares does not control for product composition. To control for HS10 product composition, I regress air shares for HS10-exporter-district-year observations on HS10 product and year fixed effects and predict the residuals. These residual air shares are orthogonal to any determinant of air shares that is common within 10-digit product categories. In Figure 3.6 I plot the air share residual averaged within country-coast and with the same country order as Figure 3.5. Even after controlling for product composition at the HS10 level, the three patterns remain closely similar to the original patterns in Figure 3.5. While the range of average air shares across countries has compressed slightly from 0.59 in Figure 3.5 to 0.44, average air shares still differ substantially across countries compared to across coasts within a country.

As previously discussed, heterogeneous product quality varies across transport modes and could be one such factor that is correlated within countries. To explore this, Figure 3.7 plots the residual air shares against the residual air value-to-weight ratios averaged within country-coast trade routes. A strong correlation is clear: after controlling for product composition, the air shares across countries are highly correlated with residual unit values. Countries that produce high quality product varieties

Figure 3.7: **Residual Air Shares and Value to Weight Ratios**



Notes: y-axis values Air Share Residual and x-axis value to weight residuals are the residuals orthogonal to year and HS10 fixed effects. Both are averaged within exporting country-U.S. coast.

have a high value to weight ratio. With shipping costs being primarily per unit, rather than ad-valorem, air shipping is relatively less expensive compared to ocean shipping the higher the value of the product. This suggests that unobserved quality is a particularly important determinant of the choice of transport mode that future models of mode choice must account for. While previous models have focused on distance, the lack of strong dependence on shipping distance suggests that estimates requiring a parameterization of distance or shipping time may be sensitive to the functional form chosen. In addition, a focus on distance and time as the key determinant misses the important factor of unobserved quality.

3.6 Conclusion

This chapter shows how prior models that consider trade differentiated by the method of transportation impose strong, and potentially unrealistic, restrictions on substitution patterns across modes and countries. I show that existing models have treated bilateral trade by air and sea as both substitutes and complements, and that this has significant quantitative and qualitative consequences for counterfactual trade patterns. This finding has broad applicability; the elasticity of substitution between modes is important in evaluating the impacts of any factor that differs across methods of transportation, including distance, freight costs, delivery times, and environmental impacts. Turning to the empirical evidence, I exploit idiosyncratic variations in freight rates to estimate reduced form own-price and cross-price elasticities of substitution for imports arriving by ocean and air transport. These estimates find little evidence for substitution between transport modes for products primarily carried by air, and some evidence for complementarity between modes for imports typically carried by ocean transport. I then show that previous cross-sectional evidence which motivates the assumption of substitution across modes, the overlap in products arriving by both modes, may instead reflect unobserved heterogeneity in product quality. This unobserved quality is a particularly important, and so far unrecognized, determinant of the choice of transportation mode across countries.

These results have important implications for future work on models of mode-specific trade, the importance of product quality and distance, and for proposed policies addressing the direct environmental impacts of international trade. As this chapter showed, models of mode choice face a trade off between quantitatively accurate approximations of substitution patterns across countries and modes, and maintaining a tractable form that is consistent with the large theoretical and empirical literature on international trade. It is important that future work further establish a set of empirical facts that must be approximated in models of endogenous mode choice. These empirical patterns are also important to understand before microfounded models are developed to explain aggregate complementarity between modes, or incorporate unobserved heterogeneity in product quality.

Of particular importance are additional estimates of the elasticities of substitution between transport modes, both within bilateral trade and across countries. These estimates are necessary to establish how many separate substitution elasticities are necessary, and what analytically convenient approximations can be made. For example, the shared model framework considered in this chapter imposes that a single elasticity governs substitution between transport modes within bilateral trade, and bilateral mode-shares are independent of relative freight costs among other countries. An alternative framework could impose that a single elasticity governs substitution across countries within a transport mode, and a country's share of imports within each mode is independent of relative freight costs among alternative transport modes. This alternative substitution pattern may be a more accurate model for overall substitution patterns, or for particular classes of products such as those with a low elasticity of substitution across transport modes.

Similarly, key determinants of mode choice–particularly unit values–need to be systematically explored. Heterogeneity in unit values may matter in two ways. The first is different papers and regression specifications use observations aggregated to different product levels, in particular HS10, HS6, and HS2 (industry). The implications, if any, of these aggregations on empirical work considering trade differentiated by mode is unknown. The second is that while prior work has exploited variation in distance for identification and considered it as a source of comparative advantage, the importance of distance to the choice of mode, relative to other determinants, is not well understood. This chapter suggests that differences in distance between U.S. coasts is not a particularly important factor, and documented a strong correlation among air shares within European and Asian countries with product unit values. As a result, unobserved product quality is highly correlated with the choice of transport mode—and may in turn be correlated with distance from the United States—potentially confounding estimates exploiting variation in mode shares and distance. This has the potential to be a widespread problem since most work to date considering separate transportation modes in international trade uses the same

U.S. Imports of Merchandise dataset used in this chapter.

While a variety of regulations on the direct carbon emissions from international aviation and shipping sources have been proposed, these efforts have considered transport modes separately and do not consider the potential interactions between them. This chapter's findings that air and ocean transport are not aggregate substitutes—and the potential for them to be complements—has two important implications for such regulations. First, air and ocean not being substitutes indicates that there is no low-cost opportunity to shift a substantial share of trade from air to ocean to reduce negative pollution externalities without reducing imports. The silver lining is this also indicates that carbon emissions from international ocean shipping can be regulated largely without concern for inducing offsetting increases through substitution to air transportation.

Conclusion

This dissertation studies topics in environmental economics and international trade: how households conserve energy in response to financial rewards, and the choice of transportation mode within international trade.

In Chapter 1, I examine households' electricity use and their participation in a voluntary program offering financial rewards in exchange for achieving energy conservation targets. I estimate households' short and long run intensive margin responses and find that while electricity use declines as households attempt to achieve their conservation target, it rebounds as they leave the program. This rebound suggests households do not make persistent changes through new capital investments or habits. Importantly, I find no evidence that households respond strategically to the voluntary program design through their decisions when to begin participating.

In Chapter 2, I use the same dataset to study households' extensive margin decisions whether to continue participating in the energy conservation program. I find that the decision to re-enroll differs little across household characteristics, yet is strongly and discontinuously predicted by their success in achieving the conservation target—and not the structure of the incentive. Exploiting this discontinuity in the probability of re-enrolling using a fuzzy regression discontinuity design, I estimate the causal effect of re-enrolling. Consistent with the results of Chapter 1 I find that additional conservation challenges cause lower electricity use. In addition, households' sensitivity to their success—and not the incentive structure—is consistent with using heuristics in making participation decisions instead of responding as a well informed agent making an optimal choice given the structure of the incentive.

Chapter 3 considers the choice of transportation mode within international trade. This chapter finds that international trade models incorporating a choice of mode have imposed potentially unrealistic restrictions on possible substitution patterns across countries and modes. In particular, I show how models have implicitly imposed transport modes to be complements, while other models have implicitly imposed modes to be substitutes. I then show that this has a substantial impact on counterfactual trade patterns and policy recommendations for any mode-specific cost change. This chapter then contributes two pieces of empirical evidence on the choice of mode. I first undertake reduced form estimates of the elasticity of substitution between modes and find no evidence transport modes are substitutes. Instead, the evidence is consistent with a limited degree of complementarity with air for products typically carried by ocean transport. I then show how unobserved quality within detailed product categories is a substantial yet largely unrecognized determinant of the choice of mode.

Bibliography

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

Allcott, H., Rogers, T., 2014. The Short-Run and Long-Run Effects of Behavioral Interventions : Experimental Evidence from Energy Conservation. American Economic Review 104 (10), 3003–3037.

Angrist, J. D., Pischke, J. S., 2008. Mostly Harmless Econometrics : An Empiricist's Companion.

Bandura, A., 1977. Self-efficacy : Toward a Unifying Theory of Behavioral Change. Psychological Review 84 (2), 191–215.

BCH, 2014. British Columbia Hydro and Power Authority Annual Report: 2014/2015.

BCH, 2016. BC Hydro Greenhouse Gas Intensities: 2007-2015. Available online at https://www.bchydro.com/about/sustainability/climate_action/greenhouse_gases.html (accessed July 15, 2018).

BCStats, 2016. B.C. Regional Statistics: Population Estimates. Available online at http://www2.gov.bc.ca/gov/content/data/statistics/people-population-community/population/population-estimates (accessed July 15, 2018).

Boomhower, J., Davis, L. W., 2014. A credible approach for measuring inframarginal participation in energy efficiency programs. Journal of Public Economics 113, 67–79.

Borusyak, K., Jaravel, X., 2016. Revisiting Event Study Designs. Available online at http://ssrn.com/abstract=2826228 (accessed July 15, 2018).

British Columbia Utilities Commission, 2017. British Columbia Utilities Commission Inquiry Respecting Site C: Preliminary Report to the Government of British Columbia. Tech. rep.

Calonico, S., Cattaneo, M. D., Titiunik, R., 2014. Robust data-driven inference in the regression-discontinuity design. Stata Journal 14 (4), 909–946.

Census, 2018. Schedule B Commodity Classification. Available online at https://www.census.gov/foreign-trade/schedules/b/2018/index.html (accessed July 17, 2018).

Census, U., 1996. U.S. Exports and Imports of Merchandise on CD-ROM: Technical Documentation. Available online through Google Books at https://books.google.com.au/books?id=3k3IArTOcqUC (accessed July 15, 2018).

Corbett, J. J., Wang, C., Winebrake, J. J., Green, E., 2007. Allocation and forecasting of global ship emissions (August 2015), 26.

Cristea, A., Hummels, D., Puzzello, L., Avetisyan, M., 2012. Trade and the greenhouse gas emissions from international freight transport. Journal of Environmental Economics and Management 65 (1), 153–173.

Dolan, P., Metcalfe, R., 2015. Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation. Available online at: http://ssrn.com/abstract=2589269 (accessed July 15, 2018).

Eaton, J., Kortum, S., 2002. Technology, Geography, and Trade. Econometrica 70 (5), 1741–1779.

ECCC, 2017. Canadian Climate Normals 1981-2010. Climate ID:1108447. Available online at http://climate.weather.gc.ca (accessed July 15, 2018).

EPA, C., 2017. California GHG Emission Inventory: California Greenhouse Gas Emissions for 2000 to 2015 - Trends of Emissions and Other Indictators. Tech. rep., Air Resources Board.

Evans, C. L., Harrigan, J., 2016. American Economic Association Distance, Time, and Specialization: Lean Retailing in General Equilibrium 95 (1), 292–313.

Feyrer, J., 2009. Distance, Trade, and Income - The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment XXXIII (2), 81–87.

Gerard, F., Costa, F. J. M., 2015. Hysteresis and the Social Cost of Corrective Policies: Evidence from a Temporary Energy Saving Program. Working Paper.

Gilbert, B., Graff Zivin, J. S., 2014. Dynamic salience with intermittent billing: Evidence from smart electricity meters. NBER Working Paper Series (Working Paper 19510).

Gneezy, U., Meier, S., Rey-Biel, P., 2011. When and Why Incentives (Don't) Work to Modify Behavior. Journal of Economic Perspectives 25 (4), 191–210.

Grassl, H., Brockhagen, D., 2007. Climate forcing of aviation emissions in high altitudes and comparison of metrics. An update according to the Fourth . . . , 1–8.

Hallak, J. C., 2006. Product quality and the direction of trade. Journal of International Economics 68 (1), 238–265.

Hallak, J. C., Schott, P. K., 2011. Estimating cross-country differences in product quality. Quarterly Journal of Economics 126 (1), 417–474.

Harrigan, J., 2010. Airplanes and comparative advantage. Journal of International Economics 82 (2), 181–194.

Harrigan, J., Venables, A. J., 2006. Timeliness and agglomeration. Journal of Urban Economics 59 (2), 300–316.

Head, K., Mayer, T., 2015. Gravity Equations: Workhorse,Toolkit, and Cookbook. Vol. 4. Elsevier B.V.

Holmes, T. J., Singer, E., 2018. Indivisibilities in Distribution.

Hummels, D., Klenow, P. J., 2005. The Variety and Quality of a Nation ' s Exports. American Economic Review 95 (3), 704–723.

Hummels, D., Lugovskyy, V., Skiba, A., 2009. The trade reducing effects of market power in international shipping. Journal of Development Economics 89 (1), 84–97.

Hummels, D., Skiba, A., 2004. Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture. Journal of Political Economy 112 (6), 1384–1402.

Hummels, D. L., Schaur, G., 2010. Hedging price volatility using fast transport. Journal of International Economics 82 (1), 15–25.

Hummels, D. L., Schaur, G., dec 2013. Time as a Trade Barrier. American Economic Review 103 (7), 2935–2959.

Imbens, G. W., Lemieux, T., 2007. Regression discontinuity designs: A guide to practice. Journal of Econometrics 142 (2), 615–635.

IPCC, 2014. Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Tech. rep.

Ito, K., 2014. Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing. American Economic Review 104 (2), 537–563.

Ito, K., 2015. Asymmetric Incentives in Subsidies: Evidence from a Large-Scale Electricity Rebate Program. American Economic Journal: Economic Policy 7 (3), 209–237.

Ito, K., Ida, T., Tanaka, M., 2015. The Persistence of Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand. NBER Working Paper Series.

Jessoe, K., Rapson, D., 2014. American Economic Association Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use 104 (4), 1417–1438.

Kassirer, J., Korteland, A., Pedersen, M., 2014. Team Power Smart Sparks Increase in Low-Priority, Repetitive Behaviors. Social Marketing Quarterly 20 (3), 165–185.

Lee, D. S., Lemieux, T., 2010. Regression Discontinuity Designs in Economics. Journal of Economic Literature 48, 281–355.

Lux, M., 2011. Defying Gravity: The Substitutability of Transportation in International Trade. MPRA (36395).

Martin, S., Rivers, N., 2016. Information provision, market incentives, and household electricity consumption: Evidence from a large-scale field trial. Working paper.

McClelland, L., Cook, S. W., 1980. Promoting Energy Conservation in Master Metered Apartments through Group Financial Incentives. Journal of Applied Social Psychology 10 (1), 20–31.

McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics 142 (2), 698–714.

Melitz, M. J., Redding, S. J., 2013. Firm Heterogeneity and Aggregate Welfare.

Micco, A., Serebrisky, T., 2006. Competition regimes and air transport costs: The effects of open skies agreements. Journal of International Economics 70 (1), 25–51.

Midden, C. J., Meter, J. F., Weenig, M. H., Zieverink, H. J., 1983. Using feedback, reinforcement and information to reduce energy consumption in households: A field-experiment. Journal of Economic Psychology 3 (1), 65–86.

Mizobuchi, K., Takeuchi, K., 2012. Using economic incentives to reduce electricity consumption: A field experiment in Matsuyama, Japan. International Journal of Energy Economics and Policy 2 (4), 318–332.

Moreira, M. J., 2003. A Conditional Likelihood Ratio Test for Structural Models. Econometrica 71 (4), 1027–1048.

Schleich, J., Faure, C., Klobasa, M., 2017. Persistence of the effects of providing feedback alongside smart metering devices on household electricity demand. Energy Policy 107 (April), 225–233.

Schott, P. K., 2004. Across-Product Versus Within-Product Specialization in International Trade. The Quarterly Journal of Economics 119 (2), 647–678.

Sexton, S., 2015. Automatic Bill Payment and Salience Effects: Evidence From Electricity Consumption. Review of Economics and Statistics 97 (2), 229–241.

Shapiro, J. S., 2016. Trade Costs, CO2, and the Environment. American Economic Journal: Economic Policy 0 (0), 220–254.

Sillard, P., Wilner, L., 2015. Constant utility index and inter-month substitution. Economics Bulletin 35 (3), 1772–1781.

Simonovska, I., Waugh, M. E., 2014. The elasticity of trade: Estimates and evidence. Journal of International Economics 92 (1), 34–50.

Strotz, R. H., 1957. The Empirical Implications of a Utility Tree. Econometrica 25 (2), 269–280.

Wichman, C. J., 2017. Information provision and consumer behavior: A natural experiment in billing frequency. Journal of Public Economics 152, 13–33.

Winett, R. A., Kagel, J. H., Battalio, R. C., Winkler, R. C., 1978. Effects of monetary rebates, feedback, and information on residential electricity conservation. Journal of Applied Psychology 63 (1), 73–80.

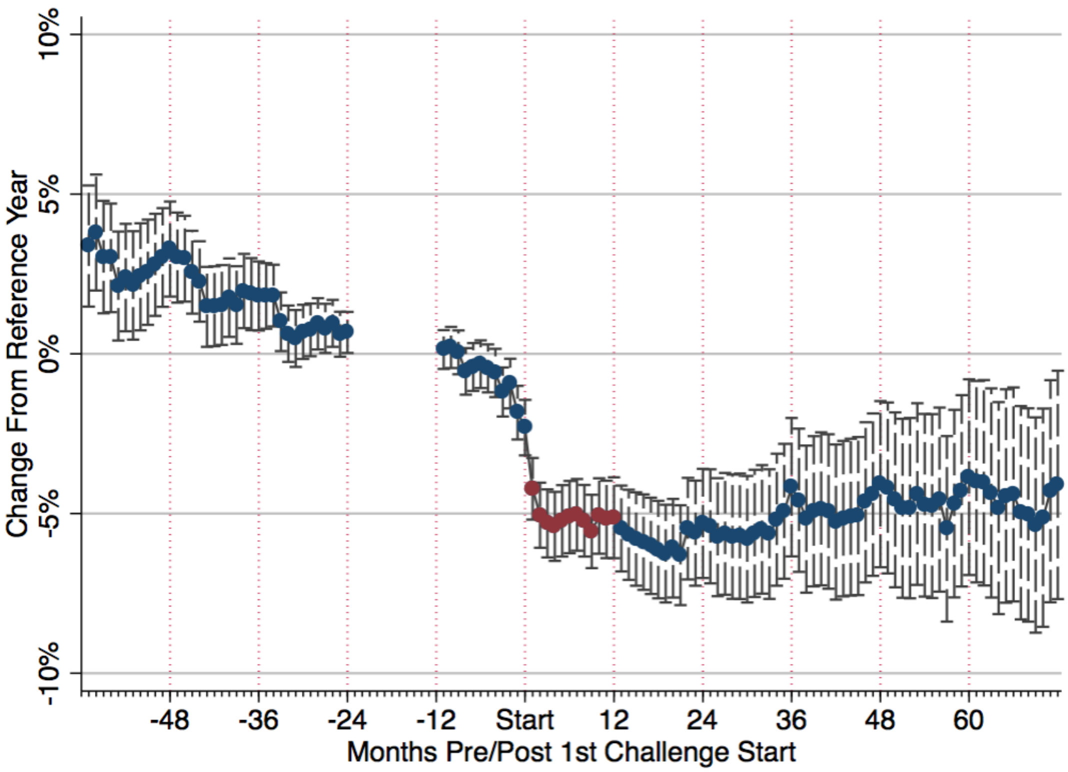
Appendix A

Appendix to Chapter 1

A.1 Event Study Estimates: Robustness Checks

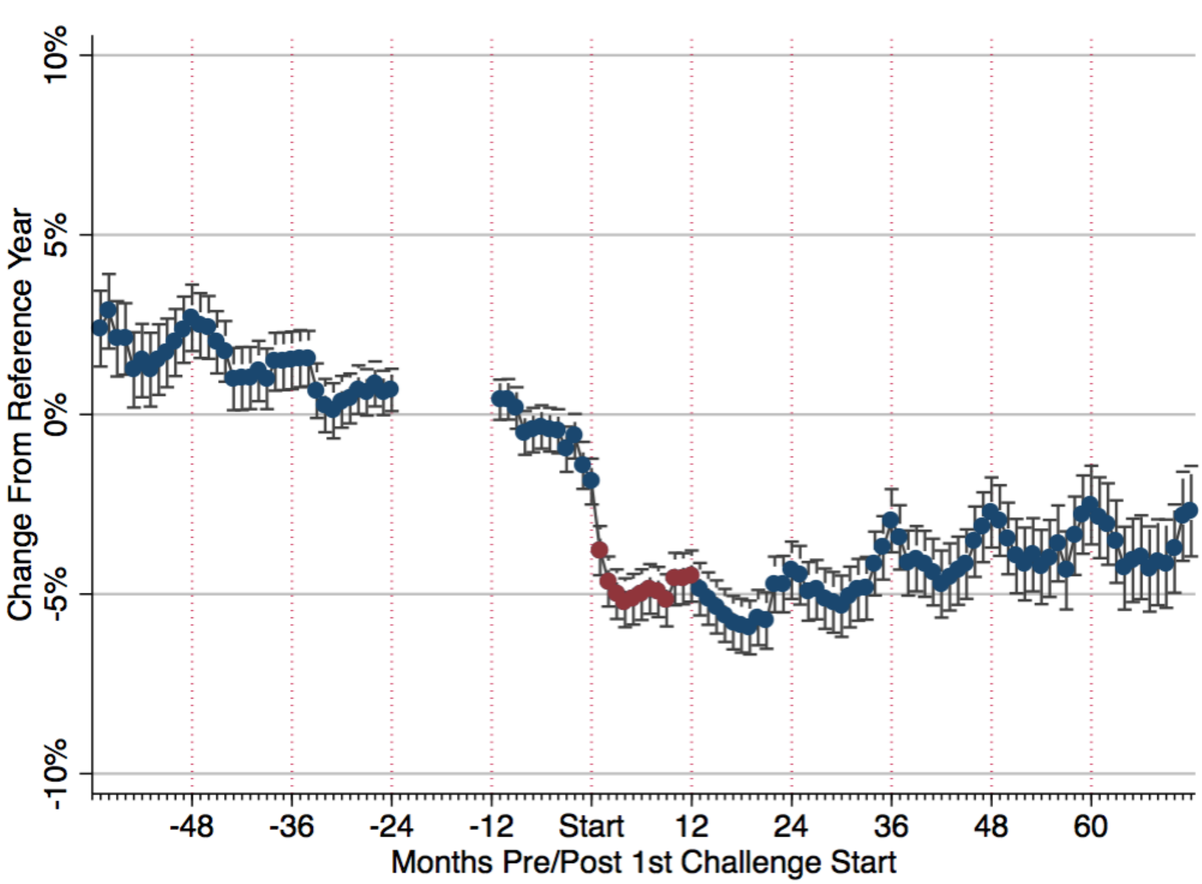
This section presents several event-study robustness checks. Figure A.1 plots the event study estimates of equation (1.1) estimated without non-participant control households. This estimation strategy identifies the program effects by exploiting the variation in timing in when a household starts their first conservation challenge. Households that start a challenge later in the panel serve as the control population for households that undertake a challenge earlier in the panel. Figure A.1 shows that the estimated pre-treatment trend is not due to diverging trends between participant and non-participant households. Confidence intervals are larger due to the lack of non-participant households for identifying the date fixed effects. Figure A.2 plots the estimated program effects including the full set of non-participant households. Figure A.3 plots estimated program effects using an alternate baseline period of the third year prior to the initial conservation challenge.

Figure A.1: **Estimated Treatment Effects For Participant Households Only**



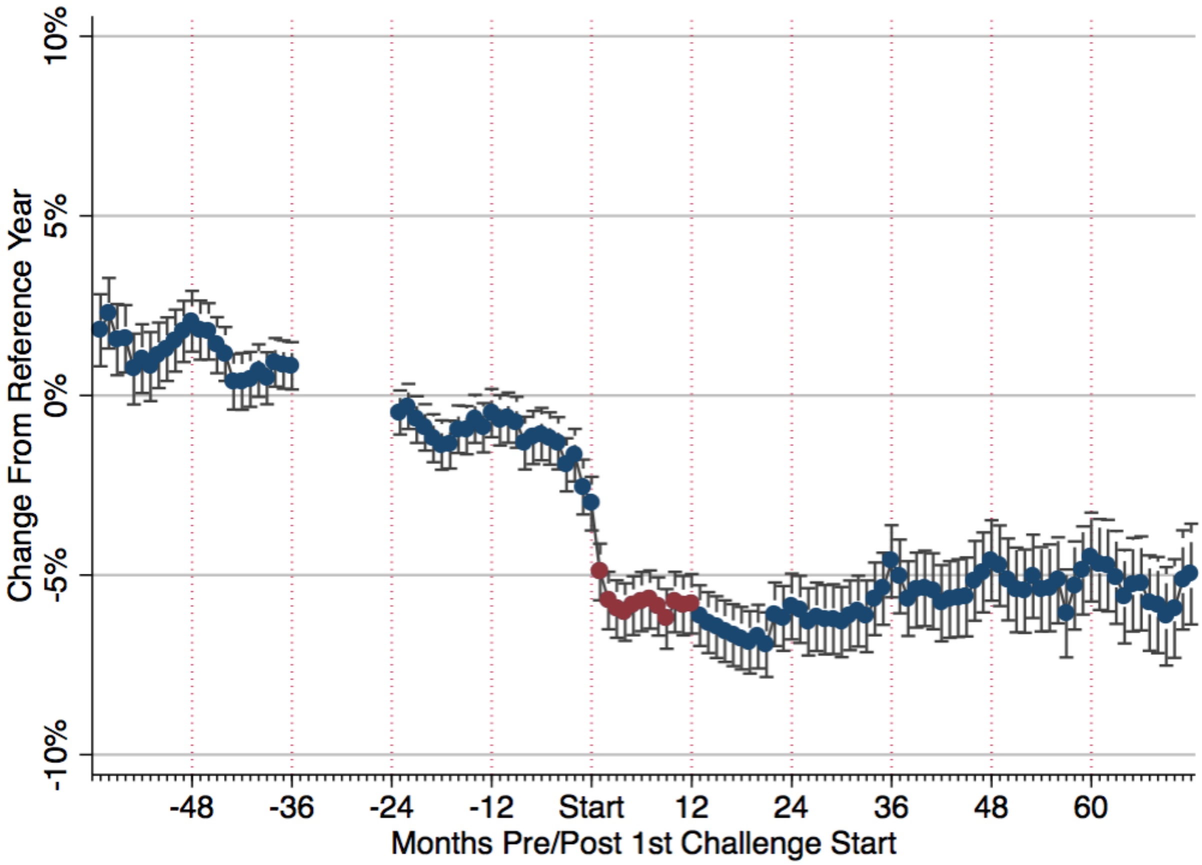
Notes: This figure plots estimates of $\hat{\beta}_\tau$ and 95% confidence intervals from specification (1.3) estimated for participant households only. Estimates $\hat{\beta}_\tau$ are ordered by event-time τ and point estimates in red denote the 12 months of the initial conservation challenge. The gap between -11 and -23 is the excluded reference period; $\hat{\beta}_\tau$ identifies the percent change in electricity use relative to the average use in this period. Point estimates are in Appendix Table D.1, specification (2).

Figure A.2: **Estimated Treatment Effects For All Households**



Notes: This figure plots estimates of $\hat{\beta}_\tau$ and 95% confidence intervals from specification (1.3) for the unbalanced set of participant and non-participant households. Estimates $\hat{\beta}_\tau$ are ordered by event-time τ. Point estimates in red denote the 12 months of the initial conservation challenge ($\tau = [1..12]$). The additional yearly variation in this figure, compared to Figure 1.6, arises due to the higher share of electric-heating households among the full set of non-participant households, compared to program participants. Electric heating households have higher seasonal variation than non-electric households which the common date fixed effects cannot fully absorb.

Figure A.3: **Estimated Treatment Effects For All Households — Alternate Baseline**



Notes: This figure plots estimates of $\hat{\beta}_\tau$ and 95% confidence intervals from specification (1.3) for all participant and non-participant households. Estimates $\hat{\beta}_\tau$ are ordered by event-time τ. Point estimates in red denote the 12 months of the initial conservation challenge ($\tau = [1..12]$). The pre-treatment period is denoted by the months prior to *Start* ($\tau \leq 0$). The visual gap in estimates between months $\tau = -24$ and $\tau = -36$ is the excluded reference period. $\hat{\beta}_\tau$ identify the percent change in electricity use relative to the average electricity use within a household during this excluded reference year.

Appendix B

Appendix to Chapter 2

B.1 Selection Into a Second Conservation Challenge

Figure B.1 plots the probability of continuing to a second conservation challenge against the reductions in billed electricity use from that household's first challenge. Larger reductions in billed electricity use are associated with a greater likelihood of continuing to a subsequent challenge. Figure B.1 also shows the fraction of households succeeding in their challenge. From this we can see some households with reductions greater than 9.5% do not pass their challenge, while other households with reductions less than 9.5% do pass. This occurs because success or failure in a challenge is evaluated from changes in weather-adjusted - not billed - electricity use.

B.2 Continuity at the Discontinuity

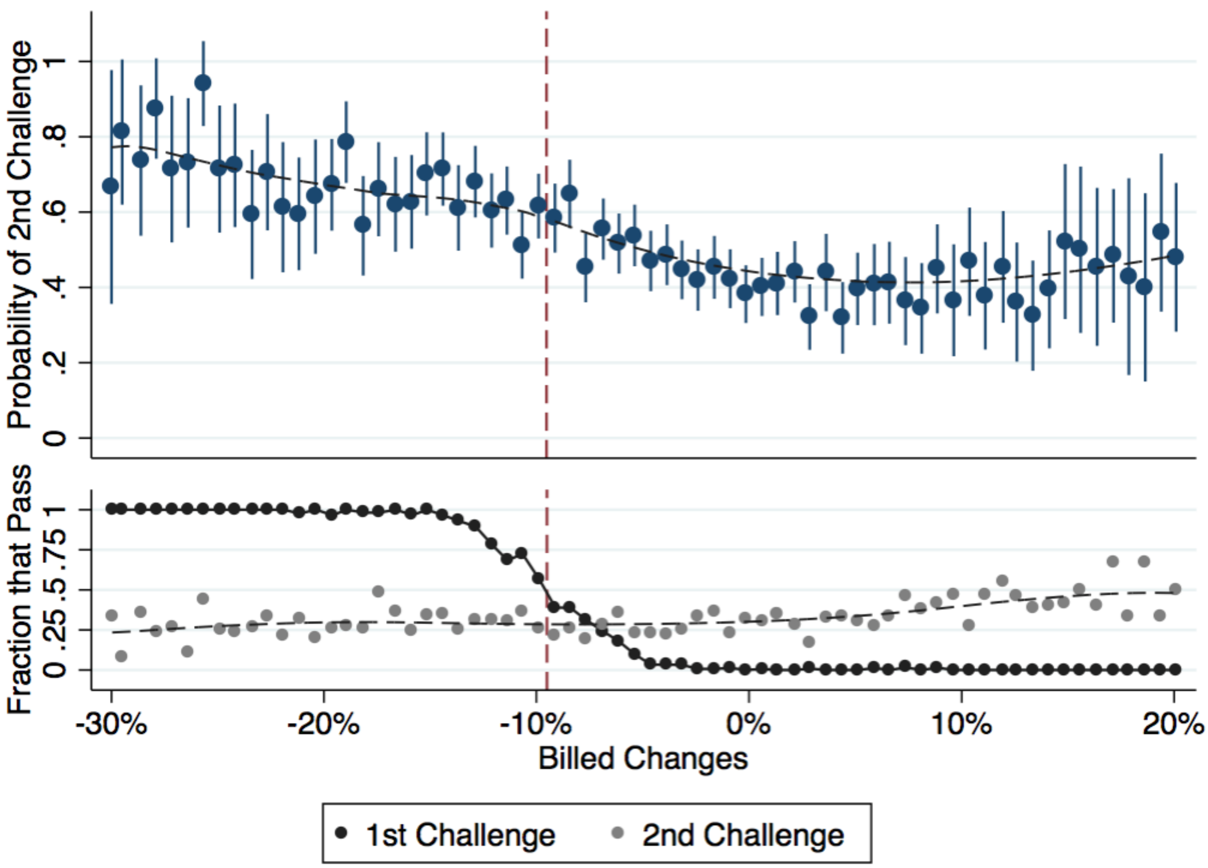
Figure B.2 shows the McCrary (2008) density test at the 9.5% threshold. This test fails to reject the null hypothesis of no sorting, supporting that if households are bunching in success, they are not doing so at the 9.5% threshold and instead only at the 10% target.

In Figure B.3 I plot the average of four household observables across the 9.5% threshold. This visually shows no sorting of households by observables around the discontinuity. Table B.1 shows the results of a linear regression discontinuity model estimated for these four and two additional observables at the threshold. This finds no statistically significant change in the density of observables and supports identifying assumption that households are not sorting around the discontinuity.

The McCrary (2008) density test, Figure B.2, suggests that there may be an increased density of households above the 10% conservation target. In a standard regression-discontinuity setup, this could indicate sorting around a threshold lead to concerns that the RD estimates may be biased by self-selection into success. However, the threshold for success in this fuzzy-RD setup is at a reduction of 9.5% — not the 10% target. If households were precisely sorting around the 10% target it is reasonable to expect this to appear as a discontinuity in either, or both, of the probability of continuing and the post-program outcome. Inspection of Figure B.4 indicates this is not happening; the only discontinuity occurs at precisely the 9.5% threshold.

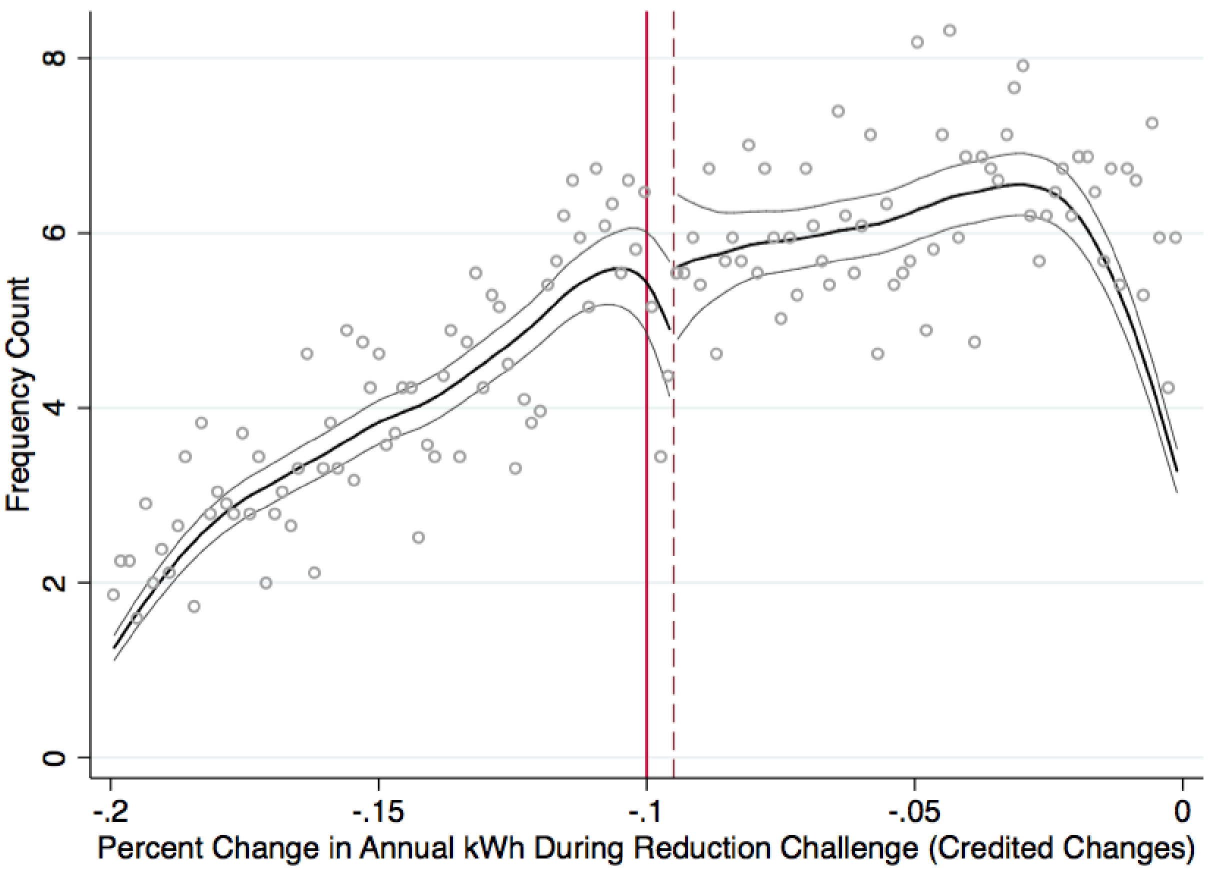
B.3 Fuzzy-RD Robustness Checks

Figure B.1: **Probability of Continuing to a Second Challenge: Billed Electricity Use**



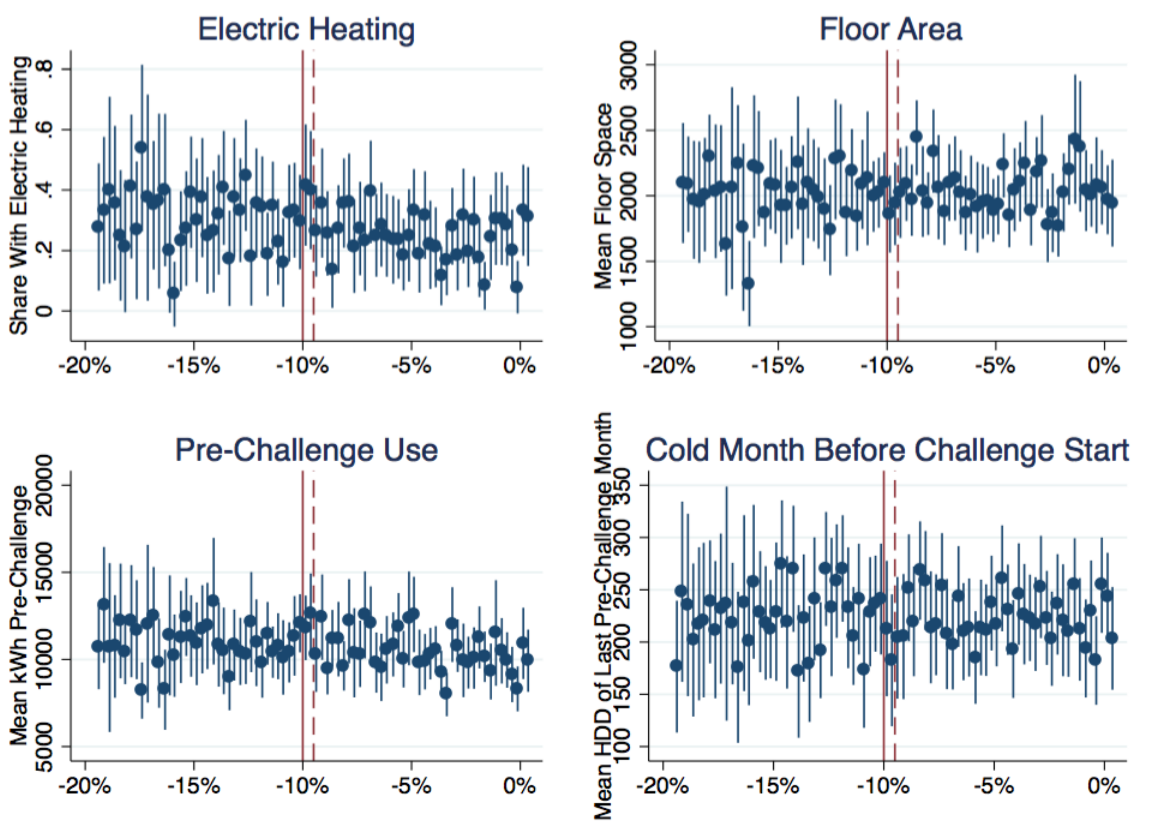
Notes: Billed changes are the percent change in billed electricity consumption from the pre-program year to the year of the first conservation challenge. The -9.5% level is shown by the vertical dashed line - note this is not the threshold for success as success was defined from credited - not billed - changes. Point estimates in the top bottom panel are the average probability of continuing to a second conservation challenge within 0.75%-wide bins of billed changes from the first conservation challenge. The dashed line in the top panel shows separate 1st order polynomial fits to households with billed changes above and below the indicated -9.5% threshold.
The bottom panel shows the corresponding fraction who pass their initial reduction challenge (dark connected line) and subsequent challenge (light grey scatter plot.) The dashed grey line in the bottom panel is a 3rd order kernel-weighted local polynomial fit to the fraction of households that pass their second conservation challenge.

Figure B.2: **Density Test of the Running Variable - 9.5% Threshold**



Notes: McCrary (2008) density test of the percent change in electricity use from a household's initial conservation challenge. The dark line is a smoothed local linear fit to the density of changes in electricity use, with 95% confidence intervals indicated by the light grey line. Point estimates of the density are grey circles. The dashed red line is the 9.5% reduction threshold, and the sold line is the 10% reduction target.

Figure B.3: **Continuity of Covariates at The Discontinuity**



Notes: Averages of four example covariates in the vicinity of the discontinuity by 0.25% wide binds of credited changes. *Electric Heating* is the share of households with electric space heating. *Floor Area* is the average floor space of a household. *Pre-Challenge Use* is the average electricity use in the year before a household begins its first challenge. *Cold Month Before Challenge Start* is the average heating degree days in the last month prior to the initial challenge. This is a measure of the last weather shock prior to the initial participation decision. The x-axis shows reductions in credited use with the dashed red vertical line denoting the 9.5% threshold for success and the solid red vertical line denoting the 10% conservation target.

Table B.1: **Discontinuity Tests of Covariates**

Dependent Variable:	Window Size				
	3	4	5	6	7
Heating	0.089	0.036	0.022	0.004	0.037
	(0.076)	(0.065)	(0.058)	(0.053)	(0.049)
Floor Area	-168.672	-118.697	-110.062	-61.026	-27.826
	(129.915)	(111.300)	(99.744)	(91.517)	(85.809)
Pre-Program kWh	1155.694	1126.179	673.470	71.468	198.537
	(853.864)	(720.725)	(663.070)	(589.560)	(552.341)
Pre-Program HDD	-23.74	-18.97	-6.499	-4.555	-2.678
	(23.32)	(19.74)	(17.53)	(15.82)	(14.67)
Property Value	-48.510	-30.153	5.379	34.891	54.255
	(68.793)	(62.081)	(54.773)	(51.611)	(47.112)
Share SFD	-0.065	-0.022	-0.042	-0.052	-0.012
	(0.078)	(0.067)	(0.060)	(0.054)	(0.050)

Notes: The table shows regression discontinuity estimates of γ_1 estimated using equation 2.1 for the listed dependent variables. The lack of statistically significant differences in covariates at the discontinuity supports that treatment is as good as randomly assigned at the discontinuity. Estimates included separate linear trends billed reductions and are not shown for conciseness. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 are listed for completeness but no coefficients are significant at a 10% level.

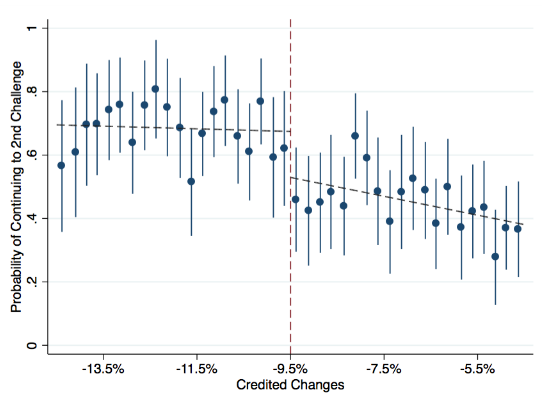
Table B.2: **2nd Order Bias-Corrected Fuzzy Regression Discontinuity Estimates**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Conventional	-0.281	-0.189	-0.211*	-0.264*	-0.319*	-0.331*	-0.281*	-0.181**
	(0.199)	(0.125)	(0.123)	(0.156)	(0.192)	(0.193)	(0.153)	(0.0794)
Bias-corrected	-0.421**	-0.315**	-0.207*	-0.149	-0.133	-0.216	-0.323**	-0.196**
	(0.199)	(0.125)	(0.123)	(0.156)	(0.192)	(0.193)	(0.153)	(0.0794)
Robust	-0.421	-0.315*	-0.207	-0.149	-0.133	-0.216	-0.323	-0.196**
	(0.264)	(0.168)	(0.166)	(0.210)	(0.259)	(0.257)	(0.201)	(0.0848)
Observations	888	1196	1475	1763	2050	2296	2543	4160
OrderPoly.(p)	2	2	2	2	2	2	2	2
OrderBias(q)	3	3	3	3	3	3	3	3
BWPoly.(h)	3%	4%	5%	6%	7%	8%	9%	18%
BWBias(b)	3%	4%	5%	6%	7%	8%	9%	33%
F-Conv.	3.1	4.9	5.5	4.5	3.9	4.1	5.3	14.8
F-Bias	1.4	3.6	6	9.2	9.1	7.3	5.8	14.8
F-Robust	.8	2.1	3.4	5.1	5	4	3.2	13

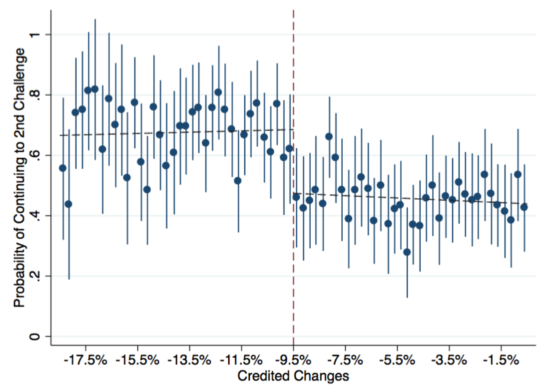
Notes: All specifications are a 2nd order polynomial estimated with a triangular kernel and restricted to households that either start their next challenge within 12 months or do not undertake an additional challenge. Specifications (1) through (7) are for $\pm$ the listed bandwidths around the threshold. Specification (8) determines the optimal polynomial and bias-correction bandwidths to be 18% and 33%, respectively. Conventional, bias-corrected, and bias-corrected and robust F-stats on the 1st stage instrument respectively denoted by F-Conv., F-Bias, and F-Robust. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure B.4: **First Stage and Reduced Form Discontinuities**

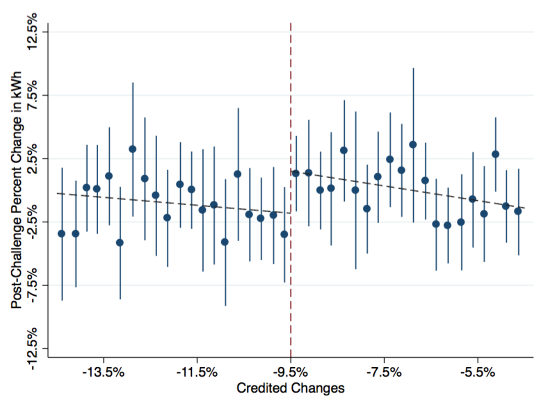
(a) First Stage - ±5% Bandwidth



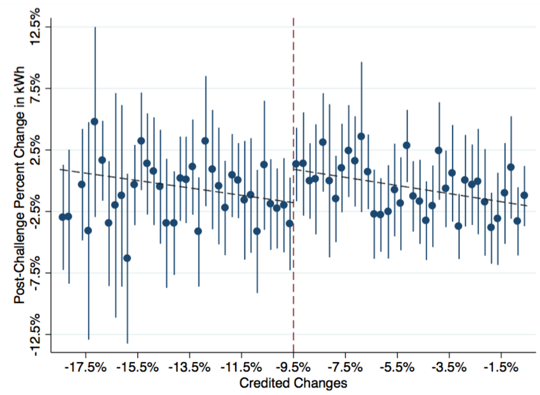
(b) First Stage - ±9% Bandwidth



(c) Reduced Form - ±5% Bandwidth



(d) Reduced Form - ±9% Bandwidth



Notes: This figure plots the first stage and reduced form discontinuities for bandwidths of ±5% and ±9% around the 9.5% reduction threshold in credited changes. Individual point estimates are the average of the outcome variable within 0.25% width bins in credited changes.

Table B.3: **Fuzzy Regression Discontinuity Estimates: 6 Month Gap**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A – First Stage							
Dependent variable: Continue to a Second Challenge C_i							
Window			±7 %	±6 %	±5 %	±4 %	±3 %
γ_1: Success Indicator			0.203***	0.195***	0.161***	0.163***	0.170**
			(0.0468)	(0.0505)	(0.0556)	(0.0622)	(0.0727)
γ_2: Cred. Reduc.			-0.775	-1.480	-2.573**	-0.510	1.794
			(0.822)	(1.016)	(1.282)	(1.818)	(2.876)
γ_3 : Success× Cred. Reduc.			1.854	3.156**	2.909	-0.527	-3.965
			(1.204)	(1.491)	(1.959)	(2.689)	(4.183)
γ_4 : Billed Reduc.			-0.576*	-0.678*	-0.367	-0.649	-0.867*
			(0.344)	(0.367)	(0.391)	(0.433)	(0.520)
γ_0 : Constant			0.435***	0.453***	0.466***	0.440***	0.419***
			(0.0317)	(0.0344)	(0.0377)	(0.0424)	(0.0495)
F-stat			18.79	14.85	8.374	6.872	5.452
Panel B – Second Stage							
Dependent variable: Percent change in post-challenge electricity use							
	OLS		Instrumental Variable Estimates				
Window		±5 %	±7 %	±6 %	±5 %	±4 %	±3 %
β_1 : Re-Enroll	-0.0171***	-0.0246***	-0.143**	-0.165**	-0.190*	-0.241*	-0.162
	(0.00444)	(0.00718)	(0.0651)	(0.0739)	(0.102)	(0.124)	(0.114)
β_2: Cred. Reduc.		-0.0829	-0.482*	-0.580	-0.864	-0.607	0.940
		(0.255)	(0.272)	(0.363)	(0.570)	(0.642)	(0.675)
β_3 : Success× Cred. Reduc.		0.451	0.423	0.544	0.827	-0.126	-1.758
		(0.456)	(0.338)	(0.446)	(0.596)	(0.905)	(1.246)
β_4 : Billed Reduc.		-0.136	-0.120	-0.197*	-0.198	-0.325**	-0.333*
		(0.105)	(0.101)	(0.112)	(0.122)	(0.162)	(0.173)
β_0 : Constant	-0.00773***	0.0145*	0.0747**	0.0879**	0.104*	0.124*	0.0731
	(0.00289)	(0.00793)	(0.0348)	(0.0407)	(0.0554)	(0.0646)	(0.0571)
N	4810	1287	1779	1535	1287	1039	775

Notes: This table reports fuzzy-RD estimates corresponding to equations (2.1) and (2.2). Estimation sample restricted to households that either start their next challenge within 6 months or do not undertake an additional challenge. Estimation window is restricted to ± the listed percent around the 9.5% threshold in credited changes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Robustness Checks - Log monthly electricity use

An alternative to defining the outcome in the fuzzy-RD approach as the post-program changes in electricity use, equation (2.3), is to use log monthly electricity use. This has the benefit of not requiring aggregating to annual changes at the cost of a less transparent estimation. Using log monthly electricity use finds similar estimates as using post-program changes from (2.3).

The first stage relationship is

$$C_i = \alpha_i + \gamma_0 D_{it,1} + \gamma_1 1\{R_i \leq \bar{R}\} \times D_{it,1} + \gamma_2 R_i \times D_{it,1} + \gamma_3 1\{R_i \leq \bar{R}\} \times R_i \times D_{it,1} + \gamma_4 B_i \times D_{it,1} + \gamma_5 X_i + \eta_{it} \tag{B.1}$$

where C_i is a binary indicator for whether a household continues to a second challenge, R_i are households' credited changes in electricity use from the first challenge, R_d is the threshold for success in the challenge and is -9.5%, $1\{R_i \leq \bar{R}\}$ is the dummy variable for success in the initial challenge, B_i are the billed changes from the initial challenge, and X_i is a vector of other controls. The instrument excluded from the second stage is $1\{R_i \leq \bar{R}\}$. $D_{it,1}$ is an indicator for if household i in month t was participating in the second challenge. The estimation sample is restricted to only observations for households undertaking their first conservation challenge or in their post-program year of a second challenge or after exiting the program.

The second-stage relationship is

$$y_{it} = \lambda_i + \beta_0 D_{it,1} + \beta_1 C_i + \beta_2 R_i \times D_{it,1} + \beta_3 1\{R_i \leq \bar{R}\} \times R_i \times D_{it,1} + \beta_4 B_i \times D_{it,1} + \beta_5 X_i + \epsilon_i \tag{B.2}$$

where y_{it} is log monthly electricity use.

Table B.4: **Fuzzy Regression Discontinuity Estimates: Log Monthly Electricity Use and 12 Month Gap**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A – First Stage				
Dependent variable: Continue to a Second Challenge C_i							
Window			±7 %	±6 %	±5 %	±4 %	±3 %
γ_1: Success$_i$			0.197***	0.185***	0.139***	0.138**	0.177***
			(0.0431)	(0.0467)	(0.0515)	(0.0574)	(0.0679)
γ_0: $D_{it,1}$			0.473***	0.477***	0.504***	0.519***	0.503***
			(0.0251)	(0.0270)	(0.0297)	(0.0324)	(0.0380)
γ_2: $R_i \times D_{it,1}$			-0.431	-0.721	-2.133**	-1.961	0.285
			(0.629)	(0.763)	(0.968)	(1.304)	(2.052)
$\gamma_3 : 1\{R_i \geq \bar{R}\}$			2.268***	2.757***	1.936*	2.066	0.0676
$\times R_i \times D_{it,1}$			(0.687)	(0.875)	(1.133)	(1.477)	(2.315)
$\gamma_4 : B_i \times D_{it,1}$			-0.288	-0.357	-0.0474	-0.288	-0.527
			(0.323)	(0.345)	(0.367)	(0.402)	(0.488)
F-stat			21.00	15.66	7.302	5.772	6.818
			Panel B – Second Stage				
Dependent variable: Log monthly electricity use							
	OLS		Instrumental Variable Estimates				
Window		±5 %	±7 %	±6 %	±5 %	±4 %	±3 %
β_1: Success$_i$	-0.0216***	-0.0279***	-0.144**	-0.184**	-0.229*	-0.299*	-0.146
	(0.00435)	(0.00732)	(0.0679)	(0.0816)	(0.127)	(0.162)	(0.109)
β_0: $D_{it,1}$	-0.0156***	-0.00267	0.0647*	0.0880*	0.115	0.161*	0.0747
	(0.00302)	(0.00539)	(0.0388)	(0.0465)	(0.0731)	(0.0953)	(0.0648)
β_2: $R_i \times D_{it,1}$			-0.289	-0.437	-0.773	-1.233	0.00274
			(0.227)	(0.295)	(0.574)	(0.808)	(0.565)
$\beta_3 : 1\{R_i \geq \bar{R}\}$			0.689**	0.863**	1.414***	2.216***	2.418***
$\times R_i \times D_{it,1}$			(0.278)	(0.349)	(0.435)	(0.591)	(0.814)
$\beta_4 : B_i \times D_{it,1}$			-0.136	-0.214*	-0.214	-0.328*	-0.367**
			(0.108)	(0.119)	(0.135)	(0.171)	(0.173)
N	130368	35400	49200	42312	35400	28704	21312
Households			2050	1763	1475	1196	888

Notes: Estimates using log monthly electricity use from Appendix section B.3. *** p<0.01, ** p<0.05, * p<0.1.

Table B.5: **Fuzzy Regression Discontinuity Estimates: Log Monthly Electricity Use and 6 Month Gap**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A − First Stage				
Dependent variable: Continue to a Second Challenge C_i							
Window			±7 %	±6 %	±5 %	±4 %	±3 %
γ_1: Success$_i$			0.187***	0.176***	0.154***	0.134***	0.145***
			(0.0391)	(0.0411)	(0.0426)	(0.0470)	(0.0523)
γ_0: $D_{it,1}$			0.420***	0.431***	0.440***	0.445***	0.439***
			(0.0234)	(0.0245)	(0.0258)	(0.0285)	(0.0318)
γ_2: $R_i \times D_{it,1}$			-0.894**	-1.047**	-1.264***	-1.817***	-1.812***
			(0.404)	(0.438)	(0.475)	(0.577)	(0.668)
$\gamma_3 : 1\{R_i \geq R_d\}$			2.526***	2.557***	2.427***	2.961***	1.370
$\times R_i \times D_{it,1}$			(0.604)	(0.736)	(0.856)	(1.008)	(1.249)
$\gamma_4 : B_i \times D_{it,1}$			0.185	0.327	0.284	0.638	1.241
			(0.444)	(0.514)	(0.611)	(0.771)	(1.062)
F-stat			22.82	18.25	13.01	8.174	7.657
			Panel B − Second Stage				
Dependent variable: Log monthly electricity use							
	OLS		Instrumental Variable Estimates				
Window		±5 %	±7 %	±6 %	±5 %	±4 %	±3 %
β_1: Success$_i$	-0.0223***	-0.0281***	-0.151**	-0.178**	-0.219**	-0.286**	-0.208*
	(0.00462)	(0.00768)	(0.0618)	(0.0723)	(0.0920)	(0.130)	(0.115)
β_0: $D_{it,1}$	-0.0156***	-0.00267	0.0569*	0.0700*	0.0884*	0.124*	0.0836
	(0.00302)	(0.00539)	(0.0316)	(0.0373)	(0.0474)	(0.0663)	(0.0584)
β_2: $R_i \times D_{it,1}$			-0.420**	-0.558***	-0.720***	-0.998**	-0.906**
			(0.164)	(0.202)	(0.262)	(0.425)	(0.406)
$\beta_3 : 1\{R_i \geq R_d\}$			0.394*	0.545**	0.748**	1.038**	1.233***
$\times R_i \times D_{it,1}$			(0.228)	(0.259)	(0.309)	(0.459)	(0.328)
$\beta_4 : B_i \times D_{it,1}$			0.156	0.240	0.389*	0.291	0.645*
			(0.136)	(0.166)	(0.213)	(0.298)	(0.374)
N	115440	30888	44976	38496	32352	26112	19536
Households			1874	1604	1348	1088	814

Notes: Estimates using log monthly electricity use from Appendix section B.3. *** p<0.01, ** p<0.05, * p<0.1.

Appendix C

Appendix to Chapter 3

C.1 Derivations

C.1.1 Mode-specific expenditure from nested CES preferences

The representative consumer has CES preferences over quantity Q_w of product variety w delivered by either mode,

$$U = \left[\sum_w (Q_w)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} \tag{C.1}$$

and preferences over aggregate quantity Q_w

$$Q_w = \left[\sum_{m=1}^{2} \lambda_w^m (q_w^m)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}} \tag{C.2}$$

where q_w^m is the quantity delivered by mode m. Elasticity of substitution across varieties and across modes are $\eta > 1$ and $\rho > 1$ and the consumer maximizes U subject to their budget constraint of total expenditure I,

$$\sum_m \sum_w p_w^m q_w^m \leq I$$

The solution to this optimization problem requires proving that expenditure is independent across varieties w and has been proved in Strotz (1957) and is equivalent to Sillard and Wilner (2015). To see the equivalence define $Q_m \equiv [\sum_i \alpha_{im} x_{im}^{\rho_m}]^{\frac{1}{\rho_m}}$ so that the maximization of Sillard and Wilner (2015) is $U(x) = \{\sum_m Q_m^{\rho}\}^{\frac{1}{\rho}}$. Transform notation from Sillard and Wilner (2015) to this chapter by

$$
\begin{aligned}
m &\rightarrow w \\
i &\rightarrow m \\
\alpha_{im} &\rightarrow \lambda_w^m \\
\rho &\rightarrow \frac{\eta-1}{\eta} \\
\rho_m &\rightarrow \frac{\rho-1}{\rho} \\
x_{im} &\rightarrow q_w^m \\
\sigma_m &\rightarrow \rho \\
\sigma &\rightarrow \eta
\end{aligned}
$$

which delivers the same optimization problem as this chapter,

$$U(x) = \left\{ \sum_w (Q_w)^{\frac{\eta-1}{\eta}} \right\}^{\frac{\eta}{\eta-1}}$$

and

$$Q_w \equiv \left[\sum_m \lambda_w^m \left(q_w^m\right)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}$$

Expenditure on variety w is given by

$$X_w = \left(\sum_m \left(\lambda_w^m\right)^\rho \left(p_w^m\right)^{1-\rho} \right)^{\frac{1}{1-\rho}}$$

and final demand becomes

$$
\begin{aligned}
q_w^m &= R \frac{\left(\frac{\lambda_w^m}{p_{mw}}\right)^\rho X_w^{\rho-\eta}}{\sum_w X_w^{1-\eta}} \\
&= R \frac{\left(\frac{\lambda_w^m}{p_{mw}}\right)^\rho \left(\sum_m \left(\lambda_w^m\right)^\rho \left(p_w^m\right)^{1-\rho}\right)^{\frac{\rho-\eta}{1-\rho}}}{\sum_w \left(\sum_m \left(\lambda_w^m\right)^\rho \left(p_w^m\right)^{1-\rho}\right)^{\frac{1-\eta}{1-\rho}}} \\
&= R \frac{\left(\frac{\lambda_w^m}{p_{mw}}\right)^\rho}{\left(P_w\right)^{1-\rho}} \left(\frac{P_w}{P}\right)^{1-\eta} \\
&= \frac{X_n}{p_w^m} \frac{(\lambda_w^m)^\rho (p_w^m)^{1-\rho}}{P_w^{1-\rho}} \left(\frac{P_w}{P_n}\right)^{1-\eta} \qquad \text{(C.3)}
\end{aligned}
$$

The Armington assumption of a continuum of differentiated varieties produces the final country-mode specific expenditure.

C.1.2 Elasticities of Trade and Mode Shares

Trade and mode shares follow with algebra from their respective definitions. $\pi_{ni} \equiv \frac{X_{ni}}{X_n}$ follows from

$$\begin{aligned} X_{ni} &= \sum_m X_{ni}^m \\ &= \sum_m X_n \frac{(\lambda_{ni}^m)^\rho (p_{ni}^m)^{1-\rho}}{P_{ni}^{1-\rho}} \left(\frac{P_{ni}}{P_n}\right)^{1-\eta} \\ &= X_n \left(\frac{P_{ni}}{P_n}\right)^{1-\eta} \frac{1}{P_{ni}^{1-\rho}} \sum_m (\lambda_{ni}^m)^\rho (p_{ni}^m)^{1-\rho} \\ &= X_n \left(\frac{P_{ni}}{P_n}\right)^{1-\eta} \end{aligned} \tag{C.4}$$

and mode shares follow from the definition and equation (3.3),

$$\gamma_{ni}^m \equiv \frac{X_{ni}^m}{X_{ni}} = \frac{\lambda_{ni}^{m\rho}(p_{ni}^m)^{1-\rho}}{P_{ni}^{1-\rho}}$$

Elasticities

Assume a shock to the transport costs for imports from country j to n by mode m', $\tau_{nj}^{m'}$. Changes CIF expenditures can be decomposed from $X_{ni}^m = X_n \pi_{ni} \gamma_{ni}^m$,

$$\frac{\partial ln X_{ni}^m}{\partial ln \tau_{nj}^{m'}} = \frac{\partial ln X_n}{\partial ln \tau_{nj}^{m'}} + \underbrace{\frac{\partial ln \pi_{ni}}{\partial ln \tau_{nj}^{m'}}}_{(A)} + \underbrace{\frac{\partial ln \gamma_{ni}^m}{\partial ln \tau_{nj}^{m'}}}_{(B)}$$

Term (A)

Term (A) starts from

$$ln \pi_{ni} = (1-\eta)(ln P_{ni}) - ln P_n^{1-\eta}$$

Taking derivatives of each term w.r.t. $\tau_{nj}^{m'}$, and starting with the first term:

$$
\begin{aligned}
(1-\eta)\frac{\partial lnP_{ni}}{\partial \tau_{nj}^{m'}} &= 0 \; if \; i \neq j, \; \forall \; m, m' \\
&= \frac{(1-\eta)}{(1-\rho)}\frac{\partial lnP_{ni}^{1-\rho}}{\partial \tau_{nj}^{m'}} \\
&= \frac{(1-\eta)}{(1-\rho)}\frac{1}{P_{ni}^{1-\rho}}\frac{\partial}{\partial \tau_{nj}^{m'}}\sum_m (\lambda_{ni}^m)^\rho (p_{ni}^m)^{1-\rho} \\
&= \frac{(1-\eta)}{P_{ni}^{1-\rho}}(\lambda_{nj}^{m'})^\rho (p_{nj}^{m'})^{-\rho} c_j \\
&= \frac{(1-\eta)}{P_{ni}^{1-\rho}\tau_{nj}^{m'}}(\lambda_{nj}^{m'})^\rho (p_{nj}^{m'})^{1-\rho} \\
&= (1-\eta)\frac{\gamma_{ni}^{m'}}{\tau_{ni}^{m'}} \; if \; i = j, \; \forall \; m, m'
\end{aligned}
$$

Solving for the second term,

$$
\frac{\partial lnP_n^{1-\eta}}{\partial \tau_{nj}^{m'}} = \frac{1}{P_n^{1-\eta}}\frac{\partial P_n^{1-\eta}}{\partial \tau_{nj}^{m'}} = \frac{1}{P_n^{1-\eta}}\frac{\partial}{\partial \tau_{nj}^{m'}}\sum_i {P_{ni}}^{1-\eta} = \frac{1}{P_n^{1-\eta}}\frac{\partial P_{nj}^{1-\eta}}{\partial \tau_{nj}^{m'}} = \frac{P_{nj}^{1-\eta}}{P_n^{1-\eta}}\frac{\partial lnP_{nj}^{1-\eta}}{\partial \tau_{nj}^{m'}} = \pi_{nj}\frac{\partial lnP_{nj}^{1-\eta}}{\partial \tau_{nj}^{m'}}
$$

Combined with the result from the first term we have

$$
\frac{\partial lnP_n^{1-\eta}}{\partial \tau_{nj}^{m'}} = (1-\eta)\pi_{nj}\frac{\gamma_{nj}^{m'}}{\tau_{nj}^{m'}}
$$

Therefore,

$$
(A) = \frac{\partial ln\pi_{ni}}{\partial ln\tau_{nj}^{m'}} = \begin{cases} (1-\eta)(1-\pi_{ni})\gamma_{ni}^{m'} & if \; i \; = \; j \; \forall \; m, m' \\ (1-\eta)(-\pi_{nj})\gamma_{nj}^{m'} & if \; i \; \neq \; j \; \forall \; m, m' \end{cases}
$$

Term (B)

Term (B) begins from $\gamma_{ni}^m = \frac{\lambda_{ni}^{m\,\rho}(p_{ni}^m)^{1-\rho}}{P_{ni}^{1-\rho}}$ and the prior derivations for changes in the price indexes,

$$
\begin{aligned}
\frac{\partial ln\gamma_{ni}^m}{\partial ln\tau_{nj}^{m'}} &= \frac{\partial ln(\lambda_{ni}^{m\,\rho}(p_{ni}^m)^{1-\rho})}{\partial ln\tau_{nj}^{m'}} - \frac{\partial lnP_{ni}^{1-\rho}}{\partial ln\tau_{nj}^{m'}} = (1-\rho)(1-\gamma_{ni}^{m'}) \; if \; i = j \; and \; m = m' \\
&= (1-\rho) - (1-\rho)\gamma_{ni}^{m'} if \; i = j \; and \; m = m' \\
(B) = \frac{\partial ln\gamma_{ni}^m}{\partial ln\tau_{nj}^{m'}} &= \begin{cases} (1-\rho)(1-\gamma_{ni}^{m'}) & if \; i = j \; and \; m = m' \\ (1-\rho)(-\gamma_{nj}^{m'}) & if \; i = j \; and \; m \neq m' \\ 0 & if \; i \neq j \; \forall \; m, m' \end{cases}
\end{aligned}
$$

Expenditure changes

Combining (A) and (B) and assuming aggregate expenditure X_n is fixed gives the change in expenditure from a change in iceberg trade costs,

$$\frac{\partial lnX^m_{ni}}{\partial ln\tau^{m'}_{nj}} = \begin{cases} (1-\eta)(1-\pi_{ni})\gamma^m_{ni} + (1-\rho)(1-\gamma^m_{ni}) & if\ i=j\ and\ m=m' \\ (1-\eta)(1-\pi_{ni})\gamma^{m'}_{ni} + (1-\rho)(-\gamma^{m'}_{ni}) & if\ i=j\ and\ m\neq m' \\ (1-\eta)(-\pi_{nj})\gamma^{m'}_{nj} & if\ i\neq j\ \forall\ m,m' \end{cases}$$

C.1.3 Difference between CIF and FOB values

Note that X^m_{ni} is defined in CIF terms, $X^m_{ni} = p^m_{ni}\cdot q^m_{ni}$, where p^m_{ni} is the CIF price $p^m_{ni} = c_i \cdot \tau^m_{ni}$. As a result, an increase in τ^m_{ni} will both change the quantity imported, q^m_{ni}, and the value of the imports.

$$\begin{aligned} \frac{\partial lnX^m_{ni}}{\partial ln\tau^{m'}_{nj}} &= \frac{\partial lnp^m_{ni}}{\partial ln\tau^{m'}_{nj}} + \frac{\partial lnq^m_{ni}}{\partial ln\tau^{m'}_{nj}} \\ &= \frac{\partial lnc_i}{\partial ln\tau^{m'}_{nj}} + \frac{\partial ln\tau^m_{ni}}{\partial ln\tau^{m'}_{nj}} + \frac{\partial lnq^m_{ni}}{\partial ln\tau^{m'}_{nj}} \end{aligned}$$

Assume Free-on-Board origin prices are independent of changes in trade costs, $\frac{\partial lnc_i}{\partial ln\tau^{m'}_{nj}} = 0$,

$$\frac{\partial lnX^m_{ni}}{\partial ln\tau^{m'}_{nj}} = \frac{\partial ln\tau^m_{ni}}{\partial ln\tau^{m'}_{nj}} + \frac{\partial lnq^m_{ni}}{\partial ln\tau^{m'}_{nj}} = \begin{cases} 1 + \frac{\partial lnq^m_{ni}}{\partial ln\tau^{m'}_{nj}} & if\ i=j\ and\ m=m' \\ \frac{\partial lnq^m_{ni}}{\partial ln\tau^{m'}_{nj}} & if\ i\neq j\ or\ m\neq m' \end{cases}$$

Consider FOB values,

$$X^{m,FOB}_{ni} = c_i * q^m_{ni} = \frac{p^m_{ni}q^m_{ni}}{\tau^m_{ni}} = \frac{X^{m,CIF}_{ni}}{\tau^m_{ni}}$$

Therefore,

$$\frac{\partial lnX^{m,FOB}_{ni}}{\partial ln\tau^{m'}_{nj}} = \frac{\partial lnX^{m,CIF}_{ni}}{\partial ln\tau^{m'}_{nj}} - \frac{\partial ln\tau^m_{ni}}{\partial ln\tau^{m'}_{nj}} = \begin{cases} \frac{\partial lnX^{m,CIF}_{ni}}{\partial ln\tau^{m'}_{nj}} - 1 & if\ i=j\ and\ m=m' \\ \frac{\partial lnX^{m,CIF}_{ni}}{\partial ln\tau^{m'}_{nj}} & if\ i\neq j\ or\ m\neq m' \end{cases}$$

Importantly, this means that the cross elasticity - which determines substitutes or complements - is the same whether imports are valued in CIF and FOB values. This is intuitive, since a change in freight costs will only affect the value of imports that include those freight costs; products redistributing to other modes and countries will face those countries freight costs and modes, which are unaffected, and therefore the percentage change in value will be the same whether measured in CIF or FOB values.

C.2 Comparison to Existing Models

C.2.1 Lux (2011)

The equivalence follows from equations (5)-(7) of Lux (2011). Let θ' and ρ' denote parameters from Lux (2011) and define the transformations $-\frac{\theta'}{1-\rho'} \equiv (1-\rho)$, $-\theta' \equiv (1-\eta)$, and $p^m_{ni} \equiv T_i^{\frac{1}{1-\eta'}} c_i \tau^m_{ni}$. Note the typo in Lux (2011) equation (6) which is missing a negative sign in the exponent $\frac{1-\rho'}{\theta'}$. This produces the same reduced form equations for mode and trade shares as the nested CES model.

C.2.2 Hummels and Schaur (2013)

Equation (1) in Hummels and Schaur (2013) is equivalent to the nested CES model equation C.3 under the assumptions $\rho = \eta$ in the CES model, which then collapses to a single CES preferences over product varieties and perceived quality from fast delivery, and using the same parameterization over fast delivery used by Hummels and Schaur (2013). That is, variety w delivered by mode m is $\lambda^m_w = \upsilon^m_w exp(-\tau \cdot days^m_w)$. See Hummels and Schaur (2013) for the definition of these terms. Similarly, equation (3.14) for the elasticity of substitution over modes is equivalent to equation (6) of Hummels and Schaur (2013).

$$ln\left(\frac{X^a_{ni,t}}{X^o_{ni,t}}\right) = \rho\tau(days^o_w - 1) + (1-\rho)\frac{c^a_{i,t}}{c^o_{i,t}} - \rho\frac{\tau^a_{ni,t}}{\tau^o_{ni,t}} + \rho ln\frac{\upsilon^a_{ni}}{\upsilon^o_{ni}} + \epsilon_{ni,t}$$

C.2.3 Shapiro (2016)

Shapiro (2016) combines mode-specific carbon taxes in a weighted average to produce an equivalent tariff on aggregate trade. After counterfactual trade patterns are determined from the model, the trade is projected back into mode shares based on the initial shares (his standard approach) or by assuming a unit mode elasticity in his endogenous mode shares robustness check. Consider a change in trade costs like a carbon price applied to both air and ocean transport within a trade route, and assume the standard form for trade and mode shares. Changes in imports by air are,

$$\begin{aligned}
dlnX^a_{ni} &= \frac{\partial lnX^a_{ni}}{\partial ln\tau^a_{ni}} dln\tau^a_{ni} + \frac{\partial lnX^a_{ni}}{\partial ln\tau^o_{ni}} dln\tau^o_{ni} \\
&= [(1-\eta)(1-\pi_{ni})\gamma^a_{ni} + (1-\rho)(1-\gamma^a_{ni})]dln\tau^a_{ni} + [(1-\eta)(1-\pi_{ni})\gamma^o_{ni} + (1-\rho)(-\gamma^o_{ni})]dln\tau^o_{ni} \\
&= (1-\eta)(1-\pi_{ni})(\gamma^a_{ni} dln\tau^a_{ni} + \gamma^o_{ni} dln\tau^o_{ni}) + (1-\rho)\gamma^o_{ni}(dln\tau^a_{ni} - dln\tau^o_{ni}) \\
&= (1-\eta)(1-\pi_{ni})dln\tau^w_{ni} + (1-\rho)\gamma^o_{ni}(dln\tau^a_{ni} - dln\tau^o_{ni})
\end{aligned}$$

where τ^w_{ni} is the mode-share weighted average change in trade cost. If we assume mode shares are fixed, then $\rho = 1$ and this is a similar change in expenditure shares from a weighted average carbon tax ($dln\tau^w_{ni}$) to that of Shapiro (2016). Alternatively, using a unit elasticity of mode shares to freight rates (equation 3.9) as assumed by Shapiro (2016) corresponds approximately to $\rho = 2$. The comparison isn't exact as Shapiro (2016) doesn't have a model of endogenous mode choice and assumes isoelastic mode shares instead of including the term γ_{ni}^{m} in the elasticity of mode shares.

C.3 Data

The U.S. Imports of Merchandise monthly data is available from the United States Census Bureau. Within each individual monthly file I drop all observations that have missing data for shipping value, weight, or freight charges. Eight of the monthly files were corrupted and are missing from the dataset. I interpolate a missing month's data using the average of the previous and subsequent months within the most detailed observation available; HS10, country, import district, and unloading district. I construct the measure of import bulk using the variable CARDS_MO, which is defined as the "Number of Detailed Records, Current Month." This variable is available within an HS10-month-origin country-import district-unlading port observation and indicates the number of distinct shipments recorded by Customs Census (1996). However, the variable aggregates over both air and ocean imports. To construct the number of separate air and ocean shipments, I first assign all observations with shipments via a single mode the corresponding CARDS_MO value. For those observations with imports observed over both modes, I use those observations with CARDS_MO of 1 or 2 only. This allows me to uniquely attribute the observed trade flow to shipments. Observations with greater than 2 shipments cannot distinguish which mode has the greater number of shipments and I cannot generate a value for import bulk for these observations. This attribution succeeds for 75% of import value. For estimating the own and cross-price elasticities I aggregate to the HS6 product level, and drop any observation with a per kilogram or ad-valorem freight rate below the 1st or above the 99th percentiles.

C.4 Estimated Elasticities

Tables C.1 and C.2 show estimated own and cross-price elasticities for air and ocean imports restricted to several subsets of observations, and which find results consistent with the main estimates.

Table C.1: **Estimated Elasticities: Air Imports**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:			Log Air Value			
Air Rate (Proxy)	-0.217***	-0.219***	-0.258***	-0.287***	-0.248***	-0.235***
	(0.0338)	(0.0200)	(0.0203)	(0.0536)	(0.0193)	(0.0177)
Sea Rate	-0.00226	0.00392	0.0128*	0.00853	-0.00746	0.0105*
	(0.0101)	(0.00502)	(0.00710)	(0.0119)	(0.00692)	(0.00584)
N	741338	712073	880487	358252	781859	862020
r2	0.399	0.409	0.370	0.402	0.362	0.388
Subset	E.U.	E.A.	East C.	West C.	< 2000	≥ 2000

Notes: This table reports estimates of equation (3.15) for air imports limited to subsets of observations. All specifications include HS6, HS6, Origin-Destination, Origin-Year, and Destination-Year. Air [Sea] Rate are log freight rates per kg. Air Rate (Proxy) is the average freight rate within the same trade route and industry excluding that observations freight rates. Subset E.U. [E.A.] are imports arriving from European Union [East Asian] countries. East C. [West C.] are imports arriving from all countries to the East Coast [West Coast] import districts of the United States. <2000 and >2000 are respectively restricted to the years 1992-1999 and 2000-2007. Standard errors are clustered at the origin-destination level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table C.2: **Estimated Elasticities: Ocean Imports**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:			Log Sea Value			
Air Rate	-0.106***	-0.159***	-0.160***	-0.152***	-0.170***	-0.139***
	(0.00817)	(0.00920)	(0.0102)	(0.00940)	(0.00838)	(0.00746)
Sea Rate (Proxy)	-0.320***	-0.191***	-0.325***	-0.312***	-0.299***	-0.305***
	(0.0394)	(0.0190)	(0.0308)	(0.0542)	(0.0255)	(0.0239)
N	756517	705370	891302	362227	787419	868452
r2	0.360	0.311	0.305	0.421	0.329	0.335
Subset	E.U.	E.A.	East C.	West C.	< 2000	> 2000

Notes: This table reports estimates of equation (3.15) for ocean imports limited to subsets of observations. All specifications include HS6, HS6, Origin-Destination, Origin-Year, and Destination-Year. Air [Sea] Rate are log freight rates per kg. Sea Rate (Proxy) is the average freight rate within the same trade route and industry excluding that observations freight rates. Subset E.U. [E.A.] are imports arriving from European Union [East Asian] countries. East C. [West C.] are imports arriving from all countries to the East Coast [West Coast] import districts of the United States. <2000 and >2000 are respectively restricted to the years 1992-1999 and 2000-2007. Standard errors are clustered at the origin-destination level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix D

Appendix to Chapter 1 (II): Table of All Estimates

D.1 Event Study Estimates For All Households

Table D.1: **Event-Study Point Estimates**

Pre	0.020***	0.030***
	(0.0051)	(0.010)
M-59	0.027***	0.034***
	(0.0055)	(0.0097)
M-58	0.032***	0.038***
	(0.0054)	(0.0093)
M-57	0.024***	0.030***
	(0.0055)	(0.0090)
M-56	0.024***	0.030***
	(0.0052)	(0.0087)
M-55	0.016***	0.021**
	(0.0054)	(0.0087)
M-54	0.019***	0.024***
	(0.0053)	(0.0086)
M-53	0.017***	0.021**
	(0.0053)	(0.0087)
M-52	0.020***	0.024***
	(0.0051)	(0.0086)
M-51	0.022***	0.026***
	(0.0049)	(0.0084)
M-50	0.024***	0.028***
	(0.0048)	(0.0082)
M-49	0.026***	0.030***
	(0.0048)	(0.0079)
M-48	0.029***	0.033***
	(0.0048)	(0.0076)
M-47	0.027***	0.030***

	(1)	(2)
	All Households	Participant Households
Dependent Variable: Ln monthly electricity use		
	(0.0047)	(0.0072)
M-46	0.026***	0.030***
	(0.0045)	(0.0069)
M-45	0.023***	0.026***
	(0.0045)	(0.0066)
M-44	0.020***	0.023***
	(0.0044)	(0.0064)
M-43	0.012***	0.015**
	(0.0045)	(0.0064)
M-42	0.013***	0.015**
	(0.0045)	(0.0064)
M-41	0.013***	0.015**
	(0.0045)	(0.0064)
M-40	0.016***	0.018***
	(0.0043)	(0.0063)
M-39	0.013***	0.015**
	(0.0043)	(0.0062)
M-38	0.018***	0.020***
	(0.0041)	(0.0059)
M-37	0.017***	0.019***
	(0.0041)	(0.0058)
M-36	0.017***	0.018***
	(0.0041)	(0.0055)
M-35	0.017***	0.018***
	(0.0041)	(0.0052)
M-34	0.017***	0.018***
	(0.0040)	(0.0049)
M-33	0.0091**	0.010**
	(0.0039)	(0.0047)
M-32	0.0054	0.0063
	(0.0038)	(0.0045)
M-31	0.0040	0.0048
	(0.0039)	(0.0045)
M-30	0.0064*	0.0069
	(0.0037)	(0.0043)
M-29	0.0071**	0.0075*

	(1)	(2)
	All Households	Participant Households
Dependent Variable: Ln monthly electricity use		
	(0.0035)	(0.0042)
M-28	0.0091***	0.0094**
	(0.0035)	(0.0041)
M-27	0.0077**	0.0080**
	(0.0033)	(0.0039)
M-26	0.0093***	0.0095**
	(0.0032)	(0.0037)
M-25	0.0061*	0.0061*
	(0.0031)	(0.0036)
M-24	0.0069**	0.0067**
	(0.0030)	(0.0033)
M-11	0.0017	0.0013
	(0.0029)	(0.0031)
M-10	0.0024	0.0020
	(0.0029)	(0.0033)
M-9	0.0011	0.00046
	(0.0030)	(0.0035)
M-8	-0.0047	-0.0055
	(0.0031)	(0.0037)
M-7	-0.0031	-0.0041
	(0.0032)	(0.0038)
M-6	-0.0021	-0.0033
	(0.0031)	(0.0038)
M-5	-0.0034	-0.0046
	(0.0031)	(0.0038)
M-4	-0.0047	-0.0060
	(0.0032)	(0.0039)
M-3	-0.011***	-0.012***
	(0.0032)	(0.0039)
M-2	-0.0080**	-0.0093**
	(0.0032)	(0.0040)
M-1	-0.017***	-0.018***
	(0.0034)	(0.0043)
M0	-0.022***	-0.023***
	(0.0033)	(0.0044)
M1	-0.041***	-0.042***

	(1)	(2)
	All Households	Participant Households
Dependent Variable: Ln monthly electricity use		
	(0.0035)	(0.0049)
M2	-0.049***	-0.051***
	(0.0036)	(0.0052)
M3	-0.051***	-0.053***
	(0.0036)	(0.0054)
M4	-0.052***	-0.054***
	(0.0035)	(0.0055)
M5	-0.050***	-0.052***
	(0.0036)	(0.0057)
M6	-0.049***	-0.051***
	(0.0037)	(0.0058)
M7	-0.048***	-0.050***
	(0.0036)	(0.0057)
M8	-0.050***	-0.052***
	(0.0036)	(0.0057)
M9	-0.054***	-0.056***
	(0.0038)	(0.0059)
M10	-0.049***	-0.051***
	(0.0038)	(0.0060)
M11	-0.050***	-0.052***
	(0.0037)	(0.0062)
M12	-0.049***	-0.051***
	(0.0037)	(0.0065)
M13	-0.053***	-0.055***
	(0.0038)	(0.0069)
M14	-0.055***	-0.057***
	(0.0038)	(0.0072)
M15	-0.056***	-0.058***
	(0.0039)	(0.0075)
M16	-0.057***	-0.059***
	(0.0039)	(0.0076)
M17	-0.058***	-0.060***
	(0.0039)	(0.0077)
M18	-0.059***	-0.061***
	(0.0039)	(0.0078)
M19	-0.060***	-0.063***

	(1)	(2)
	All Households	Participant Households
Dependent Variable: Ln monthly electricity use		
	(0.0039)	(0.0078)
M20	-0.059***	-0.061***
	(0.0040)	(0.0079)
M21	-0.061***	-0.063***
	(0.0041)	(0.0080)
M22	-0.052***	-0.055***
	(0.0041)	(0.0081)
M23	-0.053***	-0.056***
	(0.0041)	(0.0084)
M24	-0.050***	-0.053***
	(0.0042)	(0.0087)
M25	-0.051***	-0.054***
	(0.0043)	(0.0091)
M26	-0.054***	-0.057***
	(0.0043)	(0.0094)
M27	-0.053***	-0.056***
	(0.0043)	(0.0097)
M28	-0.054***	-0.057***
	(0.0044)	(0.0099)
M29	-0.054***	-0.057***
	(0.0045)	(0.010)
M30	-0.054***	-0.058***
	(0.0045)	(0.010)
M31	-0.053***	-0.056***
	(0.0045)	(0.010)
M32	-0.051***	-0.055***
	(0.0045)	(0.010)
M33	-0.053***	-0.056***
	(0.0046)	(0.010)
M34	-0.048***	-0.052***
	(0.0047)	(0.011)
M35	-0.045***	-0.049***
	(0.0047)	(0.011)
M36	-0.038***	-0.042***
	(0.0047)	(0.011)
M37	-0.042***	-0.046***

	(1)	(2)
	All Households	Participant Households
Dependent Variable: Ln monthly electricity use		
	(0.0048)	(0.011)
M38	-0.048***	-0.052***
	(0.0048)	(0.012)
M39	-0.046***	-0.049***
	(0.0048)	(0.012)
M40	-0.045***	-0.049***
	(0.0049)	(0.012)
M41	-0.046***	-0.049***
	(0.0049)	(0.012)
M42	-0.049***	-0.053***
	(0.0050)	(0.012)
M43	-0.048***	-0.052***
	(0.0050)	(0.012)
M44	-0.048***	-0.051***
	(0.0051)	(0.013)
M45	-0.048***	-0.051***
	(0.0052)	(0.013)
M46	-0.043***	-0.046***
	(0.0052)	(0.013)
M47	-0.041***	-0.044***
	(0.0052)	(0.013)
M48	-0.037***	-0.041***
	(0.0052)	(0.013)
M49	-0.039***	-0.042***
	(0.0053)	(0.014)
M50	-0.043***	-0.046***
	(0.0053)	(0.014)
M51	-0.046***	-0.048***
	(0.0055)	(0.014)
M52	-0.046***	-0.048***
	(0.0054)	(0.014)
M53	-0.042***	-0.044***
	(0.0055)	(0.015)
M54	-0.045***	-0.047***
	(0.0056)	(0.015)
M55	-0.045***	-0.048***

	(1)	(2)
	All Households	Participant Households
Dependent Variable: Ln monthly electricity use		
	(0.0058)	(0.015)
M56	-0.043***	-0.046***
	(0.0057)	(0.015)
M57	-0.052***	-0.055***
	(0.0058)	(0.015)
M58	-0.044***	-0.047***
	(0.0059)	(0.015)
M59	-0.040***	-0.043***
	(0.0059)	(0.015)
M60	-0.036***	-0.039**
	(0.0059)	(0.016)
M61	-0.038***	-0.040**
	(0.0061)	(0.016)
M62	-0.039***	-0.040**
	(0.0061)	(0.016)
M63	-0.042***	-0.044***
	(0.0062)	(0.017)
M64	-0.047***	-0.048***
	(0.0062)	(0.017)
M65	-0.044***	-0.045***
	(0.0063)	(0.017)
M66	-0.044***	-0.044**
	(0.0063)	(0.017)
M67	-0.049***	-0.050***
	(0.0064)	(0.017)
M68	-0.050***	-0.050***
	(0.0065)	(0.017)
M69	-0.053***	-0.054***
	(0.0066)	(0.017)
M70	-0.051***	-0.051***
	(0.0066)	(0.017)
M71	-0.043***	-0.043**
	(0.0067)	(0.018)
M72	-0.041***	-0.041**
	(0.0068)	(0.018)
Observations	2243267	1065236

All specifications include individual and date fixed effects. Specification (1) estimated including both participant and non-participant households. (2) estimated for participant households only. Standard errors are clustered at the household level. *** p<0.01, ** p<0.05, * p<0.1 denote significance levels where 0 is defined as the second year pre-treatment and consists of months M-12 to M-23.

D.2 Event Study Estimates By Number of Challenges

Table D.2: **Event-Study Estimates: Selection Into Challenges**

Pre	0.024***	0.020*	0.014	0.015	0.012	0.0088	0.015
	(0.0073)	(0.011)	(0.0090)	(0.018)	(0.0100)	(0.016)	(0.012)
M-59	0.035***	0.037***	0.021**	0.039**	0.012	0.0034	0.017
	(0.0076)	(0.011)	(0.0097)	(0.019)	(0.011)	(0.017)	(0.014)
M-58	0.037***	0.046***	0.019*	0.031	0.012	0.0091	0.013
	(0.0075)	(0.011)	(0.0098)	(0.020)	(0.011)	(0.017)	(0.014)
M-57	0.028***	0.031***	0.015	0.033*	0.0053	0.017	-0.0014
	(0.0076)	(0.011)	(0.0099)	(0.019)	(0.011)	(0.016)	(0.015)
M-56	0.029***	0.025**	0.019**	0.034*	0.011	0.020	0.0066
	(0.0071)	(0.011)	(0.0089)	(0.018)	(0.0100)	(0.015)	(0.013)
M-55	0.021***	0.012	0.015	0.034*	0.0049	0.0035	0.0057
	(0.0074)	(0.011)	(0.0094)	(0.019)	(0.011)	(0.016)	(0.014)
M-54	0.023***	0.016	0.017*	0.047***	0.0030	-0.0079	0.0087
	(0.0071)	(0.010)	(0.0093)	(0.018)	(0.011)	(0.016)	(0.014)
M-53	0.022***	0.016	0.015	0.040**	0.0033	-0.0089	0.0096
	(0.0073)	(0.011)	(0.0094)	(0.019)	(0.011)	(0.019)	(0.013)
M-52	0.029***	0.030***	0.017*	0.049***	0.0023	-0.0034	0.0051
	(0.0068)	(0.010)	(0.0089)	(0.018)	(0.0099)	(0.017)	(0.012)
M-51	0.031***	0.033***	0.019**	0.054***	0.0031	-0.0022	0.0055
	(0.0064)	(0.0097)	(0.0081)	(0.014)	(0.0096)	(0.016)	(0.012)
M-50	0.035***	0.039***	0.021***	0.057***	0.0046	-0.0036	0.0083
	(0.0062)	(0.0095)	(0.0080)	(0.014)	(0.0093)	(0.017)	(0.011)
M-49	0.034***	0.042***	0.018**	0.039***	0.0076	0.0030	0.0096
	(0.0062)	(0.0098)	(0.0078)	(0.015)	(0.0091)	(0.016)	(0.011)
M-48	0.035***	0.037***	0.023***	0.044***	0.012	0.015	0.010
	(0.0062)	(0.0094)	(0.0080)	(0.016)	(0.0091)	(0.015)	(0.011)
M-47	0.025***	0.021**	0.018**	0.030*	0.010	0.012	0.0092
	(0.0064)	(0.010)	(0.0080)	(0.016)	(0.0090)	(0.016)	(0.010)
M-46	0.024***	0.018*	0.018**	0.038**	0.0078	0.0046	0.0086
	(0.0060)	(0.0092)	(0.0078)	(0.015)	(0.0089)	(0.016)	(0.010)
M-45	0.025***	0.021**	0.018**	0.047***	0.0042	0.012	-0.000069
	(0.0060)	(0.0090)	(0.0077)	(0.015)	(0.0090)	(0.016)	(0.011)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	One	Two+	Two	Three+	Three	Four+
Dependent Variable: Ln monthly electricity use							
M-44	0.021***	0.021**	0.011	0.031**	0.0016	0.021	-0.0080
	(0.0058)	(0.0086)	(0.0076)	(0.015)	(0.0086)	(0.015)	(0.011)
M-43	0.014**	0.0048	0.0100	0.031*	-0.000074	0.0058	-0.0033
	(0.0059)	(0.0089)	(0.0078)	(0.016)	(0.0086)	(0.015)	(0.010)
M-42	0.017***	0.0030	0.017**	0.042***	0.0052	0.020	-0.0021
	(0.0059)	(0.0092)	(0.0075)	(0.014)	(0.0088)	(0.015)	(0.011)
M-41	0.017***	0.0065	0.015**	0.046***	0.00034	0.021	-0.0096
	(0.0058)	(0.0092)	(0.0074)	(0.015)	(0.0085)	(0.015)	(0.010)
M-40	0.022***	0.019**	0.014**	0.044***	0.00029	0.012	-0.0055
	(0.0055)	(0.0085)	(0.0071)	(0.014)	(0.0081)	(0.014)	(0.0097)
M-39	0.016***	0.019**	0.0052	0.025*	-0.0046	0.0055	-0.0098
	(0.0056)	(0.0087)	(0.0073)	(0.015)	(0.0083)	(0.014)	(0.010)
M-38	0.023***	0.024***	0.014**	0.036***	0.0029	0.0034	0.0021
	(0.0053)	(0.0084)	(0.0068)	(0.014)	(0.0077)	(0.013)	(0.0092)
M-37	0.023***	0.025***	0.013*	0.036***	0.0028	-0.0070	0.0066
	(0.0054)	(0.0082)	(0.0071)	(0.013)	(0.0084)	(0.016)	(0.0097)
M-36	0.019***	0.018**	0.012	0.029**	0.0026	-0.0033	0.0046
	(0.0054)	(0.0082)	(0.0071)	(0.013)	(0.0084)	(0.015)	(0.010)
M-35	0.017***	0.021***	0.0072	0.014	0.0023	-0.0010	0.0032
	(0.0054)	(0.0079)	(0.0071)	(0.013)	(0.0084)	(0.016)	(0.0097)
M-34	0.015***	0.015*	0.0070	0.014	0.0021	0.0040	0.00062
	(0.0053)	(0.0082)	(0.0068)	(0.013)	(0.0079)	(0.013)	(0.0097)
M-33	0.0058	0.0088	-0.0029	0.0099	-0.0096	-0.0058	-0.012
	(0.0051)	(0.0079)	(0.0066)	(0.013)	(0.0076)	(0.013)	(0.0094)
M-32	0.0014	0.0061	-0.0082	-0.0030	-0.012	0.00071	-0.018*
	(0.0049)	(0.0072)	(0.0067)	(0.014)	(0.0075)	(0.013)	(0.0092)
M-31	0.0014	-0.00049	-0.0027	0.0085	-0.0084	0.0037	-0.014
	(0.0051)	(0.0074)	(0.0070)	(0.012)	(0.0084)	(0.012)	(0.011)
M-30	0.0049	0.0033	0.00085	0.0077	-0.0030	0.00041	-0.0051
	(0.0047)	(0.0072)	(0.0063)	(0.012)	(0.0072)	(0.013)	(0.0085)
M-29	0.0061	0.0094	-0.0014	0.0057	-0.0053	-0.000065	-0.0082
	(0.0046)	(0.0067)	(0.0062)	(0.013)	(0.0070)	(0.013)	(0.0084)
M-28	0.011**	0.016**	0.0017	0.014	-0.0040	0.0029	-0.0075
	(0.0044)	(0.0065)	(0.0059)	(0.012)	(0.0067)	(0.012)	(0.0082)
M-27	0.0065	0.0081	0.0014	0.011	-0.0030	0.0028	-0.0060
	(0.0042)	(0.0065)	(0.0055)	(0.010)	(0.0064)	(0.011)	(0.0079)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	One	Two+	Two	Three+	Three	Four+
Dependent Variable: Ln monthly electricity use							
M-26	0.0082**	0.013**	0.0011	0.0067	-0.0018	-0.0048	-0.0010
	(0.0041)	(0.0064)	(0.0053)	(0.010)	(0.0062)	(0.011)	(0.0075)
M-25	0.0054	0.010*	-0.0020	0.0032	-0.0048	-0.0079	-0.0038
	(0.0040)	(0.0060)	(0.0054)	(0.010)	(0.0064)	(0.010)	(0.0079)
M-24	0.0060	0.0069	0.0025	0.0093	-0.00068	-0.0039	0.00053
	(0.0039)	(0.0056)	(0.0054)	(0.010)	(0.0063)	(0.010)	(0.0078)
M-11	0.0045	0.0042	0.0063	0.012	0.0046	-0.00089	0.0069
	(0.0035)	(0.0049)	(0.0048)	(0.0095)	(0.0056)	(0.010)	(0.0067)
M-10	0.0045	0.0031	0.0073	0.014	0.0050	0.0058	0.0047
	(0.0036)	(0.0050)	(0.0051)	(0.0093)	(0.0061)	(0.0095)	(0.0076)
M-9	0.0011	0.00014	0.0036	0.0090	0.0017	0.0065	-0.00016
	(0.0037)	(0.0052)	(0.0053)	(0.010)	(0.0061)	(0.011)	(0.0074)
M-8	-0.0048	-0.0086	0.00041	0.0058	-0.0014	0.0051	-0.0039
	(0.0040)	(0.0058)	(0.0055)	(0.011)	(0.0063)	(0.011)	(0.0078)
M-7	-0.0013	-0.0018	0.00086	0.0027	0.00057	0.0047	-0.00081
	(0.0040)	(0.0055)	(0.0058)	(0.012)	(0.0066)	(0.011)	(0.0080)
M-6	0.00059	0.0016	0.0012	-0.0083	0.0057	0.0100	0.0043
	(0.0040)	(0.0054)	(0.0059)	(0.012)	(0.0066)	(0.011)	(0.0081)
M-5	0.00023	0.0021	0.00017	-0.0034	0.0022	0.0054	0.0012
	(0.0040)	(0.0054)	(0.0059)	(0.012)	(0.0068)	(0.011)	(0.0084)
M-4	-0.0033	-0.0029	-0.0018	0.0055	-0.0044	-0.014	-0.000036
	(0.0040)	(0.0056)	(0.0056)	(0.011)	(0.0065)	(0.012)	(0.0077)
M-3	-0.0087**	-0.0054	-0.010*	0.0054	-0.016**	-0.035***	-0.0080
	(0.0041)	(0.0057)	(0.0058)	(0.012)	(0.0067)	(0.013)	(0.0077)
M-2	-0.0039	0.0017	-0.0076	-0.0039	-0.0087	-0.025**	-0.0021
	(0.0040)	(0.0055)	(0.0057)	(0.011)	(0.0064)	(0.012)	(0.0076)
M-1	-0.012***	-0.0010	-0.020***	-0.028**	-0.017**	-0.038***	-0.0080
	(0.0042)	(0.0058)	(0.0060)	(0.011)	(0.0072)	(0.012)	(0.0088)
M0	-0.015***	0.0019	-0.029***	-0.028***	-0.029***	-0.047***	-0.021***
	(0.0041)	(0.0058)	(0.0058)	(0.011)	(0.0067)	(0.012)	(0.0081)
M1	-0.029***	-0.0069	-0.049***	-0.043***	-0.051***	-0.074***	-0.042***
	(0.0044)	(0.0062)	(0.0060)	(0.011)	(0.0069)	(0.013)	(0.0080)
M2	-0.035***	-0.012**	-0.056***	-0.047***	-0.059***	-0.077***	-0.052***
	(0.0043)	(0.0061)	(0.0059)	(0.011)	(0.0069)	(0.012)	(0.0082)
M3	-0.039***	-0.014**	-0.062***	-0.054***	-0.064***	-0.071***	-0.062***
	(0.0044)	(0.0061)	(0.0060)	(0.011)	(0.0072)	(0.013)	(0.0085)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	One	Two+	Two	Three+	Three	Four+
Dependent Variable: Ln monthly electricity use							
M4	-0.043***	-0.016***	-0.067***	-0.056***	-0.071***	-0.069***	-0.072***
	(0.0043)	(0.0060)	(0.0059)	(0.010)	(0.0070)	(0.013)	(0.0083)
M5	-0.039***	-0.012*	-0.063***	-0.052***	-0.068***	-0.070***	-0.066***
	(0.0045)	(0.0061)	(0.0063)	(0.011)	(0.0074)	(0.013)	(0.0088)
M6	-0.037***	-0.0100*	-0.061***	-0.064***	-0.060***	-0.073***	-0.054***
	(0.0046)	(0.0060)	(0.0067)	(0.013)	(0.0075)	(0.015)	(0.0086)
M7	-0.035***	-0.0056	-0.062***	-0.059***	-0.062***	-0.071***	-0.059***
	(0.0045)	(0.0062)	(0.0062)	(0.011)	(0.0073)	(0.013)	(0.0086)
M8	-0.038***	-0.0043	-0.068***	-0.060***	-0.072***	-0.079***	-0.069***
	(0.0045)	(0.0064)	(0.0061)	(0.011)	(0.0071)	(0.013)	(0.0084)
M9	-0.045***	-0.011*	-0.077***	-0.061***	-0.083***	-0.089***	-0.081***
	(0.0048)	(0.0069)	(0.0063)	(0.012)	(0.0073)	(0.014)	(0.0085)
M10	-0.039***	-0.0032	-0.073***	-0.062***	-0.077***	-0.082***	-0.075***
	(0.0048)	(0.0066)	(0.0066)	(0.013)	(0.0076)	(0.013)	(0.0091)
M11	-0.040***	-0.0026	-0.074***	-0.062***	-0.079***	-0.089***	-0.075***
	(0.0046)	(0.0063)	(0.0064)	(0.012)	(0.0075)	(0.014)	(0.0089)
M12	-0.038***	0.00075	-0.073***	-0.049***	-0.083***	-0.086***	-0.082***
	(0.0046)	(0.0062)	(0.0065)	(0.012)	(0.0075)	(0.014)	(0.0088)
Gap 1	-0.053***	0	-0.076***	-0.061***	-0.081***	-0.073***	-0.084***
	(0.0056)	(.)	(0.0062)	(0.011)	(0.0073)	(0.013)	(0.0087)
M13	-0.050***	-0.0084	-0.089***	-0.059***	-0.10***	-0.094***	-0.10***
	(0.0048)	(0.0065)	(0.0066)	(0.012)	(0.0077)	(0.014)	(0.0090)
M14	-0.053***	-0.013**	-0.090***	-0.053***	-0.11***	-0.100***	-0.11***
	(0.0050)	(0.0067)	(0.0069)	(0.014)	(0.0079)	(0.014)	(0.0094)
M15	-0.049***	-0.0076	-0.088***	-0.037***	-0.11***	-0.10***	-0.11***
	(0.0050)	(0.0066)	(0.0070)	(0.013)	(0.0082)	(0.015)	(0.0095)
M16	-0.047***	-0.0076	-0.085***	-0.036***	-0.11***	-0.096***	-0.11***
	(0.0049)	(0.0065)	(0.0069)	(0.013)	(0.0080)	(0.015)	(0.0094)
M17	-0.044***	-0.0081	-0.077***	-0.032**	-0.097***	-0.092***	-0.099***
	(0.0049)	(0.0066)	(0.0069)	(0.013)	(0.0078)	(0.014)	(0.0092)
M18	-0.048***	-0.013*	-0.081***	-0.043***	-0.096***	-0.089***	-0.099***
	(0.0050)	(0.0067)	(0.0071)	(0.014)	(0.0081)	(0.014)	(0.0096)
M19	-0.049***	-0.012*	-0.083***	-0.042***	-0.10***	-0.10***	-0.10***
	(0.0051)	(0.0067)	(0.0072)	(0.013)	(0.0084)	(0.015)	(0.0099)
M20	-0.050***	-0.012*	-0.086***	-0.046***	-0.10***	-0.10***	-0.10***
	(0.0051)	(0.0069)	(0.0070)	(0.013)	(0.0081)	(0.015)	(0.0096)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	One	Two+	Two	Three+	Three	Four+
Dependent Variable: Ln monthly electricity use							
M21	-0.053***	-0.011	-0.093***	-0.060***	-0.11***	-0.095***	-0.11***
	(0.0052)	(0.0068)	(0.0073)	(0.015)	(0.0082)	(0.014)	(0.0098)
M22	-0.043***	0.0025	-0.086***	-0.055***	-0.100***	-0.088***	-0.10***
	(0.0052)	(0.0067)	(0.0073)	(0.015)	(0.0081)	(0.015)	(0.0096)
M23	-0.044***	-0.0013	-0.084***	-0.038***	-0.10***	-0.090***	-0.11***
	(0.0052)	(0.0068)	(0.0073)	(0.014)	(0.0085)	(0.015)	(0.010)
M24	-0.040***	0.0016	-0.080***	-0.018	-0.11***	-0.095***	-0.11***
	(0.0052)	(0.0069)	(0.0073)	(0.013)	(0.0085)	(0.016)	(0.0099)
Gap 2	-0.062***	0	-0.085***	0	-0.099***	-0.083***	-0.11***
	(0.0082)	(.)	(0.0088)	(.)	(0.0092)	(0.017)	(0.011)
M25	-0.042***	0.00020	-0.083***	-0.022*	-0.11***	-0.078***	-0.12***
	(0.0052)	(0.0069)	(0.0073)	(0.013)	(0.0085)	(0.015)	(0.010)
M26	-0.045***	-0.0016	-0.086***	-0.020	-0.11***	-0.088***	-0.13***
	(0.0052)	(0.0068)	(0.0074)	(0.013)	(0.0086)	(0.015)	(0.010)
M27	-0.042***	0.00018	-0.082***	-0.0093	-0.11***	-0.081***	-0.13***
	(0.0053)	(0.0069)	(0.0074)	(0.013)	(0.0086)	(0.014)	(0.010)
M28	-0.041***	-0.00085	-0.080***	-0.0071	-0.11***	-0.073***	-0.13***
	(0.0053)	(0.0069)	(0.0074)	(0.014)	(0.0085)	(0.015)	(0.010)
M29	-0.040***	-0.000075	-0.078***	-0.0094	-0.11***	-0.074***	-0.12***
	(0.0054)	(0.0071)	(0.0076)	(0.014)	(0.0086)	(0.015)	(0.010)
M30	-0.042***	-0.0052	-0.078***	-0.016	-0.10***	-0.075***	-0.12***
	(0.0055)	(0.0073)	(0.0076)	(0.014)	(0.0087)	(0.016)	(0.010)
M31	-0.044***	-0.0029	-0.084***	-0.016	-0.11***	-0.076***	-0.13***
	(0.0056)	(0.0072)	(0.0079)	(0.014)	(0.0093)	(0.017)	(0.011)
M32	-0.040***	0.0021	-0.082***	-0.028**	-0.11***	-0.070***	-0.12***
	(0.0055)	(0.0072)	(0.0077)	(0.014)	(0.0091)	(0.017)	(0.011)
M33	-0.042***	0.0020	-0.085***	-0.034**	-0.11***	-0.058***	-0.13***
	(0.0057)	(0.0075)	(0.0078)	(0.014)	(0.0091)	(0.017)	(0.011)
M34	-0.038***	0.0054	-0.082***	-0.023	-0.11***	-0.044***	-0.13***
	(0.0056)	(0.0073)	(0.0079)	(0.014)	(0.0091)	(0.016)	(0.011)
M35	-0.036***	0.0090	-0.082***	-0.012	-0.11***	-0.051***	-0.13***
	(0.0056)	(0.0072)	(0.0079)	(0.015)	(0.0092)	(0.017)	(0.011)
M36	-0.030***	0.018**	-0.080***	-0.0027	-0.11***	-0.051***	-0.13***
	(0.0057)	(0.0073)	(0.0081)	(0.014)	(0.0095)	(0.017)	(0.011)
Gap 3	-0.050***	0	-0.076***	0	-0.091***	0	-0.098***
	(0.010)	(.)	(0.011)	(.)	(0.012)	(.)	(0.012)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	One	Two+	Two	Three+	Three	Four+
Dependent Variable: Ln monthly electricity use							
M37	-0.034***	0.0092	-0.080***	-0.0017	-0.11***	-0.037**	-0.14***
	(0.0058)	(0.0075)	(0.0083)	(0.014)	(0.0097)	(0.017)	(0.011)
M38	-0.038***	0.0060	-0.084***	0.0052	-0.12***	-0.045***	-0.15***
	(0.0060)	(0.0074)	(0.0086)	(0.015)	(0.010)	(0.017)	(0.012)
M39	-0.037***	0.0094	-0.086***	-0.00075	-0.12***	-0.047***	-0.15***
	(0.0059)	(0.0074)	(0.0084)	(0.015)	(0.0098)	(0.016)	(0.012)
M40	-0.034***	0.0086	-0.079***	0.00053	-0.11***	-0.053***	-0.13***
	(0.0060)	(0.0075)	(0.0086)	(0.015)	(0.010)	(0.018)	(0.012)
M41	-0.033***	0.011	-0.080***	-0.012	-0.11***	-0.051***	-0.13***
	(0.0061)	(0.0076)	(0.0088)	(0.016)	(0.010)	(0.018)	(0.012)
M42	-0.033***	0.0073	-0.077***	-0.011	-0.11***	-0.041**	-0.13***
	(0.0061)	(0.0077)	(0.0088)	(0.015)	(0.010)	(0.019)	(0.012)
M43	-0.032***	0.0065	-0.074***	-0.0091	-0.10***	-0.038*	-0.12***
	(0.0061)	(0.0077)	(0.0088)	(0.014)	(0.011)	(0.020)	(0.012)
M44	-0.031***	0.0100	-0.076***	-0.015	-0.10***	-0.031	-0.13***
	(0.0062)	(0.0077)	(0.0090)	(0.016)	(0.011)	(0.020)	(0.012)
M45	-0.030***	0.011	-0.076***	-0.012	-0.11***	-0.023	-0.13***
	(0.0063)	(0.0077)	(0.0093)	(0.016)	(0.011)	(0.020)	(0.013)
M46	-0.027***	0.015*	-0.076***	-0.017	-0.10***	-0.024	-0.13***
	(0.0063)	(0.0079)	(0.0092)	(0.016)	(0.011)	(0.020)	(0.013)
M47	-0.028***	0.015*	-0.078***	-0.017	-0.11***	-0.048**	-0.12***
	(0.0064)	(0.0080)	(0.0093)	(0.017)	(0.011)	(0.021)	(0.012)
M48	-0.024***	0.018**	-0.074***	-0.0039	-0.11***	-0.059**	-0.12***
	(0.0065)	(0.0080)	(0.0096)	(0.016)	(0.011)	(0.024)	(0.013)
Gap 4	-0.050***	0	-0.080***	0	-0.097***	0	-0.11***
	(0.015)	(.)	(0.016)	(.)	(0.016)	(.)	(0.017)
M49	-0.029***	0.012	-0.077***	-0.0028	-0.11***	-0.068***	-0.13***
	(0.0065)	(0.0080)	(0.0094)	(0.016)	(0.011)	(0.022)	(0.013)
M50	-0.028***	0.011	-0.076***	-0.0069	-0.11***	-0.061***	-0.12***
	(0.0066)	(0.0082)	(0.0097)	(0.016)	(0.011)	(0.019)	(0.013)
M51	-0.024***	0.012	-0.069***	0.00069	-0.10***	-0.059***	-0.12***
	(0.0067)	(0.0083)	(0.0097)	(0.016)	(0.012)	(0.020)	(0.014)
M52	-0.025***	0.011	-0.072***	-0.0011	-0.11***	-0.057***	-0.12***
	(0.0067)	(0.0083)	(0.0098)	(0.016)	(0.012)	(0.020)	(0.014)
M53	-0.025***	0.014	-0.076***	-0.029	-0.100***	-0.038*	-0.12***
	(0.0072)	(0.0084)	(0.011)	(0.022)	(0.013)	(0.020)	(0.015)

	(1) All	(2) One	(3) Two+	(4) Two	(5) Three+	(6) Three	(7) Four+
Dependent Variable: Ln monthly electricity use							
M54	-0.024***	0.016*	-0.077***	-0.022	-0.10***	-0.055**	-0.12***
	(0.0070)	(0.0083)	(0.011)	(0.017)	(0.014)	(0.027)	(0.015)
M55	-0.026***	0.017*	-0.082***	-0.027	-0.11***	-0.029	-0.13***
	(0.0071)	(0.0085)	(0.011)	(0.018)	(0.013)	(0.023)	(0.016)
M56	-0.028***	0.014	-0.083***	-0.032*	-0.11***	-0.022	-0.13***
	(0.0073)	(0.0088)	(0.011)	(0.019)	(0.014)	(0.025)	(0.016)
M57	-0.031***	0.0070	-0.082***	-0.040**	-0.10***	-0.031	-0.13***
	(0.0074)	(0.0089)	(0.012)	(0.020)	(0.014)	(0.026)	(0.016)
M58	-0.029***	0.010	-0.083***	-0.033*	-0.11***	-0.052	-0.13***
	(0.0076)	(0.0090)	(0.012)	(0.020)	(0.015)	(0.032)	(0.016)
M59	-0.032***	0.0085	-0.087***	-0.025	-0.12***	-0.065*	-0.13***
	(0.0075)	(0.0088)	(0.012)	(0.019)	(0.015)	(0.038)	(0.016)
M60	-0.025***	0.017*	-0.085***	-0.013	-0.12***	-0.079**	-0.14***
	(0.0076)	(0.0089)	(0.012)	(0.019)	(0.015)	(0.038)	(0.016)
Gap 5	-0.050***	0	-0.083***	0	-0.10***	0	-0.11***
	(0.018)	(.)	(0.019)	(.)	(0.019)	(.)	(0.020)
M61	-0.023***	0.016*	-0.077***	0.0083	-0.12***	-0.046*	-0.15***
	(0.0076)	(0.0091)	(0.012)	(0.019)	(0.015)	(0.027)	(0.017)
M62	-0.022***	0.016*	-0.078***	0.0015	-0.12***	-0.041	-0.14***
	(0.0077)	(0.0091)	(0.012)	(0.019)	(0.015)	(0.029)	(0.017)
M63	-0.024***	0.014	-0.079***	0.0042	-0.12***	-0.032	-0.15***
	(0.0077)	(0.0092)	(0.012)	(0.019)	(0.015)	(0.028)	(0.018)
M64	-0.026***	0.011	-0.081***	0.0036	-0.13***	-0.046	-0.15***
	(0.0079)	(0.0094)	(0.013)	(0.019)	(0.016)	(0.031)	(0.018)
M65	-0.024***	0.011	-0.077***	0.014	-0.13***	-0.032	-0.16***
	(0.0080)	(0.0094)	(0.013)	(0.018)	(0.016)	(0.027)	(0.019)
M66	-0.021***	0.011	-0.068***	-0.00079	-0.11***	-0.021	-0.13***
	(0.0080)	(0.0094)	(0.013)	(0.019)	(0.016)	(0.028)	(0.019)
M67	-0.020**	0.0065	-0.057***	0.011	-0.093***	-0.0025	-0.12***
	(0.0082)	(0.0098)	(0.013)	(0.019)	(0.016)	(0.027)	(0.019)
M68	-0.022***	0.0080	-0.066***	0.0019	-0.10***	-0.017	-0.13***
	(0.0082)	(0.0097)	(0.013)	(0.019)	(0.017)	(0.032)	(0.019)
M69	-0.024***	0.0033	-0.062***	-0.0059	-0.093***	-0.028	-0.12***
	(0.0086)	(0.010)	(0.014)	(0.021)	(0.018)	(0.033)	(0.020)
M70	-0.024***	0.0064	-0.070***	-0.020	-0.097***	-0.057*	-0.11***
	(0.0085)	(0.010)	(0.014)	(0.021)	(0.018)	(0.030)	(0.021)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	One	Two+	Two	Three+	Three	Four+
Dependent Variable: Ln monthly electricity use							
M71	-0.022***	0.012	-0.076***	-0.0046	-0.12***	-0.064**	-0.14***
	(0.0085)	(0.0100)	(0.014)	(0.019)	(0.019)	(0.030)	(0.023)
M72	-0.027***	0.0089	-0.086***	-0.016	-0.13***	-0.079**	-0.15***
	(0.0085)	(0.0100)	(0.014)	(0.019)	(0.018)	(0.034)	(0.021)
Post	-0.020***	0.0072	-0.064***	-0.0087	-0.10***	-0.053*	-0.12***
	(0.0077)	(0.0087)	(0.013)	(0.018)	(0.017)	(0.030)	(0.019)
Observations	1513669	1184992	1190510	958670	1093673	928673	1026833

All specifications include individual and date fixed effects, participant and non-participant households, and are restricted to participant households that begin their subsequent challenges within 12 months of finishing their previous challenge. In addition, specifications have the following restrictions: (1) has no further restrictions; (2) is households that undertake a single challenge only; (3) is households that undertake two or more challenges; (4) households that undertake two challenges only; (5) households undertaking three or more challenges; (6) three challenges only; (7) Four or more challenges. Standard errors are clustered at the household level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$ denote significance levels where 0 is defined as the second year pre-treatment and consists of months M-12 to M-23.