

2014

THREE ESSAYS ON THE INCENTIVE STRUCTURE OF ENERGY CONSERVATION PROGRAMS

Edson Ogochukwu Okwelum
University of Rhode Island, edson_okwelum@my.uri.edu

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss



Part of the [Natural Resource Economics Commons](#)

Terms of Use

All rights reserved under copyright.

Recommended Citation

Okwelum, Edson Ogochukwu, "THREE ESSAYS ON THE INCENTIVE STRUCTURE OF ENERGY CONSERVATION PROGRAMS" (2014). *Open Access Dissertations*. Paper 269.
https://digitalcommons.uri.edu/oa_diss/269

This Dissertation is brought to you by the University of Rhode Island. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu. For permission to reuse copyrighted content, contact the author directly.

THREE ESSAYS ON THE INCENTIVE STRUCTURE OF
ENERGY CONSERVATION PROGRAMS

BY

EDSON OGOCHUKWU OKWELUM

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
NATURAL RESOURCE ECONOMICS

UNIVERSITY OF RHODE ISLAND

2014

DOCTOR OF PHILOSOPHY DISSERTATION
OF
EDSON OGOCHUKWU OKWELUM

APPROVED:

Dissertation Committee:

Major Professor James Opaluch

Corey Lang

Marion Gold

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2014

ABSTRACT

Three Essays on the Incentive Structure of Energy Conservation Programs

By

Edson Ogochukwu Okwelum

Doctor Philosophy in Environmental and Natural Resource Economics

University of Rhode Island

Professor James Opaluch, Chair

This dissertation is comprised of three related essays examining the potential effectiveness of government energy efficiency programs from both the producer and consumer perspectives. The first chapter is based on a paper I coauthored with Corey Lang. In this manuscript, I address the question of whether strategic behavior by consumers could result in the erosion of energy savings in a demand response program. Understanding how the strategic behavior of consumers affects the net benefits from a demand response program has policy implications because of the increasing importance that demand response has come to play in utility load and reliability management during peak times. Using data from a large field experiment in California in 2007, we test the hypothesis that under a technology program, consumers' strategic behavior results in outcomes that are opposite what is obtainable under a program with price incentive or based of behavior. Chapter II is also an empirical study which explores how the preferences of consumers for large and heavy vehicles imposes costs on society in the form of external costs of accident. This chapter looks at how fleet changes in weight distribution due to corporate average fuel

economy and consumer demand for heavier vehicles results in fatalities. It is important to understand how consumer behavior affects the accident rates so that one can obtain unbiased estimates of accident costs that go into benefit-cost analysis of the impact of regulations in automobiles. Chapter three addresses how unobserved heterogeneity and sorting affect the estimates of the consumer willingness to pay for reduction in future gasoline costs. This tradeoff is important to policy makers and manufacturers because it could help explain why manufacturers fail to adopt technologies for which the fuel savings far outweigh the costs. The remainder of the abstract provides a more detailed outline of the three essays.

Chapter 1 explores strategic behavior by consumers in demand response programs.

The chapter looks at how the confluences of consumer strategic behavior, technology and price incentives in demand response lead to conflicting outcomes. Demand response has become an important tool for utilities to manage load during peak periods. While the effects of demand response programs on peak load reductions are well studied and intuitive, assessments typically fail to recognize the potential for off-peak behavioral responses that may mitigate the total benefits of the program. Using smart meter consumption data on residential air conditioning units enrolled in a direct load control program, this paper examines changes in consumption prior to and after curtailment events. The manuscript rigorously estimates both peak and off peak changes in consumption to better understand net benefits of DR programs. We examine Pacific Gas and Electric's (PG&E) SmartAC program, which is designed to reduce peak cooling load by directly controlling air conditioning units of participants.

Usually, program evaluation uses the availability of a control group to obtain unbiased estimates of program impacts. However, we do not have the luxury of a control group in our data. Instead we estimate unit-specific non-linear consumption models and then compare load on event days to predicted load. Importantly, we use data from the summer 2007 training period from a stratified random sample of 294 participating AC units. In contrast to subsequent summers and normal DR designs, during the training period curtailment days were called for many different temperature levels, not just the hottest days. This aspect means we are not predicting out of sample, and we can have greater confidence in our econometric evidence. Our methodology is validated by a falsification test in which we find no changes in consumption during peak or off-peak times on non-curtailment days that match the temperature and timing profile of actual curtailment days.

Several key results emerge from this analysis. First, we confirm that the SmartAC program, like other DR and DLC programs, reduces peak load during event days. In this case, peak consumption was reduced 19% on average. However, we also find substantial increases in consumption in the hours preceding and the hours following an event by 8% and 7% respectively. Essentially, load is being displaced from peak to off-peak times. These behavioral changes mitigate the monetary benefits of the program by 41%.

Chapter II addresses the question of how much accident risks would change if consumers who purchase larger, heavier vehicles of the existing fleet were to behave strategically given regulatory constraints imposed by corporate average fuel economy. The paper estimates both the risk of dying given a crash and the crash frequencies of

different vehicles in the fleet. We use a unique data set that combines data from fatality analysis reporting system (FARS) and NASS¹ General Estimates Systems. The estimation strategy used corrects for selection bias due to moral hazard problems. The two equations are estimated simultaneously using simulated maximum likelihood without the need for exclusion restrictions (Green 2003, 2007). This approach allows us to obtain estimates that are causal.

One of the main issues that have been left unanswered in earlier works is the existence of Peltzman-type moral hazard problems and endogeneity of vehicle choice. In the presence of Peltzman-type moral hazard problems, drivers in heavier vehicles will find it advantageous to change their behavior in ways that have opposite effect to regulations. From a public policy perspective if increasing the weight of a vehicle increases the probability of its being involved in collisions, then we might be interested in more than the impact of heavier vehicles conditional on a crash occurring. We find that increasing the weight of any vehicle increases the probability of it being the heavier vehicle in a collision and the probability of it being involved in a fatal accident. The external costs translate to about 38 cents/gallon of gasoline.

Chapter III explores how individual unobserved heterogeneity in tastes could lead to different customers sorting into different classes of fuel efficient vehicles and therefore affect estimates of consumer's ability to tradeoff vehicle costs and discounted future gasoline costs. The research question is important for several reasons. First, it is interesting from a theoretical basis if consumers make this trade off optimally. Many theoretical models in economics make the fundamental assumption that consumers equally weigh current and future events when making decisions today.

¹ National Automotive Sampling System

However, plenty empirical and laboratory evidence conclude the contrary. Second, Regulatory Impact Analyses (RIA) of regulations on energy efficiency standards usually find that the benefits of these programs are in many times larger than the costs, with the fuel savings over the life of the vehicle accounting for the majority of the benefits. However, if fuel savings are large relative to costs, then the question is why are manufacturers not incorporating these technologies in their product offerings? Therefore understanding energy paradox is critical for evaluating the standards and regulations so that policy makers can identify baseline for assessment in cost-benefit analyses.

In this paper, we contribute to the literature on *Energy Paradox in Fuel Economy* by providing empirical evidence that unobserved consumer heterogeneity could result in different consumers sorting into different classes of fuel efficient cars which results in consumers undervaluing fuel economy. Unlike existing literature in this topic, this paper accounts for sorting bias due unobserved heterogeneity by using a random coefficient logit with error components (Train and Winston 2007). We pair the random coefficient discrete choice model with a supply side model in which firms compete in Bertrand Nash fashion where price depends on elasticity of demand and marginal costs. With the random utility model, we can allow taste to vary across consumers. We treat the discounted operating costs and vehicle costs as random variables.

We find that a substantial portion of upper the 95% of households in our sample correctly value fuel economy (61%) as they are willing to pay \$0.99 to reduce obtain a \$1.00 discounted future gasoline costs over the lifetime of the vehicle. And 29% of the upper 95% overvalue fuel economy as they are willing to pay \$1.57. These results

are in contrast to those reported elsewhere by Alcott and Wozny (2009) but support results from the simulation study by Bento et al 2012 who find that not accounting for sorting bias and consumer heterogeneity could lead to a conclusion that consumers undervalue fuel economy.

ACKNOWLEDGMENTS

I must acknowledge several individuals who taught, advised, and supported me in different ways in the last four years. Professor Jim Opaluch has been an outstanding advisor during this period. His thoughtfulness, insights and careful consideration have helped me through the graduate program, especially the dissertation process. I cannot express how grateful I am in words alone. I am eternally grateful for the opportunity that he has provided me.

I would like to express my sincere gratitude to Professor Corey Lang for his guidance and willingness to be part of my dissertation committee and involve me in some of his research. I learnt a great deal from working with Corey, and I am the better for it today. That I am able to complete my dissertation process, owes a lot him. I have learned so much from Professor Hiro Uchida. I am really grateful for his insights and ideas as well as willingness to discuss research ideas. I gained a lot of economic intuition discussing with Hiro, and I am profoundly grateful for that. I am also grateful to Emi Uchida for all the help and guidance during my graduate program. I cannot thank her enough for the interdisciplinary training and all the support that she has provided to me in these four years.

Part of my graduate research was funded by the Outreach Centre. I am very grateful to the people at the Outreach Centre for their financial support and allowing me to participate in “Renewable Energy Siting Partnership” project. I am especially grateful to Marion Gold for the financial support that was provided by the Outreach Centre and for agreeing to serve in my dissertation committee given her busy schedule.

Professor John Burkett provided comments and suggestions to earlier draft versions of the second and third manuscripts. I am very grateful for taking time to review the drafts and provide insightful comments. I am also indebted to Nathaniel Merritt for all the support, friendship and reading draft copies of my dissertation. His comments and suggestions were invaluable in writing this dissertation, and I cannot thank him enough.

The data set for the first manuscript was generously provided by Dr Abigail Anthony of Environmental North East. I am very grateful to her for allowing us the use of the datasets. I would also like to thank the following colleagues and friends at ENRE who have helped in several ways during my graduate studies: Kyle Montanio, Eleanor, Pratheesh, Ganesh, Brandon Elsner, Andy Boslet, Matt Altonji. They all provided insights and ideas on how to move forward with different challenges. I am grateful for all the help.

A very special thank you to Judy Palmer for all the support throughout my graduate program at ENRE. I am also grateful to the graduate student body at ENRE; I would not have done this without them.

Kemi and Alex have been a tremendous source of inspiration, support and love during this period and I am eternally grateful for their patience and understanding. I could not have done this without their full understanding and cooperation. And lastly, my mother, sisters and brothers, I am grateful to all the support and encouragement during my graduate studies.

PREFACE

This dissertation is written in three manuscript form. The first manuscript is based on a paper I coauthored with Corey Lang and has been submitted to the Energy Economics Journal for publication.

Manuscript 1: The Mitigating Effect of Strategic Behavior on the Net Benefits of a Direct Load Control Program

Manuscript 2: Fuel Economy and Vehicle Safety: Selection on Moral Hazard and Estimates of External Accident Costs

Manuscript 3: Heterogeneity and Sorting in Consumer Valuation of Energy Efficiency Investments: Empirical Evidence from US Automobile Market

Table of Contents

ABSTRACT ii

ACKNOWLEDGMENTS viii

PREFACE..... x

Table of Contents..... xi

LIST OF TABLES xiii

LIST OF FIGURES xiv

List of Abbreviations and Acronyms xv

INTRODUCTION..... 1

MANUSCRIPT 1 8

Abstract 9

Introduction 10

Background on the SmartAC Program 13

Methodology..... 16

Data 19

Results..... 21

Policy Discussions 31

Conclusion 35

REFERENCES 36

Figures and Tables..... 39

Online Appendix 47

MANUSCRIPT 2..... 63

Introduction 65

Theoretical Background..... 70

Data 75

Econometric Model..... 79

Results..... 82

Policy Simulations..... 89

Conclusion	95
REFERENCES	98
Figures and Tables	101
MANUSCRIPT 3	112
Abstract.....	113
Introduction	114
Methodology and Conceptual Framework	120
Data	127
Estimation and Identification Strategy	132
Results	136
Conclusions	141
REFERENCES	143
CONCLUSION	154

LIST OF TABLES

MANUSCRIPT 1: The Mitigating Effect of Strategic Behavior on the Net Benefits of a Direct Load Control Program

Table 1: Average Precooling, Load Impact Response & Rebound during Curtailment Days by all Units.....	43
Table 2: Heterogeneity in Response to SmartAC Program.....	44
Table 3: Average Precooling, Load Response and Rebound on Proxy days.....	45
Table 4: Net Savings with and without Precooling and Rebound.....	46

MANUSCRIPT 2: Fuel Economy and Vehicle Safety: Selection on Moral Hazard and Estimates of External Accident Costs

Table 1: Summary Statistics for Two Vehicle Collision.....	104
Table 2: Regression Estimates of Number Occupant Fatalities- No Selection.....	105
Table 3: First-Stage Regression Results.....	106
Table 4: Second Stage Regression Results.....	107
Table 5: Second Stage Regression Estimates for Serious Injury Model.....	108
Table 6: Alternative Specification.....	109
Table 7 Falsification Test.....	110
Table 8: Accident-Related Externalities.....	111

MANUSCRIPT 3: Heterogeneity and Sorting in Consumer Valuation of Energy Efficiency Investments: Empirical Evidence from US Automobile Market

Table 1: Description of Sample Households who Bought New Vehicles.....	148
Table 2: Summary Statistics of Sample Vehicle Characteristics.....	149
Table 3: ARIMA (1,1,1) Model Annual US Retail Gas Price Data 1994-2005.....	150
Table 4: Parameter Estimates of Consumer Demand for Fuel Economy.....	151
Table 5: Cost Side Parameter Estimates.....	152
Table 6: Robustness of Consumer Valuation of Fuel Economy.....	153

LIST OF FIGURES

MANUSCRIPT 1: The Mitigating Effect of Strategic Behavior on the Net Benefits of a Direct Load Control Program

Figure 1: Density of Load Distribution with Quartiles.....	39
Figure 2: Density of Average Daily Temperature Values and Curtailment Days.....	40
Figure 3-a: Distribution of Estimated First Breakpoint Parameter.....	41
Figure 3-b: Distribution of Estimated Second Breakpoint Parameter.....	41
Figure 4: Aggregate Precooling over Time.....	42

MANUSCRIPT 2: Fuel Economy and Vehicle Safety: Selection on Moral Hazard and Estimates of External Accident Costs

Figure 1-A: Mean-Variance Relationship for Number of Fatal Crashes in Months by State.....	99
Figure 1-B: Mean Variance Relation for Number of Fatalities in Struck Vehicle...	99
Figure 2: Predicted and Observed probabilities of Fatal Accidents.....	100
Figure 3: Predicted probability of Fatality in Struck Vehicle against Striking Vehicle Weight.....	101

MANUSCRIPT 3: Heterogeneity and Sorting in Consumer Valuation of Energy Efficiency Investments: Empirical Evidence from US Automobile Market

Figure 1: Plot of Natural Log Real Gasoline Price.....	146
Figure 2: Plot of Natural log of Real Gasoline Price and Forecasts.....	147
Figure 3: Distribution of WTP to Reduce Future Operating Costs.....	148

List of Abbreviations and Acronyms

AC	Air-conditioner
CAISO	California Independent System Operator
CPP	Critical Peak Pricing
CEX	Consumer Expenditure Survey
DLC	Direct Load Control
DR	Demand Response
EIA	Energy Information Administration
EPA	Environmental Protection Agency
ENRE	Environmental and Natural Resources Economics
FARS	Fatality Analysis Reporting System
GHG	Green House Gases
IEA	International Energy Association
KWh	Kilowatt Hour
KWh/h	Kilowatt Hour per Hour
LMP	Locational Marginal Pricing
MWh	Megawatt Hour
MWTP	Marginal Willingness to Pay
NASS	National Automotive Sampling System
NERC	North American Electric Reliability Corporation
NOAA	National Oceanic and Atmospheric Administration
NRC	National Research Council
PCT	Programmable Communicating Thermostats
PG&E	Pacific Gas & Electric
PTR	Peak Time Pricing
RIA	Regulatory Impact Analyses
RTP	Real Time Pricing
SMG	SmartAC Monitoring Group
TOU	Time of Use

INTRODUCTION

The dissertation explores the potential effectiveness of government energy efficiency programs and regulations from both the producer and consumer perspectives. The last three decades has seen the deployment of various energy conservation policies as part of governments' energy policies. These energy efficiency programs are driven in part by the desire to reduce energy consumption, lower greenhouse gas (GHG) emissions in order to mitigate the impacts of climate change, reduce criteria air pollutants, and to provide affordable energy services. While it is widely acknowledged that there is still a large potential for energy and greenhouse gas savings, the design of effective policies to realize these potentials is challenging. For example, estimates from the International Energy Association (IEA 2001) suggests that over 34% of the expected carbon emissions under a scenario limiting the long term concentration of greenhouse gases could be obtained solely from direct end use efficiency measures alone.

However, achieving this target presupposes that individuals will invest in energy efficiency measures and firms will invest in innovations to commercialize energy efficient technologies.

Consumer decision making and/or behavior involving energy consumption and investment in energy efficiency projects have attracted increased scrutiny recently. Policy design and evaluation have increasingly incorporated consumer behavioral responses to policy interventions. Given the inter-relationship between consumer behavior, energy and climate policy, it stands to reason that if consumers can be nudged to change their behavior and reduce their energy consumption or be more responsive to demand for weather induced changes in energy supply, we could

significantly reduce the cost of compliance to several policies. However, the general consensus is that consumers are generally not good optimizers with respect to the costs and benefits of energy conservation. Consumers generally place different life-cycle value on investments in energy efficiency than manufacturers and policymakers. Behavioral economics offers insights and different perspectives regarding how consumers value energy efficiency products or regard energy conservation efforts. Behavioral economics concepts, supported by anecdotal evidence from experimental economics suggests a set of reasons other than costs and benefits of why consumers are not good optimizers: consumers do not consider energy efficiency issues to be priority (work, family, healthcare, education and other family related matters are more important than energy efficiency; households exhibit preference and perceptual invariance in consumer choice situations; energy efficiency decision scenarios are complicated, with multiple options and difficult mathematical calculations; and people simply tend to procrastinate a lot (Kahneman and Tversky. 1979, Dhar and Glazer 1996; Abadie and Gay. 2006; Allcott and Mullainathan 2010, DellaVigna 2009). Also households show significant differences in behavior which can affect energy consumption, suggesting that intervention strategies to promote sustainable behaviors should result in significant energy savings in the future.

If the foregoing is true, then we need to answer the following important questions:

1. As consumers invest in energy efficiency projects, they will obtain economics savings over the life of such projects, an important question is how consumers will utilize the economic savings over the lifetime of their investment from an energy efficient appliance?

2. What is the impact of strategic behavior by agents on the overall impact of energy conservation programs?
3. If changes in consumers' consumption patterns occur, producers are likely to adjust their production functions and what will be the effect on the overall social welfare?
4. Are the incentives in some energy policies such that they incentivize agents to behave in ways that are in conflict with the goals of regulations? Are these incentives such that could lead to unintended consequences?
5. If consumers are different and choose durables according to some taste distributions, are baseline estimates of the marginal willingness to pay for reductions in future operating costs biased because of sorting?

How consumers weigh temporal effects of future fuel costs have policy implications. This is because the nature of the temporal weighting helps in determining if market failure exists and helps indicate whether policy prescriptions that affect initial vehicle costs such as gas guzzlers tax will reduce fuel consumption at lower costs than gasoline tax.

This dissertation answers these questions and more in three different essays. In the first manuscript, we examine how consumers' strategic behavior could erase the net benefits from a DR program. We rigorously estimate both peak and off peak changes in consumption to better understand how strategic behavior by consumers affect the net benefits of demand response (DR) programs. The main contributions of this manuscript are twofold. First, it shows that DR energy policies can lead to unintended consequences based on design and implementation. Second, we provide limited

evidence that consumers learn to erase the benefits from DR programs with time under a technology program without price incentives. We use data from a stratified random sample of 294 participating AC units. Because there is no data for a control group, we estimate unit-specific non-linear consumption models and then compare load on event days to predicted load on non-curtailment days. Importantly, we use data from the summer 2007 training period. In contrast to subsequent summers and normal DR designs, during the training period curtailment days were called for many different temperature levels, not just the hottest days. This aspect means we are not predicting out of sample, and we can have greater confidence in our econometric evidence.

We use a falsification test in which we find no changes in consumption during peak or off-peak times on non-curtailment days that match the temperature and timing profile of actual curtailment days to validate our results. First, we confirm that the SmartAC program, like other DR and DLC programs, reduces peak load during event days. In this case, peak consumption was reduced 19% on average. However, we also find substantial increases in consumption in the hours preceding and the hours following an event by 8% and 7% respectively. Essentially, load is being displaced from peak to off-peak times. These behavioral changes mitigate the monetary benefits of the program by 41%.

In the second manuscript, we also explore how the interaction of consumer behavior and energy policy could result in unintended consequences by looking at how consumers' choices of vehicle weight could result in increased fatalities under regulatory constraints. In this manuscript, , we estimate the external costs of vehicle weight in traffic accidents using a random sample of police reported accidents by

estimating the number of fatally or seriously injured occupants in two vehicle crashes. Our main challenge is that unobserved driver and occupant attributes could induce risky drivers to cluster around higher vehicles, and also cause them to switch across models as they re-optimize according to the conflicting incentives placed by gasoline and safety policies. This is because, there are interrelated human factors, which are not easily quantifiable, but are however related to vehicle weight and fatality/injury risk. We use an estimation strategy that is robust enough to correct for such biases. The estimation strategy is semiparametric and based around baseline gamma density function. The estimation strategy is able to account for unobserved driver heterogeneity. Unlike earlier studies (Jacobsen 2013; Anderson and Auffhammer 2013) our unique data set and methodology allows us to adjust for the fact that heavier vehicles are involved in more fatal accidents than lighter vehicles ensuring that the coefficients of weight are causal. We find that it is vehicle weight dispersion rather than the absolute weight of vehicles involved in multiple vehicle crashes that is the key factor determining injury severity. In a two-vehicle crash involving a lighter vehicle and a heavier vehicle, the external costs imposed on the lighter vehicle increases with weight of the heavier vehicle and decreases with own weight. Specifically, we find that increasing the weight of the heavier vehicle over the lighter vehicle by 1000 pounds increases the expected number of fatalities in the lighter vehicle by about 39%. Second, increasing the weight of any vehicle increases the probability of it being the heavier vehicle and increases the probability to be the heavier vehicle in a collision and the probability of it being involved in a fatal accident. The external costs translate to about 34 cents/gallon of gasoline.

Finally, we combine these empirical results with a theoretical framework develop simple tests to show that selection bias due to drivers of heavier vehicles involvement in more fatal/serious injury accidents. The paper contributes to two lines of literature. First is the rich set of literature on individual choice of activity on self protection and averting behavior. The second area is the literature on the protective and aggressive effects of vehicle weight on occupant safety (NRC 2002; Peltzman 1975).

In the final manuscript, I explore how sorting due to individual unobserved heterogeneity could bias estimates of MWTP for lower fuel costs. The paper contributes to the existing literature on Energy Paradox by providing empirical evidence that unobserved consumer heterogeneity could result in different consumers sorting into different classes of fuel efficient cars which results in consumers undervaluing fuel economy. Unlike existing literature in this topic, this paper accounts for sorting bias due to unobserved heterogeneity by using a random coefficient logit with error components. The random utility model allows taste to vary across consumers. I treat the discounted operating costs and vehicle costs as random variables. This allows us to obtain a distribution of households' preferences for fuel economy across the population.

Unlike other recent papers that identify consumers' response to fuel economy based on monthly variation on gasoline prices alone, this manuscript supplements information on gasoline prices with information on vehicle demand and elasticity extracted from observed vehicle prices. I use data on household vehicle holdings from the automobile extract of the consumer expenditure survey (CEX) covering the period 2002-2005. This dataset contains information on the vehicle purchases and driving

habits of a random sample of US households. The CEX contains actual data on the number of miles travelled by each household, what make of vehicle they buy, when they buy and what vehicle they trade-in. The individual household data allows me to match vehicles with the household demographic characteristics of those who own them and estimate how different segments of the population value fuel economy differently.

I find significant and substantial evidence of sorting and consumer heterogeneity in the estimates of MWTP. The heterogeneity arises from huge differences in the amount of miles travelled by consumers and heterogeneity in taste and time preferences. We find that a substantial portion of the upper 95% of households in our sample correctly value fuel economy (61%) as they are willing to pay \$0.99 to reduce obtain a \$1.00 discounted future gasoline costs over the lifetime of the vehicle. These results are in contrast to those reported elsewhere by Alcott and Wozny (2009) but support results from the simulation study by Bento et al 2012 who find that not accounting for sorting bias and consumer heterogeneity could lead to a conclusion that consumers undervalue fuel economy. In that light, our results might not necessarily imply that consumers overvalue/undervalue fuel economy, but can be interpreted as being heterogeneous in their valuation of fuel economy.

This dissertation is divided into three main parts. The first part, this section provides a general introduction to the whole body of work. The second section is composed of three different chapters, with each chapter containing a different manuscript. The last section provides a general conclusion and tries to tie all the manuscript together.

MANUSCRIPT 1
**The Mitigating Effect of Strategic Behavior on the Net Benefits of a
Direct Load Control Program**

By

Corey Lang and Edson Okwelum*

Department of Environmental and Natural Resource Economics

University of Rhode Island

Submitted for Publication to the Energy Economics Journal (Revise and

Resubmit)

The Mitigating Effect of Strategic Behavior on the Net Benefits of a Direct Load Control Program

Corey Lang and Edson Okwelum*

Department of Environmental and Natural Resource Economics
University of Rhode Island

October 22, 2014

Abstract

Demand response has become an important tool for utilities to manage load during peak periods. While the effects of demand response programs on peak load reductions are well studied and intuitive, assessments typically fail to recognize the potential for off-peak behavioral responses that may mitigate the total benefits of the program. Using smart meter consumption data on residential air conditioning units enrolled in a direct load control program, this paper examines changes in consumption prior to and after curtailment events. The results suggest substantial increases in off-peak consumption, which reduce monetary benefits of the program by 41%.

Keywords: Demand Response, Precooling, Snapback, Strategic Behavior, Breakpoints

*Email: (Lang) clang@mail.uri.edu, (Okwelum) edson_okwelum@my.uri.edu. We graciously thank Dr. Abigail Anthony of Environmental Northeast for providing us with the energy consumption and weather datasets. We also thank Nate Merrill for his comments on an earlier draft. All remaining errors are our own.

Introduction

Many electricity utilities are challenged to meet demand during peak² consumption, usually during summer afternoons. Further, peak demand has been increasing at an average annual growth rate of 0.8% over the last 10 years (NERC 2013), heightening the need for solutions. Importing capacity from other areas and firing peak generators are two costly solutions; these can increase the cost of generation by \$29.2-80/MWh. (EIA 2013). Alternatively, demand response (DR) programs offer the possibility of no or low cost reductions in peak load, and utilities are increasingly implementing these programs. While DR programs come in several forms, they are generically classed into two groups. Programs with price incentives- Time of Use (TOU), Critical Peak Pricing (CPP), Real Time Pricing (RTP) and Peak Time Pricing (PTR); and technology only programs with no price incentive- direct load control (DLC) programs in which the utility can control usage of appliances for a few hours per day during critical event days.

In general, all these programs seek to reduce peak load³ by changing consumption prices or offering incentives to consumers. Impact evaluations have demonstrated large benefits of DR programs (e.g., Herter 2007, Herter et al. 2007, Faruqui and George 2002, Crew et al. 1995, Matsukawa 2001, Binswanger 2001, Berkout et al. 2000, Gillingham et al. 2013, Thomas et al. 2013), and comparison of load reduction during event hours show that DLC programs are more effective than programs with time-varying price structures (Wolak 2011, Newsham et al 2010). However most of

² Although, CAISO generally defines peak period as being from Hour Ending (HE) 0700 Pacific Prevailing Time (PPT) through HE 2200 PPT, the peak period we refer to in this paper narrowly defined to in line with PG&E's Summer Peak periods which is from 12:00 noon to 8:00 PM

³ Although not part of the program design, some of these programs also reduce energy use during off-peak hours

these assessments tend to only examine benefits in terms of peak reductions, thereby ignoring behavioral responses that could be happening in technology only and potentially overestimating the total benefits. Although behavioral responses in programs with time-varying price incentives have been looked at in terms of load shifting, the closely related concept of precooling and snapback (two types of behavioral responses in technology only) less so (Newsham et al 2010, Jessoe and Rapson 2014, Wolak 2011).

The purpose of this paper is to rigorously estimate both peak and off peak changes in consumption to better understand net benefits of DLC programs with no price incentive or behavioral aspect. We examine Pacific Gas and Electric's (PG&E) SmartAC program, which is designed to reduce peak cooling load by directly controlling air conditioning units of participants.

We use data from a stratified random sample of 294 participating AC units. Because there is no data for a control group, we estimate unit-specific non-linear consumption models and then compare load on event days to predicted load. Importantly, we use data from the summer 2007 training period. In contrast to subsequent summers and normal DR designs, during the training period curtailment days were called for many different temperature levels, not just the hottest days. This aspect means we are not predicting out of sample, and we can have greater confidence in our econometric evidence. Our methodology is validated by a falsification test in which we find no changes in consumption during peak or off-peak times on non-curtailment days that match the temperature and timing profile of actual curtailment days.

Several key results emerge from our analysis. First, we confirm that the SmartAC program, like other DR and DLC programs, reduces peak load during event days. In this case, peak consumption was reduced 19% on average. However, we also find substantial increases in consumption in the hours preceding and the hours following an event by 8% and 7% respectively. Essentially, load is being displaced from peak to off-peak times. These behavioral changes mitigate the monetary benefits of the program by 41%.

The main contribution of this paper is to show DR energy policies can lead to unintended consequences. We contribute to the emerging body of work on the perverse incentives in energy policies (e.g., Fowle 2009, Davis and Kahn 2010, Goulder and Stavins 2011). Particularly relevant to this paper are works by Holladay et al. (2013), and Jessoe and Rapson (2014). Holladay et al. (2013) examine changes in energy production when media outlets relay emergency calls from utilities for electricity reduction during off-peak hours on curtailment days. They estimate that behavioral responses lead to an increase in generation cost of about \$43.70/MWh to \$61.70/MWh in the late morning and early evening on emergency days. We estimate a smaller increase in generation cost, around \$21/MWh, but the results largely corroborate each other. In contrast, to Jessoe and Rapson (2014) who examine the added benefit of load reduction during peak hours on non-event days due to learning, we look at behavioral responses on curtailment days in non-event hours in a technology only program.

Background on the SmartAC Program

PG&E's SmartAC program is a DLC program designed to reduce participating households cooling load during peak load times. The program works by directly controlling air conditioning during peak hours on designated curtailment days. Curtailment events can only occur between the hours of 11 am to 7 pm from May 1 to October 31 and only for a total of six hours per day and for no more than 100 hours per cooling season. Before and after each event, the customer has complete control over their AC unit(s).

Typically, curtailment days are called when system load is expected to be burdensome, usually the hottest days of the year. The California Independent System Operator (ISO) may declare a system emergency when an electric-resource generation facility reaches or exceeds a certain heat rate (usually heat rates of 15,000 British thermal units per kW). However, PG&E can also call events to test the devices and for other discretionary reasons such as transmission or distribution system overload. The Appendix gives a more detailed description of the process of initiating a control event.

In contrast to the normal objective of the program, summer 2007 was a training period, and curtailment days were called on days with a variety of temperatures, not just the hottest days of the year. The inclusion of typical summer temperature days is important because it allows the utility to test the operational performance of the control devices.

The SmartAC Program uses paging signals to reduce cooling load of enrolled air conditioning units during times of peak system demand. The control devices are either programmable communicating thermostats (PCTs) and/or direct load control adaptive switches (switches). Both technologies receive signals through a paging

device but differ by how they lower air conditioning load. Unlike the adaptive switch control devices, the thermostats provide additional functionality to participants.

Demand reduction during curtailment events are achieved by either adjusting thermostat temperature set points or limiting the duty cycles of switches in the units. The PCTs implement load reduction by increasing the cooling set point temperature on the thermostat controls when an event signal is received. Increasing the PCT set point ensures that temperature increases are equitably distributed across the population irrespective of house, occupant temperature preference and air conditioner attributes. The switches on the other hand, reduce air conditioner load by directly controlling how the unit's compressor operates.⁴The SmartAC program was structured such that 70% of the program participants had control switches and 30% smart thermostats. Customers with either device can only opt out of an event by either calling a toll-free number or accessing the program website⁵.

There were two schedules by which load was reduced for those households with PCTs, and PG&E alternated between them. The first schedule set thermostats back one degree every other hour, resulting in a maximum setback of three degrees in each 6 hour curtailment event. The second schedule set thermostats back one degree every hour for the first four hours of the event and then held constant. This last strategy was done in an attempt to evenly spread the load reduction resulting in a gradual temperature decline as against the steep temperature from the first strategy.

⁴ Switches generally control air conditioner load by limiting the compressor run-time or duty cycle to a maximum over a period. When activated on curtailment days, the switches use adaptive algorithm to reduce air conditioner load to a percentage of the load on "learning days". The "learning days" are days with similar characteristics as potential curtailment days and are chosen by the program administrator. The observed duty cycle on these learning days provides an estimate of expected duty cycle on a curtailment day.

⁵ While participants had the option to opt of the program, this option was never exercised by the households (KEMA 2008).

For units with adaptive switches, load reduction during curtailment event was achieved by limiting compressor duty-cycle to a maximum amount in the curtailment period.

Using information from focus groups and similar program evaluation surveys in other locations, participants were recruited into the SmartAC program through direct mail campaigns. Participants into the SmartAC program were first enlisted in the spring of 2007 in San Joaquin County. However this has been expanded to include all service areas covered by the utility. As at the beginning of July 2007 when the first curtailment event was called, the program had enrolled over 2,857. The 297 households sampled for this evaluation, are located in the city of Stockton and it's surrounding areas (San Joaquin County). And as of August 2007, which was the system peak, there were 8,193 participants in the program, which was almost composed exclusively of residential customers with less than 1% commercial customers. The program has subsequently grown in popularity to such an extent that by January 2008, there were 26,000 households enrolled in the program with an additional 22,000 waiting to be supplied control technology. All program participants received a one-off incentive payment of \$25 in return for up to 100 hours/year of load control. In addition, participants in the program evaluation group were given a maximum payment of \$110, based on the number of questionnaires answered in a survey conducted at the completion of the program's first year. Also, participants in the PCT group were given the thermostats free, with a market value of \$200 (KEMA, 2008). The goal of the program is to enroll a sufficient number of participants to achieve load reduction of 305 MW with over 95% coming from residential accounts.

Methodology

We develop a methodology that estimates air conditioner load usage as a function of weather control (temperature and humidity) and time of day, parameterizing each unit separately. Then we use these predictions to compare actual usage on curtailment days to expected usage. While there are several methods for estimating baseline consumption patterns and load comparisons, we use a regression-based baseline modeling approach. The wide scale adoption of smart grid meters and availability of high-resolution, hourly or 15-minute energy consumption data has contributed a great deal to improvements in regression-based baseline models (Armel et al. 2013; Newsham et al. 2011; Santin and Itard 2010; Santin et al. 2009). Mathieu et al. (2011) find that the regression-based baseline model performs better than most models used in evaluating demand response programs. Given that the regressions coefficients are from a least squares estimator, Mittelhammer et al. (2000) show that coefficients from these models should be unbiased and consistent as long as we have controlled for all the relevant predictor variables (Newsham et al 2011). In addition, the regression based model allows for customer-specific response estimates to be used separately for decision making or easily averaged to any degree of aggregation.

We model each air conditioner's baseline electric load as a function of time of the day, relative humidity, cooling degree days, heat index and a 631 index. By estimating parameters for each unit, we better capture unique occupant cooling preferences and differences in physical properties of residences. We use 15-minute-interval fixed effects, which captures the pattern of usage of each unit. Cooling degree days enters into the model as a 2-breakpoint demand function.⁶ The first breakpoint is the unit's

⁶ Figure 1 depicts the intuition of the 2-breakpoint model.

base or reference temperature, which marks the temperature above which air conditioning is used. The second breakpoint represents the temperatures point above which the air conditioner runs at 100% duty cycle. This specification allows for three different regimes of sensitivity to temperature changes. The heat index is a function of relative humidity and temperature and gives an idea how hot it given the relative humidity and actual air temperature. We use an equation that reproduces NOAA's heat index chart to estimate the heat index for all participating units. The 631 index is a weighted average of temperature and is estimated as 60% of a given day, 30% of prior day and 10% of two days prior average temperature readings (CAISO 2013)⁷. Specifically, we estimate the baseline model using only non-curtailement days for each unit as follows:

$$L_{ihtd} = \alpha_{it} + \gamma_i B_{ih} + \delta_i C_{ih} + \eta_i R_{ih} + \theta_i X_{ih} + \zeta_i Y_i + \varepsilon_{ihtd} \quad (1)$$

where L_{ihtd} is the observed load in kWh of unit i in time t of day d , α_{it} is unit-15 minute fixed effect, and B_{ih} and C_{ih} are the degrees Fahrenheit above the first and second breakpoints, respectively. R_{ih} is unit i 's relative humidity at hour h , X_{ih} is heat index for unit i at hour h and Y_i is the 631 temperature index for unit i . The breakpoints are defined as follows:

$$\left. \begin{aligned} B_{ih} &= \max(T_h - b_i, 0) \\ C_{ih} &= \max(T_h - c_i, 0) \end{aligned} \right\} \quad (2)$$

where T_h is the hourly average temperature of day d , and b_i and c_i are the unit specific breakpoints. These equations are similar in structure to the calculation of cooling

⁷ http://www.caiso.com/Documents/2013SummerLoads_ResourcesAssessment.pdf

degree days, except that cooling degree days is calculated with a set breakpoint, usually 65°F, whereas here we allow the breakpoints to take a range of values. α , γ , β , δ , η , ζ , and θ are parameters to be estimated.

For each unit, we implement a grid search to find the breakpoint values that fit the consumption patterns best. We estimate the model for values of b_i in the range of 65°F to 84°F and for values of c_i in the range of 78°F to 90°F, both in increments of 1° and with the constraint that $b_i < c_i$. The optimal combination of b_i and c_i is chosen based on which model maximizes R-squared.⁸

One drawback to using the regression-based approach is that out-of-sample predictions could perform poorly due to the parameterization. Regression-based baseline models assume that the functional relationship between load and weather (cooling degree-days) is the same on curtailment days as on the non-curtailment days (KEMA 2008). There are two reasons to be optimistic that this is not a large problem in our case. First, our empirical approach allows for non-linearity in the load-weather relationship, which should enable the upper tail of the distribution to be modeled better. Second, as explained in more detail in the next section, curtailment days were not temperature outliers, and so our out-of-sample weather has strong statistical overlap with the in-sample days.

From the parameters of the baseline models determined with non-curtailment days, we then predict demand on curtailment days. The difference between actual consumption and predicted consumption is the key statistic for determining the effect of the program for both hours during the event and hours before and after the event.

⁸ While R-squared is typically not thought to be a useful judge of predictive power, in this case we are only comparing models with the same number of explanatory variables. So R-squared would indicate the same breakpoints as optimal as would AIC or BIC.

Data

This section details the multiple sources of data used for the analysis. The main dataset is the 15 minute interval consumption data covering the period June 12 through October 31, 2007. We additionally use weather data to model consumption and data on the marginal cost of generating 1MWh in the wholesale market to monetize the benefits of changes in consumption on curtailment days.

Electric Load Data

As part of the SmartAC program load impact assessment, air conditioner load loggers were installed on a sample of 352 participating units. These units were contained in a stratified random sample of 297 (146 PCTs and 151 switches) program participants from a population 2,917.⁹ The sample was stratified by type of device (PCT or switch), total number of cooling tons from all units (less than four tons, four tons or more), and multiple units (yes or no). The sample design, and stratification as well as design sample sizes, and installed sample sizes are presented in Table A1 in the Appendix. We removed one unit for having an anomalously high load of 253 kW (over 10 times larger than the mean load) and four units for having 0 kW consumption throughout the study period. 53 units were dropped because of missing data, which left us with a final sample of 254 participants and 294 loggers.

Figure A5 in the online appendix is a density plot of load with the dashed vertical lines representing the quartile positions. About fifty percent of the participating units' rated KW loads were less than 20KW with the lowest unit having a rated load of 2.8KW and the maximum unit with rated KW load of 35KW.

Electricity consumption for the participating units were recorded by the data loggers in 15-minute ampere readings, which recorded the average instantaneous amp

⁹ Some participants had more than one unit.

reading for each minute during that period. The amperage was converted to kW using the voltage levels measured at the units. The average instantaneous readings were then converted to kWh by summing each participating unit's KW consumption over the four 15 minute interval readings within that hour. We discarded data on weekends and holidays, as curtailment days are only called on normal weekdays. In total, we had over 3.9 million observations at the 15-minute-unit level. The mean KW consumption was 0.26KW with a standard deviation of 0.85, while the maximum KW value was 7.66.

Temperature Data

Daily average temperatures were obtained from the NOAA data web site.¹⁰ While we do not know the location of each sample unit, PG&E mapped each unit to the nearest weather station and we use this to match each sample unit with weather data.

Figure2 is a plot of the density of the average daily temperature data used in the analysis with vertical lines representing the average daily temperature at which curtailment events were called. Due to the experimental nature of this program in 2007, many curtailment days were called for average heat days when peak load reductions were unlikely needed. Calling curtailment events on days with mild temperature is important because it ensures that our out-of-sample predictions will have strong statistical overlap with the in-sample days. In addition, it allows program administrators to obtain data on the performance of the technologies over the whole period of the cooling season and get a good performance of the technologies on learning days.

¹⁰<http://www.ncdc.noaa.gov/>

Marginal Cost of Generating Electricity

In order to monetize the net savings from the SmartAC program, we obtained data on the marginal cost of generating 1MWh of electricity from California ISO website.¹¹ We use the locational marginal pricing (LMP) at Stockton, CA. We believe that the conditions existing in 2007 in the California ISO (CAISO) means that the price data we use are indicative of both the relative scarcity and abundance of the product and the sensitivity of the demand to variations in price. The CAISO market uses balancing energy market in allocating energy to match load less than 10 minutes ahead forecasting scenarios. This market has two important features, namely an hourly, paid-as-bid market that is conducted at the intertie points with neighboring control areas, and a five-minute, single price auction market for internal generators. Prices for the balancing energy market are settled at a market determined price that is set by the most expensive energy dispatched by the system in any given interval using a uniform price auction. The LMP reflects the energy value at a specific location at a specific time, fluctuating with time of day, and can surge significantly with extreme temperature and other unpredictable events. The LMP data we use is hourly. During peak load hours on curtailment days, the mean price is \$100/MWh, with a maximum of \$230/MWh and a minimum of \$60/MWh. During off-peak times, the mean LMP value is \$58/MWh.

Results

DR programs have historically played significant roles in lowering the costs of meeting peak demand reduction and reducing environmental impacts. DR resources

¹¹<http://www.ferc.gov/market-oversight/mkt-electric/california/caiso-archives.asp>

offset the need for building additional peaking generation units and supply some low carbon flexible capacity utilities need to ensure real-time balance and reliability of the system. Given that the overall demand is an important determinant of use of fossil fuels, DR resources support the integration of increasing levels of nonrenewable resources.

Baseline model

While our primary interest is changes in consumption patterns induced by curtailment days, we first examine the baseline model estimation for indications that it captures demand well. In terms of model fit, R-squared averages 0.76 with a 90% confidence interval of [0.53, 0.97], which is an improvement over papers using temperature regressions or single breakpoint models (Herter 2007, Reed *et al.* 2009, Newsham *et al.* 2011). In addition, the average participants R-squared show that the model is able to account for over 94% of the variation in cooling load.

The breakpoints resulting from the optimization showed substantial heterogeneity.¹² Figures 3a and 3b are density plots of the estimated first and second breakpoints, respectively, for participating AC units. The mean value of the first breakpoint is 71°F with a 90% confidence interval of [66, 82], which seems like reasonable temperatures for AC units to be switched on. The distribution of the second breakpoint appears bimodal, with peaks at 80°F and 88°F. The two figures illustrate how different units have different breakpoints because of the differences in the buildings mass, outside temperature and user preferences. This indicates the importance of modeling this heterogeneity, which may be lost with a more traditional panel estimator.

¹² For some of the units (less than 3%), the second breakpoint was not used in estimating the baseline model. However, all the units had the first breakpoint in the baseline model.

In addition, the coefficient estimates on the breakpoint parameters were highly statistically significant. 88.4% were significant at the 0.1% level, and the remaining were significant at either 5% or 10% level. In contrast, when the models are estimated with uniform base temperature values of 65°F and 72°F for the first and second breakpoints, respectively, slightly less than 65% of the coefficient estimates were significant at the 10% level. Thus, our grid search method and 15-minute fixed effects substantially improves model fit.

Changes in consumption on curtailment days

In Figure 4, we plot a single unit's load profile for a typical curtailment day and non curtailment on the same axes. The dark, solid lines represent actual load profile while the red, dashed lines are for predicted load on a typical curtailment day. The Figure shows a trough in between two peaks. The two peaks represent precooling and snapback loads, while the trough represents load reduction during peak period. Next, we quantify these figures. Table 1 presents the results of our analysis of estimated change in consumption before, during and after a curtailment event on a curtailment day. Results are presented for each of the 13 curtailment days as well as an overall mean. The values in parenthesis represent standard errors computed by dividing the standard deviation obtained from a t-test of the difference in means between predicted and observed load with the square root of the number of observations. All values are significant at the 1% level except for the amount of precooling on July 17, which is negative and not significant. Therefore, the value for July 17 might just be random noise.

The first two columns give the date and average daily temperature of each curtailment day. During the curtailment event, the results indicate that the average sample AC unit reduced load between 0.020 kWh/h to 0.754KWh/h, with an average reduction of 0.410 kWh across all curtailment days. (We normalize kWh reductions by per hour because the length of curtailment varies between four and six hours). These estimates are consistent with prior studies that have estimated the load impact of demand response programs and found values ranging from 0KWh/h to 0.75KWh/h (KEMA 2007, 2008, Herter et al. 2007).

Table 1 also displays average changes in consumption for the three hours before and three hours after an event. Our results suggest that program participants do increase AC usage both before and after an event.¹³Precooling occurred in 11 of the 13 curtailment days. Precooling over the period ranged from 0.040 kWh/h to 0.320 kWh/h with an average value of 0.127 kWh/h. The amount of snapback ranges from 0.009 kWh/h to 0.240 kWh/h with mean of 0.115KWh/h. On a per hour basis, pre-cooling equates to 31% of the reduction during the curtailment period, and snapback equates to 28%.

Taken together, these figures imply that participating residents are consuming between 25-35% more in the hours immediately before and after a curtailment event than they would ordinarily have consumed. These results quantify behavior taken by households to maintain comfort during curtailment days.

¹³ An alternative version of the model was estimated in which hours of the day fixed effect and curtailment day variable were interacted. The positive sign on the interaction term pre-curtailment hours and curtailment day variable indicates precooling while a positive sign between post-curtailment hours with curtailment day variables indicate significant snapback event respectively. A test of joint significance applied to the results from the curtailment period and the surrounding hours indicate that the SmartAC impacts are statistically significant and in the expected direction across the curtailment period, pre-curtailment hours, and post-curtailment hours respectively.

Learning over time

Table 1 indicates that households behave strategically to pre-cool their homes before curtailment. We wanted to further investigate if the amount of pre-cooling increased over time, suggesting that customers were learning how to best pre-cool their homes to cope with curtailment. If there is a pattern of learning, then the later days in the summer will be more indicative of the amount of pre-cooling to expect as this program and others like it mature.

First, a comparison of the amount of precooling on the curtailment days that occur early in the summer reveals that that pre-cooling is increasing over time. Second, customers precool their homes more in the late, hotter summer days than curtailment days with lower temperature. This last fact implies that there exists a positive correlation between the amount of precooling and temperature. This make sense because one would more pre-cooling will be needed to maintain comfort levels. This last fact presents challenges in trying to identify the evidence of learning from temperature effects.

To help us decipher whether the larger precooling amount in later, hotter summer days are due to learning by customers, and not a reflection of those days having higher temperatures, we undertake additional analysis. First, we look at amount of precooling in late hot summers days that are non-curtailment but occurring after two consecutive curtailment days. And then compare this to the amount of precooling occurring in non-curtailment days following consecutive curtailment days in July. Second, we also look at the load shape a day before curtailment and a day after curtailment. However, given the limited number of curtailment days in our sample, these two approaches could not

conclusively reveal additional information about whether or not households learn to precool their homes over the course of the program¹⁴.

The direction and size of precooling and rebound we find are in line with figures reported elsewhere from experiments about using building thermal mass to reduce peak cooling load (Xu 2006, Xu et al 2004¹⁵). However, unlike Jessoe and Rapson (2014) who found that households learn to conserve energy beyond the conservation periods during peak hours under demand response programs with price incentives and behavioral intervention; our results show that whatever learning (if it does occur) that occurs under a direct load control program without price incentives is in the opposite direction: households learn to erase gains made during peak periods through precooling of their homes given comfort constraints. The two conflicting results provide an insight into how consumers' behavior responds to non-monetary intervention such as technology and a monetary intervention.

Heterogeneity by technology and load

We are also interested in how curtailment affects units with different installed technologies (PCT vs. switch) and units with different loads. Table 2 shows the average changes in consumption averaged across all curtailment days splitting the sample various ways. Panel A splits the sample by choice of technology. The results

¹⁴ In addition, we examined whether conditional on temperature, the curtailment days that occur on dates later in the summer have positive coefficients which could indicate some evidence of learning. We estimated a regression similar to equation 1, but using all days (both curtailment and non-curtailment days) and including variables to account for curtailment event and number of prior curtailment days in the summer experienced by households. The results of such regressions are presented in the online appendix in Table 7. We see a positive coefficient for the interaction terms of hours prior to curtailment period and the number of consecutive prior curtailment events experienced. This is just an additional indication that learning could be occurring and should be explored additionally if one has additional data.

¹⁵ Xu 2006, precooling could shift about 20-50% of peak load to non-peak load using building thermal mass.

show the same general pattern of decreases in consumption during the event and increases in consumption before and after the event for both groups. However, the changes for the PCT group are larger in every stage – more reductions during the event and more pre-cooling and snapback. One possible explanation for the heterogeneous change in consumption between the PCT units relative to those with adaptive switches could be related to how the two technologies effect load reduction during curtailment days. Another reason could be that the larger changes in consumption by thermostat compared to the adaptive switches is a reflection of the higher variation in thermostat load compared to switches. The algorithms for the two technologies are such that load reduction by units with adaptive switches is predictable and more evenly distributed over the curtailment period compared to units with PCTs (Agnew et al. 2008).

Also, within PCT group, the two different strategies for achieving load reduction results in different outcomes during the curtailment events, but do not seem to affect the amount of precooling and snapback. On the average, PCT units exposed to the steep temperature declines reduce about 13% more load given that all other conditions are controlled for.

Panel B partitions units into quartiles based on load size. Similar to Panel A, we see the same pattern across all four quartiles. However, another pattern emerges that as load grows so does the magnitude of consumption change. This finding is consistent with intuition because units with a larger load are likely to be used more and thus have larger decreases during curtailment. Further, larger units are likely to pre-cool more if occupants want to maintain the same comfort level during the curtailment event.

Falsification test

In order to alleviate concerns that the estimated changes in consumption during event days were a spurious result, we conducted a falsification test with proxy days treated as curtailment days. To implement this, we sought a match for each curtailment day that had the same average temperature, was within two days of the curtailment day, and was not a curtailment day itself. Matches meeting these criteria were found for eight of 13 curtailment days. Our definition of proxy days uses only non-holiday weekdays, and accounts for daily maximum and minimum temperature. We also attempted to define proxy days using humidity; however it was difficult matching on several covariates. We then re-estimated our baseline model excluding both the true curtailment days and the proxy days, and then calculated the difference between actual and expected consumption for the proxy days. The results, shown in Table 3, suggest no evidence of systematic increases or decreases in consumption during any of the three periods (before, during, and after the event). Magnitudes in all three columns tend to be smaller than the numbers in Table 1 and flip sign from day to day.

In addition, we relaxed the criteria for selecting the proxy days and used all non-curtailment days as proxy. Doing this allows us to relax the temperature control and possibly account for humidity. We performed a Monte Carlo exercise, first excluding each non-curtailment day and then estimate the model and predicting demand for that proxy day. We did this for each non-curtailment day in the sample (including weekends and public holidays) and plotted the average observed load and predicted load. The results are presented in Figure 7. These results validate our methodology and results.

Calculating peak load savings and net savings

We now examine the complete effect of the SmartAC program and the mitigating effect that pre-cooling and snapback have on total benefits. Table 4 reports the results. Columns A and B present the peak load reduction benefits of the program in terms of kWh savings and monetary savings per unit. Columns D and E present net benefits of the program, additionally accounting for pre-cooling and snapback.¹⁶

As expected and as foreshadowed in earlier results, on average, units reduced consumption by a total of 23.218 kWh during the summer's peak events. Monetizing these reductions indicates that PG&E saved each participant nearly \$41. However, after accounting for precooling and snapback, kWh savings are reduced by 41% and monetary savings are reduced by 42%.¹⁷ These results make clear the importance of accounting for behavioral responses to DLC programs when determining program benefits.

Column G presents the cost of conservation, which is calculated as the marginal cost of generating the extra electricity needed to meet the increased off-peak demand due to conservation during peak times. The values indicate that reductions during peak times cost the utilities an additional \$21.07/MWh on average, with values ranging from \$5.40/MWh to \$43.41/MWh. Of course, this is cheaper than the cost of generation during peak times, so the net benefit is still positive. But it is important to know that the cost is not zero. These estimates are similar though smaller to those of

¹⁶The values in column A were obtained by averaging hourly savings during the curtailment period for all units. Those in column B were obtained by first multiplying the hourly cost of generating 1MWh with the change in consumption at that hour, and thereafter computing the average value. In the next column, C, we add the precooling and snapback values to the change in consumption during the curtailment period. Finally, the values in column D were obtained by summing the product the change in consumption in each hour with the marginal cost of generating 1MWh during each hour.

¹⁷The discrepancy between percentage reductions is due to the higher generation costs during peak times.

Holladay et al. (2013), who estimate conservation costs ranging from \$43.70/MWh to \$61.70/MWh.

Despite the large reductions in net benefits from pre-cooling and snapback, the program still offers substantial benefits. Projecting the net savings onto the entire 8,193 customers enrolled in the program leads to an estimated \$194,000 in utility bill savings. Further, the hotter days in the latter part of the summer yield the largest monetary savings, due both to larger kWh reductions and higher marginal prices. This result suggests that per-unit savings are likely to be larger for summers in which curtailment occurs mostly on very hot days.

Estimates of Environmental Impacts

We use the load reduction estimates from the program to estimate the overall environmental impacts of the programs. We make simplifying assumptions that increased consumption increases in the period before and after curtailment occurs when natural gas is at the margin and the consumption reductions during the curtailment period occur when coal is the margin. While these assumptions may be very simplistic, they ensure that whatever environmental impact we attribute to the program are lower bounds. In addition, it does reflect what is obtainable since natural gas is the non-base load source in the Western Interconnection area. We use the hourly marginal emissions intensities from Graff-Zivin, Kotchen and Mansur (2012) to multiply the changes in consumption to estimate the environmental impacts.

The result from the analysis is presented in columns C and F of Table 4. In Column C, we report the values of reduced emission due to load curtailment attributable to the program. Column C reports estimates of emissions reduction due to avoided or displaced generation from coal of about 0.78 *metric tons of CO₂e per unit*

during the curtailment period. Finally, in Column F, we report net environmental effects by aggregating across the periods: precooling, curtailment and snapback periods. The program results in a net reduction of pollution even after accounting for increased emissions due to behavioral response by households of *0.42 metric tons of CO_{2e} per unit*. Given the high intra-day volatility in the marginal fuel, the emissions reductions we attribute to the program are lower than would be expected.

Policy Discussions

Faruqui and Sergici (2009) and Newsham et al (2010) provide reviews of several studies on evaluation of DR programs and how these results vary. These papers provide reviews of programs using technology-based standards compared to programs with price incentives and information feedback. A third type of program considers the customer behavior other than pricing incentives like goal setting, commitment, social approval, consumption feedback, etc (Allcott and Rogers 2014, Jessoe and Rapson 2014).

Program with Price Incentives

Programs with price incentives achieve peak load reduction by encouraging either the elimination of on-peak energy consumption activities or shifting of such activities to other periods through economic incentives. The economic incentives are usually in the form differential pricing in residential electricity rate structure based on time of delivery of electricity. That is, utilities encourage behavioral modifications by charging higher prices during periods of high system-wide demand and lower prices in other periods. Alternatively, utilities could compensate households for every kWh reduced during periods of high system-wide demand. These programs come in various

forms. While programs with price incentives have been effective at reducing peak load, their relative effectiveness depends largely on how the price incentives are structured. Generally, CPP programs seem to be very effective at reducing peak loads, whereas TOU less so (Newsham et al 2010). Two main reasons have been proffered to account for the higher on-peak load reduction of CPP programs compared to TOU programs: price and frequency of occurrence. The intuition is that in the absence of enabling technologies, households are more apt to respond to a CPP program given the small number of events that are called over each summer. However, under the TOU pricing, households have to change their behavior daily, which could be cumbersome. In addition, CPP programs have been shown to reduce energy during peak periods on non-event days, implying that behaviors formed on event days can be transferred to non-event days (Newsham et al 2010, Faruqui and Sergici 2008).

PTR programs are less effective at reducing peak load than CPP programs even though they are similar (e.g., Newsham et al 2010). However, unlike CPP programs, PTR programs are designed to provide rebates on bills for reduction in peak load. This result is in line with evidence from Psychology which found that people respond more to “sticks” than “carrots” (Kahneman and Tversky 1984). This also suggests that behavior-based programs might make price-based policies more effective by making people more aware of potential savings

Behavior Based Programs

A basic tenet of behavior change is providing consumers with regular feedback on their energy consumption to take action and achieve greater levels of energy savings over and above what traditional price based incentive programs can accomplish. Behavior based programs do not use monetary benefits to encourage load

reduction. The literature on behavior change is rich and most are rooted in social sciences (Jesoe and Rapson 2014, Allcott and Rogers 2014, Allcott and Mullainathan 2010).

Allcott and Rogers use large data sets from a series of programs by OPOWER to study how providing households' feedback information on their energy consumptions induces conservation behavior. They find that such behavioral programs result in about 1.1 to 2.8% reduction in energy consumption relative to baselines. Jesoe and Rapson 2014 use a field experiment to study the additional impact of high-frequency information about households' residential electricity usage on the price elasticity of demand in addition to household's response to price incentives. They find that households that were not provided information reduce peak demand by about 2-6% in response to price incentives alone. Additionally, they find that households that were provided with the high frequency information on price are more likely to reduce peak load by 8-22% over and above the amount due to price incentives alone.

Technology Only Programs

Technology only programs are designed to reduce peak load through the use of equipment to modify the operation of appliances during peak period (Newsham 2010, CAISO 2013, Faruqui and Wood 2008). DLC programs are designed to reduce peak use on a small number of event days without time-varying price incentives or behavior interventions. From a review of literature on program evaluation on DR programs, DLC based programs are the most efficient (Newsham et al 2010, Faruqui and Wood 2008). A review of peer reviewed literature shows that DLC programs reduce approximately 0.25KWh/h to 1.4KWh/h per household during curtailment periods.

Our findings indicate technology only programs result in larger load reductions during the curtailment periods. This is because technology only programs provide fewer margins for adjustment (Newsham et al 2010). This result is expected from an automated system because technology provides opportunity for participating households to choose their preferred mode of response once during the program lifetime and not be bothered to remember when curtailment events are due. However, if we account for precooling and rebound, the net savings are comparable to load reduction from other programs with price incentives.

Although, we find that the result of customers' behavioral response and the interaction of technology under a non-monetary intervention results in the diminution of conservation gains, the fact that customers may learn to precool their homes at higher temperatures could also provide policy benefits. First, the fact that households precool their homes under a DLC only program is an indication that residential households may be averse to relinquishing complete control of their air conditioners to utilities during summer peak times. Perhaps giving consumers back some type of control during curtailment hours could be one way of improving the performance as it could provide additional avenues for adjustment. Second, the evidence of precooling suggests that utilities should bundle some sort of time-varying price incentive into DLC program to make them more efficient.

Taken together, our results suggests a hybrid policy program that incorporate some form of automation and time-varying price structure with behavioral aspects will perform better and reduce the most amount of peak load while still maintaining consumer comfort.

Conclusion

Deregulation has greatly impacted the ability for the electricity industry in the US to make profit. Specifically, additional generation capacity is no longer guaranteed a rate of return. Because of the costliness of new capacity and the growing pressure peak demand places on the system, there has been increasing implementation of demand response programs as a way to curb peak load. DR can also be applied for energy conservation in general. Recent literature has quantified the benefits of DR by measuring the amount of load reduced during peak periods, and the impacts are substantial. However, little work has been done to quantify the impact of behavioral responses to such programs. This paper seeks to contribute to that gap in the literature by examining the strategic response of participants in one such DLC demand response program and how that response may undermine its effectiveness.

Our results suggest that program participants increase consumption before and after event hours relative to the counterfactual. These results are intuitive, but are also of critical importance as DR policies move forward and the need to understand net benefits and forecast demand grows. For example, the SmartAC program, now in its seventh year, has over 100,000 enrolled customers. To be clear, the results still support DR and DLC programs as an effective way to curb peak demand; but, the true conservation cost is not zero.

REFERENCES

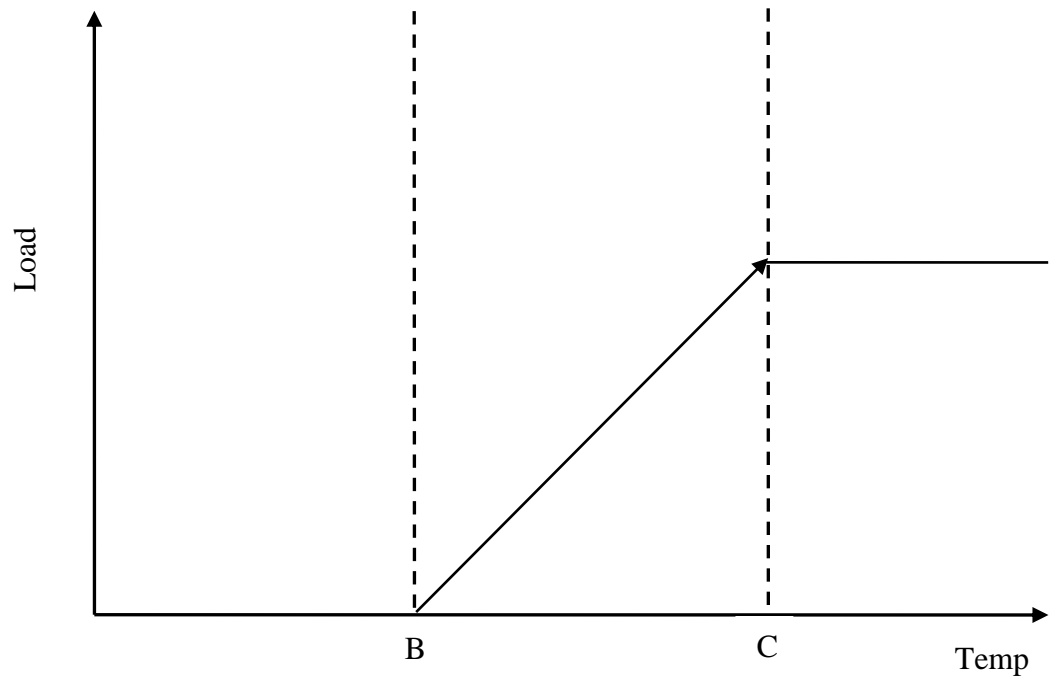
- Anthony, W Abigail, "Climate change and peak demand for electricity: Evaluating policies for reducing peak demand under different climate change scenarios", Unpublished PhD Thesis. 2009
- Armel, K Carrie., Abhay Gupta, Gireesh Shrimali, and Adrian Albert. "Is Disaggregation the Holy Grail of Energy Efficiency? The Case of Electricity." *Energy Policy* (2012).
<http://www.sciencedirect.com/science/article/pii/S0301421512007446>.
- Azevedo, Ines L., Marco Sonnberger, Brinda A. Thomas, Granger Morgan, and Ortwin, Renn. "Developing Robust Energy Efficiency Policies while accounting for consumer behavior," International Risk Governance Council (IRGC) report.2012.
- Berkhout, Peter H.G., Jos C. Muskens, and Jan W. Velthuisen. "Defining the Rebound Effect." *Energy Policy* 28, no. 6–7 (June 2000): 425–432. doi: 10.1016/S0301-4215(00)00022-7.
- Binswanger, Mathias. "Technological Progress and Sustainable Development: What about the Rebound Effect?" *Ecological Economics* 36, no. 1 (January 2001): 119–132. doi: 10.1016/S0921-8009(00)00214-7.
- Birt, Benjamin J., Guy R. Newsham, Ian Beausoleil-Morrison, Marianne M. Armstrong, Neil Saldanha, and Ian H. Rowlands. "Disaggregating Categories of Electrical Energy End-Use from Whole-House Hourly Data." *Energy and Buildings* 50 (2012): 93–102.
- Borenstein, Severin, and Lucas W. Davis. "The Equity and Efficiency of Two-Part Tariffs in U.S. Natural Gas Markets." *Journal of Law and Economics* 55, no. 1 (2012): 75 – 128.
- Borenstein, Severin, and Stephen Holland. "On the Efficiency of Competitive Electricity Markets with Time-Invariant Retail Prices." *RAND Journal of Economics* 36, no. 3 (2005): 469–493.
- Borenstein, Severin, Michael Jaske, and Arthur Rosenfeld. "Dynamic Pricing, Advanced Metering, and Demand Response in Electricity Markets." *Center for the Study of Energy Markets* (October 31, 2002).
<http://escholarship.org/uc/item/11w8d6m4>.
- Brookes, L. G. "Energy Efficiency Fallacies: The Debate Concluded." *Energy Policy* 21, no. 4 (1993): 346–347.
- Brookes, Len. "The Greenhouse Effect: The Fallacies in the Energy Efficiency Solution." *Energy Policy* 18, no. 2 (March 1990): 199–201. doi:10.1016/0301-4215(90)90145-T.
- Callaway, Duncan S. "Tapping the Energy Storage Potential in Electric Loads to Deliver Load Following and Regulation, with Application to Wind Energy." *Energy Conversion and Management* 50, no. 5 (May 2009): 1389–1400. doi:10.1016/j.enconman.2008.12.012.
- Costa, Dora L., and Matthew E. Kahn. "Energy Conservation 'nudges' and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment." *Journal of the European Economic Association* 11, no. 3 (2013): 680–702. doi:10.1111/jeea.12011.

- Coughlin, Katie, Mary Ann Piette, Charles Goldman, and Sila Kiliccote. "Statistical Analysis of Baseline Load Models for Non-Residential Buildings." *Energy and Buildings* 41, no. 4 (April 2009): 374–381. doi:10.1016/j.enbuild.2008.11.002.
- Crew, Michael A., Chitru S. Fernando, and Paul R. Kleindorfer. "The Theory of Peak-Load Pricing: A Survey." *Journal of Regulatory Economics* 8, no. 3 (1995): 215–48.
- Davis, L. W., & Kahn, M. E. (2010). International trade in used vehicles: The environmental consequences of NAFTA. *American Economic Journal: Economic Policy*, 2 (4), 58{82.
- Della Vigna, Stefano. *Psychology and Economics: Evidence from the Field*. National Bureau of Economic Research, 2007. <http://www.nber.org/papers/w13420>.
- Faruqui, Ahmad, Hung-po Chao, Vic Niemeyer, Jeremy Platt, and Karl Stahlkopf. "Analyzing California's Power Crisis." *The Energy Journal* 22, no. 4 (January 1, 2001): 29–52.
- Faruqui, Ahmad, Hung-Po Chao, Victor Niemeyer, and Jeremy Platt. "Getting out of the Dark: Market-Based Pricing Could Prevent Future Crises." *Regulation* 24 (2001): 58.
- Faruqui, Ahmad, and Stephen George. "Quantifying Customer Response to Dynamic Pricing." *The Electricity Journal* 18, no. 4 (May 2005): 53–63. doi:10.1016/j.tej.2005.04.005.
- Faruqui, Ahmad, and Stephen S. George. "The Value of Dynamic Pricing in Mass Markets." *The Electricity Journal* 15, no. 6 (2002): 45–55.
- Fowle, M. L. (2009). Incomplete environmental regulation, imperfect competition, and emissions leakage. *American Economic Journal: Economic Policy*, 1 (2), 72{112.
- Gillingham, Kenneth, Matthew J. Kotchen, David S. Rapson, and Gernot Wagner. "Energy Policy: The Rebound Effect Is Overplayed." *Nature* 493, no. 7433 (2013): 475–476.
- Gillingham, Kenneth, Richard G. Newell, and Karen Palmer. *Energy Efficiency Economics and Policy*. National Bureau of Economic Research, 2009..
- Goulder, L. H., and Stavins, R. N. (2011). Interactions between state and federal climate change policies. *In The Design and Implementation of US Climate Policy*, NBER Chapters,(pp. 109{121). National Bureau of Economic Research, Inc.
- Greening, Lorna A, David L. Greene, and Carmen Difiglio. "Energy Efficiency and Consumption—the Rebound Effect—a Survey." *Energy Policy* 28, no. 6 (2000): 389–401.
- Guerra Santin, Olivia, Laure Itard, and Henk Visscher. "The Effect of Occupancy and Building Characteristics on Energy Use for Space and Water Heating in Dutch Residential Stock." *Energy and Buildings* 41, no. 11 (2009): 1223–1232.
- Heffner, Grayson. "Demand Response Valuation Frameworks Paper" (2010). <http://escholarship.org/uc/item/401781d4.pdf>.
- Herter, Karen. "Residential Implementation of Critical-Peak Pricing of Electricity." *Energy Policy* 35, no. 4 (2007): 2121–2130.

- Herter, Karen, Patrick McAuliffe, and Arthur Rosenfeld. "An Exploratory Analysis of California Residential Customer Response to Critical Peak Pricing of Electricity." *Energy* 32, no. 1 (2007): 25–34.
- Herter, Karen, and Seth Wayland. "Residential Response to Critical-Peak Pricing of Electricity: California Evidence." *Energy* 35, no. 4 (2010): 1561–1567.
- Holladay, J. Scott, Michael Price, and Marianne Wanamaker. "The Perverse Impact of Calling for Energy Conservation" (2012). http://www.ncsu.edu/cenrep/workshops/documents/Holladay_Scott.pdf.
- Khazzoom, J. Daniel. "Economic Implications of Mandated Efficiency in Standards for Household Appliances." *The Energy Journal* 1, no. 4 (1980): 21–40.
- Kiliccote, Sila. "Open Automated Demand Response Communications in Demand Response for Wholesale Ancillary Services" (2010). <http://escholarship.org/uc/item/1f4923j2.pdf>.
- Kiliccote, Sila, Mary Ann Piette, Greg Wikler, Joe Prijyanonda, and Albert Chiu. *Installation and Commissioning Automated Demand Response Systems*. Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA (US), 2008. http://www.osti.gov/energycitations/product.biblio.jsp?osti_id=927888.
- Matsukawa, Isamu. "Household Response to Optional Peak-Load Pricing of Electricity." *Journal of Regulatory Economics* 20, no. 3 (2001): 249–267.
- Moser, Susanne C., and Amy Lynd Luers. "Managing Climate Risks in California: The Need to Engage Resource Managers for Successful Adaptation to Change." *Climatic Change* 87, no. 1 (2008): 309–322.
- Muggeo, Vito M. R. "Estimating Regression Models with Unknown Break-Points." *Statistics in Medicine* 22, no. 19 (2003): 3055–3071. doi:10.1002/sim.1545.
- Newsham, Guy R., and Brent G. Bowker. "The Effect of Utility Time-Varying Pricing and Load Control Strategies on Residential Summer Peak Electricity Use: A Review." *Energy Policy* 38, no. 7 (2010): 3289–3296.
- Pedersen, Linda, Jacob Stang, and Rolf Ulseth. "Load Prediction Method for Heat and Electricity Demand in Buildings for the Purpose of Planning for Mixed Energy Distribution Systems." *Energy and Buildings* 40, no. 7 (2008): 1124–1134.
- Sorrell, Steve. "Jevons' Paradox Revisited: The Evidence for Backfire from Improved Energy Efficiency." *Energy Policy* 37, no. 4 (2009): 1456–1469.
- Thomas, Brinda A., and Inês L. Azevedo. "Estimating Direct and Indirect Rebound Effects for US Households with Input–output Analysis Part 1: Theoretical Framework." *Ecological Economics* 86 (2013): 199–210.
- Wellinghoff, H., D. Morenoff, James Pederson, and M. Tighe. "Creating Regulatory Structures for Robust Demand Response Participation in Organized Wholesale Electric Markets." *Proceedings of ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA* (2008). http://www.aceee.org/files/proceedings/2008/data/papers/5_83.pdf.

Figures and Tables

Figure 1: Density of Load Distribution with Quartiles



Notes: Point B on the horizontal axis represents the first breakpoint, the point at which cooling starts. Point C, represents the point at which air conditioner runs at 100% duty circle.

Figure 2: Density of Average Daily Temperature Values and Curtailment Days

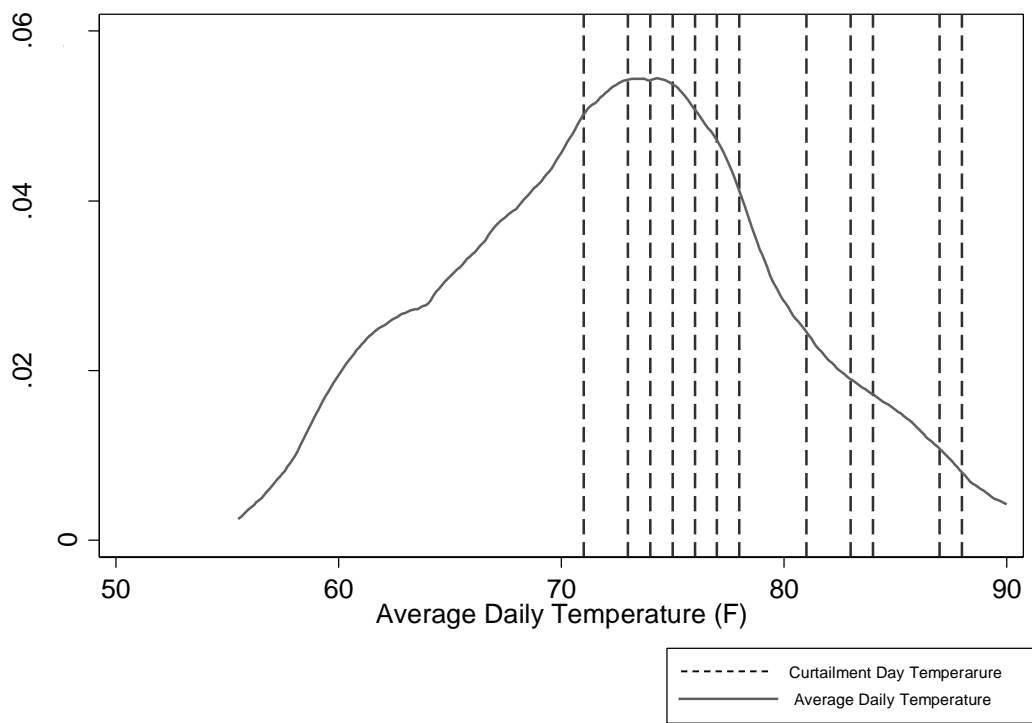


Figure 3-a: Distribution of Estimated First Breakpoint Parameter

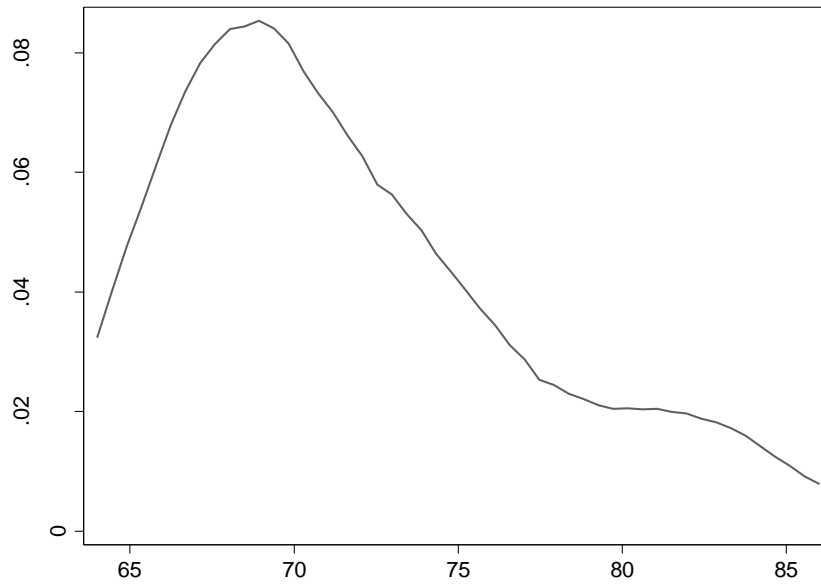
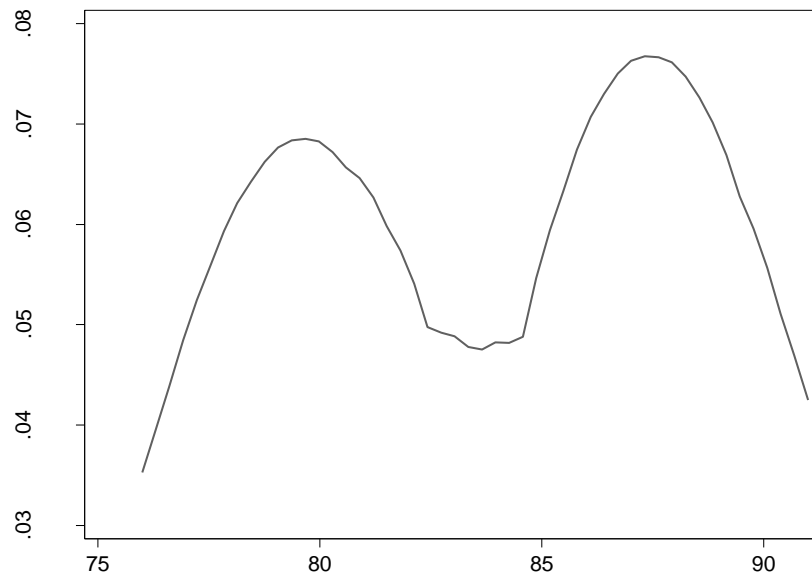


Figure 3-b: Distribution of Estimated Second Breakpoint Parameter



Notes: Figure 3a is a kernel density estimate of the first breakpoint for all the participating AC units, while Figure 3b is the kernel density estimates for the second breakpoint temperature for all participating units. Both densities were estimated with the epanechnikov kernel and optimal smoothing constant of 2.0

Figure 4: Aggregate Precooling over Time

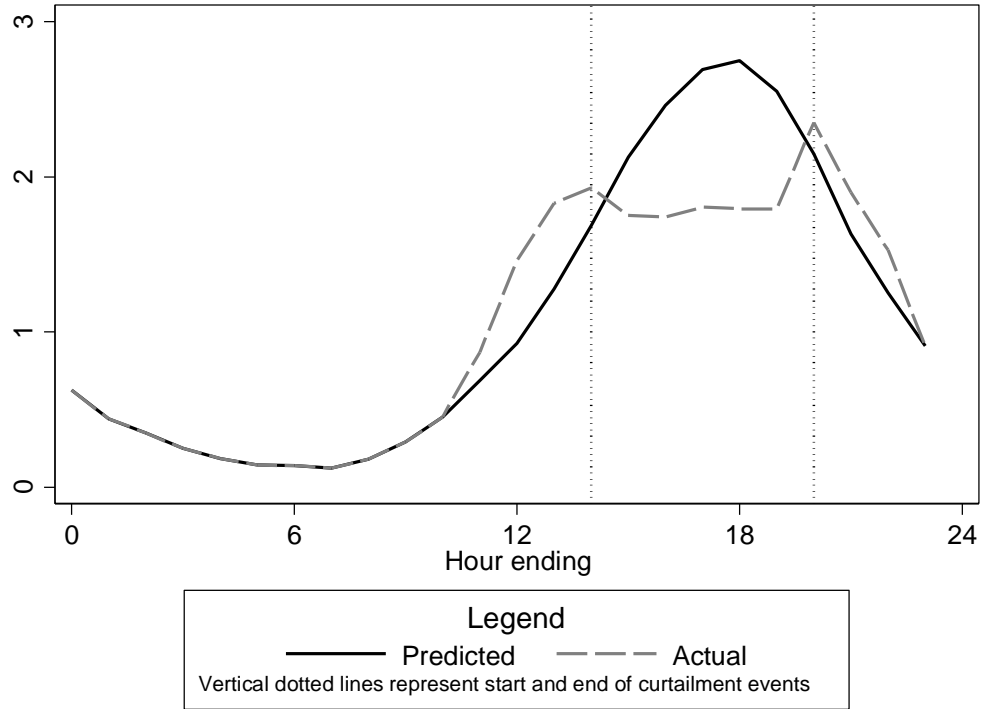


Table 1: Average Precooling, Load Impact Response & Rebound during Curtailment Days by all Units

Date	Average Daily Temperature (°F)	Average Change in Consumption (KWh/h)		
		Before Event	During Event	After Event
12-Jul	71	0.000 (0.000)	-0.331 (0.004)	0.045 (0.011)
17-Jul	73	-0.090 (0.300)	-0.020 (0.006)	0.050 (0.013)
23-Jul	81	0.320 (0.002)	-0.720 (0.013)	0.090 (0.021)
26-Jul	74	0.040 (0.001)	-0.120 (0.001)	0.109 (0.015)
27-Jul	76	0.029 (0.005)	-0.140 (0.018)	0.119 (0.017)
1-Aug	77	0.110 (0.052)	-0.615 (0.012)	0.177 (0.020)
9-Aug	75	0.060 (0.003)	-0.217 (0.013)	0.009 (0.002)
10-Aug	78	0.099 (0.004)	-0.320 (0.001)	0.240 (0.016)
21-Aug	84	0.221 (0.001)	-0.564 (0.030)	0.101 (0.020)
22-Aug	83	0.159 (0.001)	-0.487 (0.001)	0.180 (0.021)
28-Aug	84	0.173 (0.038)	-0.553 (0.001)	0.120 (0.022)
30-Aug	87	0.310 (0.000)	-0.515 (0.001)	0.129 (0.023)
31-Aug	88	0.230 (0.000)	-0.754 (0.034)	0.133 (0.022)
Mean		0.127	-0.410	0.115

Notes: The values in columns 3,4, and 5 were computed by averaging the difference between the predicted mean hourly demand and observed demand on curtailment days. The baseline model was estimated using hourly observations on non-curtailment days alone. Impacts are calculated for each 15-minute interval. The curtailment day loads are estimates is the same for each 15-minute interval in the hour. The impacts are calculated separately for each observed 15-minute average. The average precooling and rebound values were computed by averaging the differences between the observed and predicted cooling loads in the three hours prior to and after the curtailment event respectively. Standard errors are in parenthesis. The standard errors were computed by dividing the standard deviation obtained from a t-test of the difference in means between predicted and observed loads with the number of observations.

Table 2: Heterogeneity in Response to SmartAC Program

	Average Peak Load on non- curtailment days	Average Change in Consumption (KWh)		
		Before Event	During Event	After Event
<u>Panel A: Technology</u>				
PCT Group	2.41	0.190 (0.001)	-0.370 (0.001)	0.273 (0.001)
Switch Group	1.83	0.090 (0.000)	-0.310 (0.001)	0.140 (0.001)
<u>Panel B: Classification by Load Size</u>				
1st Quartile	0.68	0.069 (0.032)	-0.150 (0.005)	0.090 (0.029)
2nd Quartile	1.83	0.114 (0.001)	-0.261 (0.005)	0.073 (0.002)
3rd Quartile	2.56	0.150 (0.003)	-0.470 (0.006)	0.119 (0.005)
4th Quartile	2.67	0.318 (0.020)	-0.570 (0.007)	0.191 (0.001)

Notes: Panel A reports the load impact response by technology type and panel B reports the load impact response by load size in KWh. The reported values are mean values averaged over time and specific class. Standard errors are in parenthesis .

Table 3: Average Precooling, Load Response and Rebound on Proxy days

Date	Change in Consumption (Kwh/h)		
	Before Event	During Event	After Event
11-Jul	0.06	-0.01	0.00
18-Jul	0.01	-0.01	0.00
19-Jul	-0.03	-0.03	0.07
25-Jul	-0.06	-0.01	-0.05
31-Jul	-0.04	0.00	-0.03
2-Aug	0.07	0.00	-0.02
27-Aug	0.00	-0.01	0.00
29-Aug	0.04	0.11	0.00

Notes: Proxy event days are defined as days with similar temperature profile as the curtailment days; and also fall within a few days of the curtailment days (1 or 2 days before or after) the curtailment days that is not a weekend, public holiday, past curtailment day. Model results are calculated for all the curtailment days, however, diagnostics are calculated only for the set of proxy event days.

Table 6: Net Savings with and without Precooling and Rebound

Date	Per Unit Event Savings			Per Unit Net Savings			Conservation Cost (\$/MWh) (G)
	KWh (A)	\$ (B)	lbs of CO ₂ e (C)	KWh (D)	\$ (E)	lbs of CO ₂ e (F)	
12-Jul	1.156	0.724	0.031	1.036	0.339	-0.053	7.24
17-Jul	0.140	0.730	0.133	0.290	0.154	0.109	-2.65
23-Jul	2.884	2.866	1.379	1.654	1.343	0.832	17.14
26-Jul	0.600	3.219	0.237	0.150	0.812	0.033	23.3+2
27-Jul	0.564	2.919	0.281	0.114	0.547	0.072	35.14
1-Aug	2.476	3.053	1.065	1.636	1.908	0.667	30.65
9-Aug	0.864	1.604	0.388	0.635	1.564	0.279	11.85
10-Aug	1.283	2.731	0.589	0.263	1.143	0.106	49.62
21-Aug	2.795	4.692	1.283	1.835	3.771	0.816	53.83
22-Aug	2.398	3.855	0.855	1.378	1.952	0.360	24.17
28-Aug	2.224	3.855	1.247	1.354	2456	0.698	58.23
30-Aug	2.084	4.747	0.923	0.764	2.676	0.323	61.11
31-Aug	3.750	5.986	1.709	2.670	4.957	1.200	46.00
Sum	23.218	41.040	10.12	13.798	23.622	5.442	

Notes: Columns A and B report the savings over the curtailment period in KWh and Dollar values respectively. Column B is computed by multiplying KWh savings by the number of hours and the total KW signed up to the program as at August 31 2007 (number of units multiplied by the tonnage per unit) as well as Net Marginal Price at that particular hour in \$/MWh. Columns D and E report the savings accounting for the precooling and rebound effects in KWh and dollars respectively. Column D was computed by subtracting the amount of precooling and snapback from total reduction during the curtailment period. Column E was computed the same way as Column B, however the net reduction was used. Columns C and F were computed by multiplying hourly change in consumption with hourly emission factors as reported in literature. Column G is the marginal cost of generation for the changes in off-peak consumption.

Online Appendix

for
The Mitigating Effect of Strategic Behavior on the Net Benefits of a Direct Load
Control Program
Corey Lang and Edson Okwelum

(Not for publication)

This appendix provides results that supplement, but are not critical to the analysis in the main paper.

Figures A1-A and A1-B represents plots of precooling over time and temperature respectively. The vertical lines are the average precooling in both graphs. Both figures imply that precooling increases over time and higher temperatures

Figure A2 presents a plot of precooling as function of temperature and date on which curtailment event was called. The vertical line represents daily average temperature while the horizontal axis represents the date on which a particular curtailment corresponding to that average temperature was called. The size of the balls reflects the amount of precooling in KWh/h. The figure reflects the relationship between amount of precooling, average daily temperature and possible effect of learning by participating households.

Figure A3 is a density plot of change in consumption from a Monte Carlo exercise in which we use each non-curtailment day as a proxy event day. We then estimate our model without that day, predict the consumption for that proxy event day and compute the change in consumption between the observed and predicted loads. We repeat for these for all non curtailment days and plot the density in A3.

Figure A4-1 & 2 plots both the observed load and estimated load on the same axes for a single participating unit on a typical curtailment day. The dashed red lines represent the estimated load profile while the continuous blue line is the observed load profile for a single unit. The vertical red lines are the points at which the curtailment event started and stopped respectively.

Figure A5 is a density plot of load with the dashed vertical lines representing the quartile positions.

Figure A6 is a plot of deviation (the difference between predicted and observed loads) one day before and one day after curtailment event.

Figure A7 is a plot of (the difference between predicted and observed loads)

Table 1-A is the hourly change in consumption on all curtailment days studied. The values were computed by subtracting the 15 minute observed load from the 15 minute estimated load. While the baseline model was estimated using hourly observations, the change in consumption was however calculated for each 15-minute interval. The estimated cooling load is the same for each 15-minute interval within each hour. All colored values are significant except on July 12.

Typical Curtailment Event Initiation and Process

Typical process for a control event involves the following steps:

- The evaluators decide what day and time period should be used for the curtailment event based on the weather forecast, and the mix of curtailment and non-curtailment days that were already experienced.
- If the evaluators call the curtailment event, they alternate the two smart thermostat

subgroups within the special monitoring group that received the population treatment and the alternative treatment. If the event is called by PG&E -- i.e., the curtailment applies to all SmartAC participants and not just those in the special monitoring group -- then the entire special monitoring group received the population treatment.

- The day prior to the curtailment event, the evaluators submit these instructions to the PG&E SmartAC Program staff who then relay them to Cannon Technologies which manages the Yukon system that allows the curtailment events to be initiated.
- Third party technology solution provider writes a computer program that will implement the control instructions requested by the evaluators. These programs are tested.
- PG&E's Transmission and Operations staff receive the evaluator instructions and initiate the curtailment events.

Sampling Design and Stratification

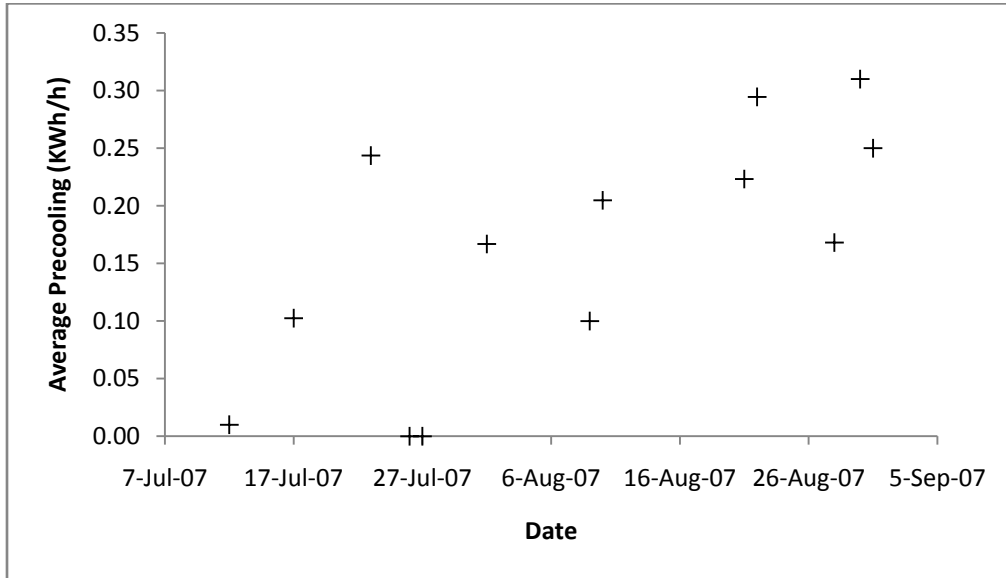
Table 2-A shows the sample design. Sampling for the Special Monitoring Group was stratified into 8 strata on the basis of device type, tons of air conditioning per household and the presence or absence of multiple AC units, as well as other demographic factors.

Description of Curtailment Events Involving the SmartAC Monitoring Group (SMG)

During the period June 07-October 07, PG&E initiated 13 curtailment events affecting Program participants from the SMG. Of these 13 events, all the program

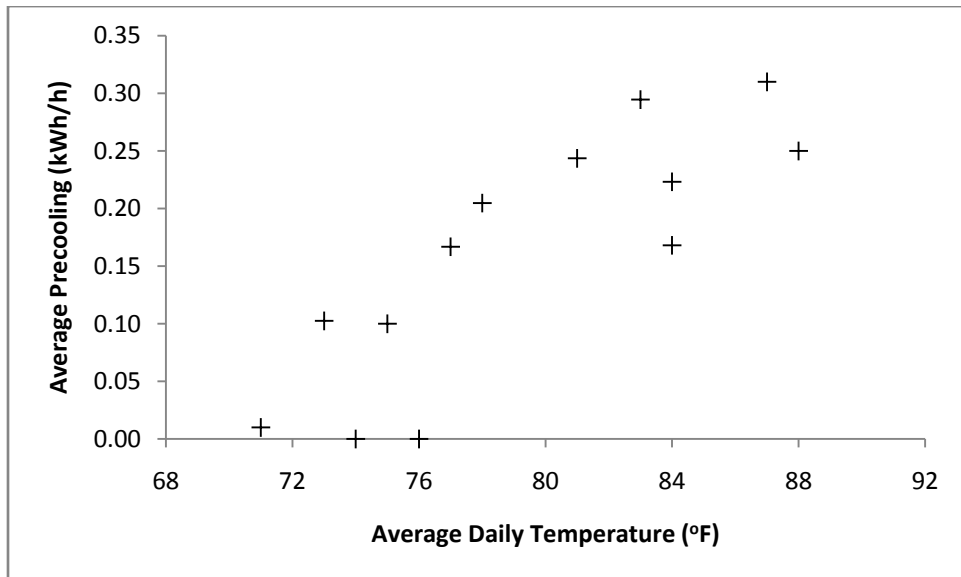
participants were exposed to two of the curtailment events compared to the 13 that the SMG was exposed to: one in July 2007 and another in September 2007. All these curtailment events were initiated by PG&E. Households with smart thermostats were further divided into two subgroups-A & B based on the type of curtailment day and type of treatment received on that date. On the basis of the type treatment, three “Day Types” are recognized: Day Type 0, Day Type 1 & Day Type 2. Day type 0 is used to describe curtailment day in which all the smart thermostats received the same type of treatment. On such days, PG&E set back the thermostat one degree every other hour. We refer to this as the “population treatment”. “Day Type 1” refers to curtailment days when subgroup A received the population treatment and subgroup B received the setback treatment. The setback treatment involved setting back the thermostat one degree every hour for four hours. Day Type = 2 refers to a day when subgroup B received the population treatment and subgroup A received the setback treatment.

Figure A1-A: Aggregate precooling over Time



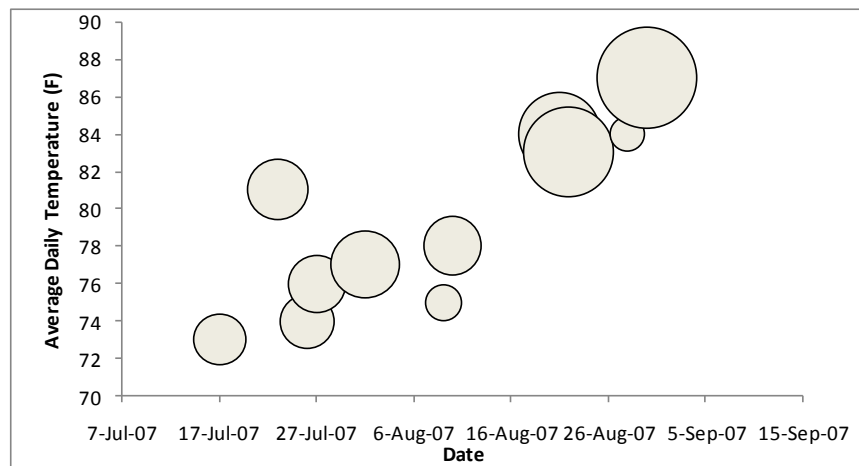
Notes: Vertical axis is the average precooling on curtailment days for all participating units obtained by averaging the difference between the observed 15 minute load demand and predicted 15 minute load estimate. The horizontal axis plots the specific curtailment day.

Figure A1-B: Aggregate Precooling across Daily Average Temperature



Notes: This figure also plots average precooling on the vertical axis; however, the horizontal axis denotes average daily temperature on curtailment days.

Figure A2: Aggregate precooling with respect to time and temperature



Notes: The size of the bubbles represents the size of the average precooling on a specific date and average temperature value on that curtailment day.

Figure A3: Monte Carlo Estimate of Precooling using Non-Curtailment Days

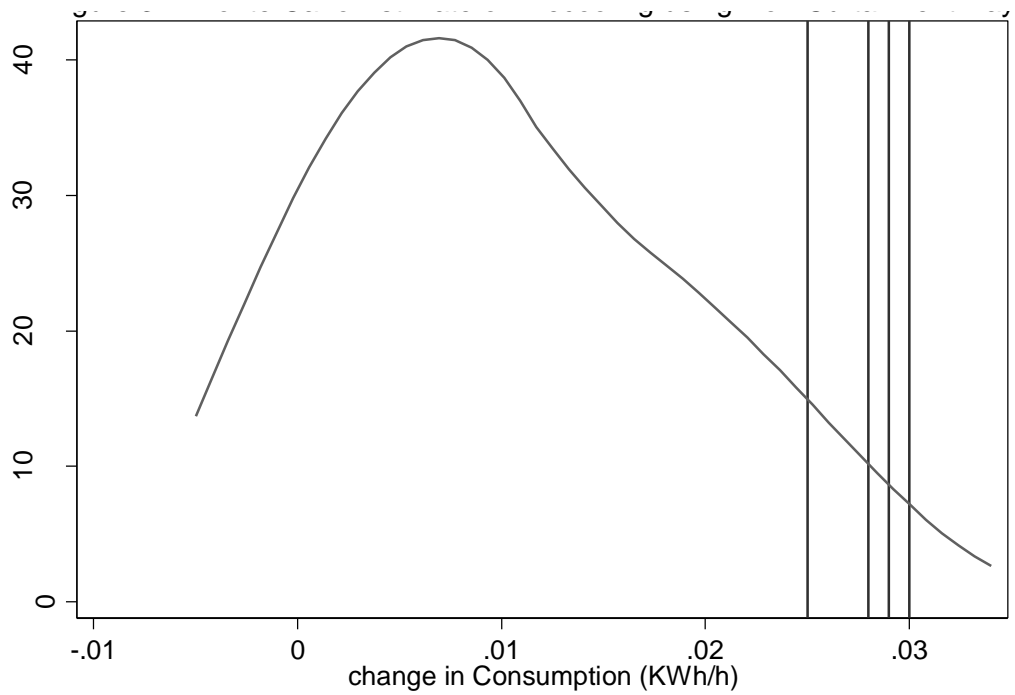


Figure A4-1: Curtailment Day Observed Load vs. Estimated Load



Figure A4-2: Curtailment day Observed Vs. Estimated Load-Individual

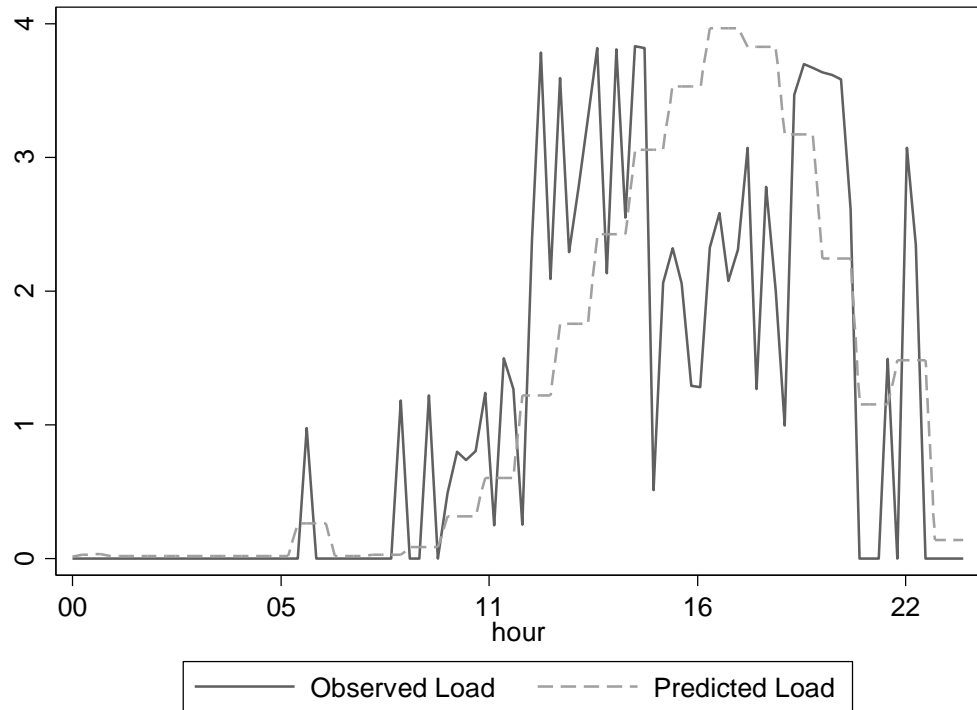


Figure A5: Density of Load Distribution with Quartiles

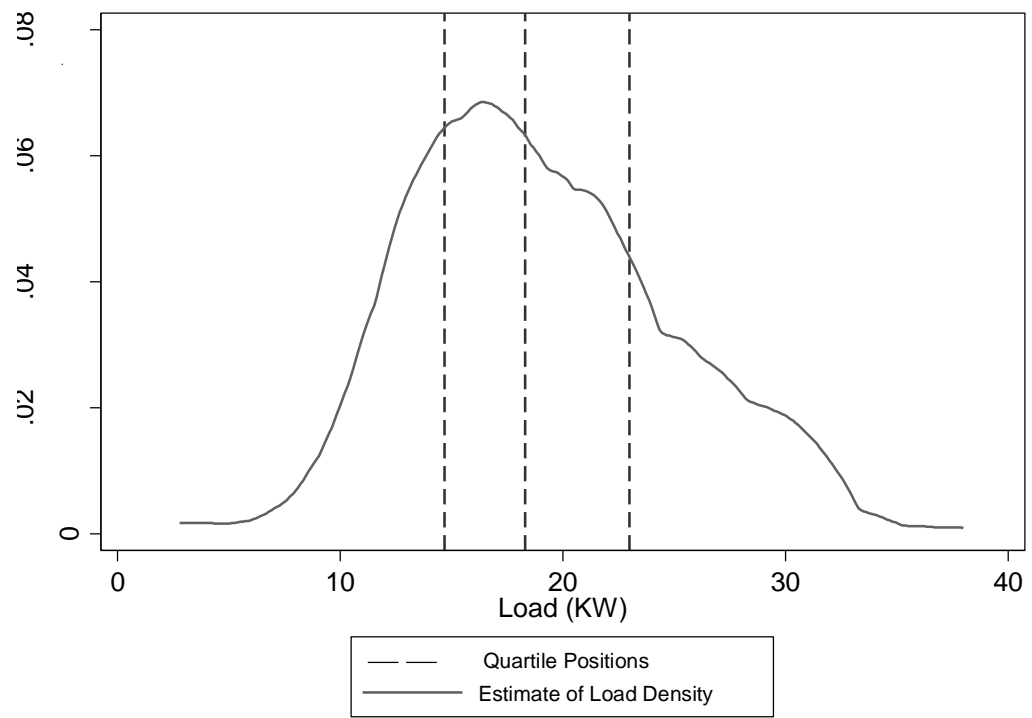


Figure A6: Deviation of Predicted from Observed Load

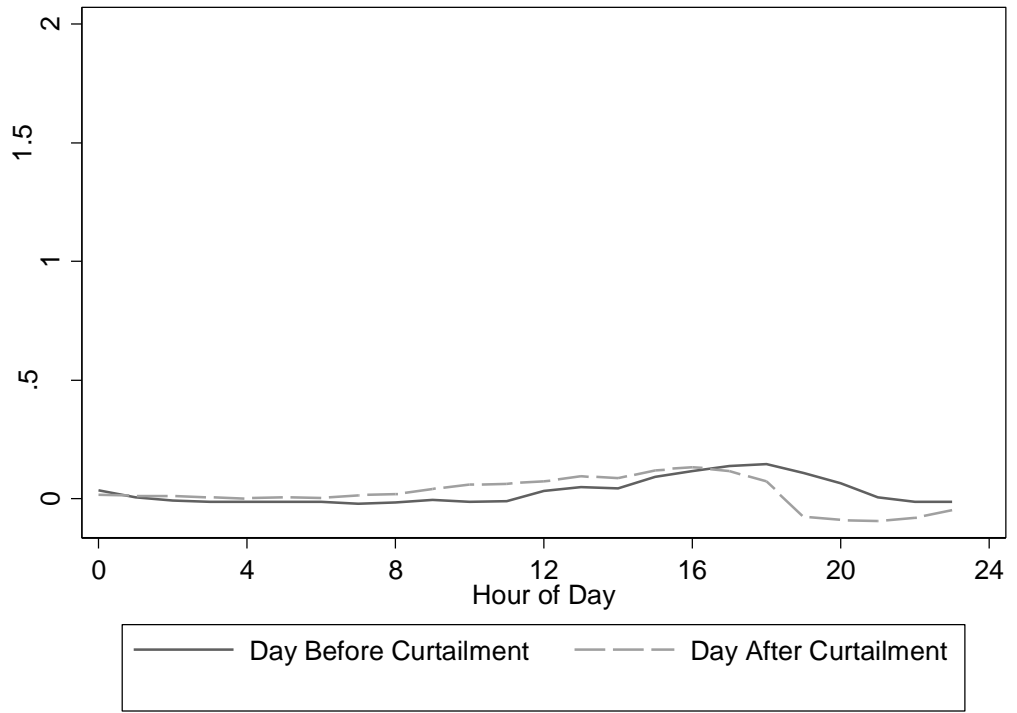


Figure A7: Deviation of Predicted Load from Observed Load

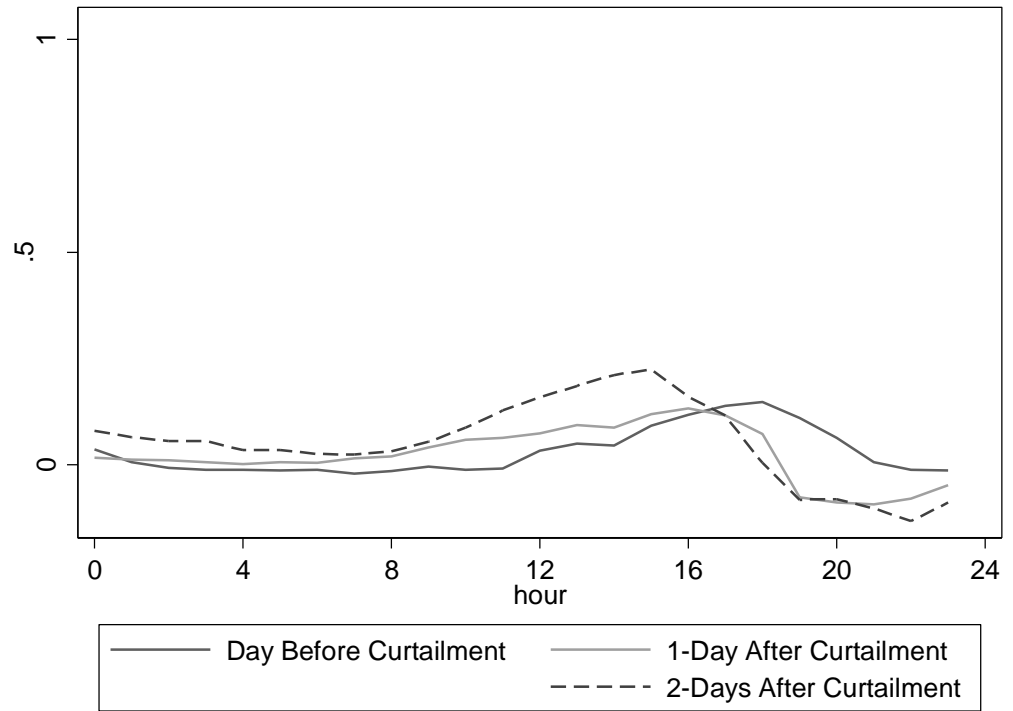


Table 7-A Hourly Changes in Consumption on Curtailment Days

Hour	0.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	12.00	13.00	14.00	15.00	16.00	17.00	18.00	19.00	20.00	21.00	22.00	23.00	
Date																									
12-Jul-07	-0.298	-0.050	-0.033	-0.017	-0.011	-0.011	-0.009	0.023	-0.065	-0.030	-0.053	-0.056	-0.069	-0.110	0.023	-0.423	-0.382	-0.236	-0.195	-0.029	0.010	-0.020	-0.095	-0.056	
17-Jul-07	0.013	-0.039	-0.018	-0.020	-0.014	-0.013	-0.013	-0.014	-0.105	0.128	0.058	0.022	-0.027	-0.068	0.007	-0.051	-0.025	-0.069	-0.157	-0.229	-0.037	0.063	0.146	-0.047	
23-Jul-07	0.821	0.028	0.012	0.014	0.031	-0.010	-0.063	-0.096	-0.121	-0.016	0.233	0.071	0.118	0.147	-0.176	-0.040	-0.394	0.048	-0.026	0.536	0.179	0.144	0.211	0.059	
26-Jul-07	-0.331	-0.011	-0.017	-0.012	-0.008	0.001	0.010	0.036	-0.350	-0.212	0.074	0.098	0.055	0.071	-0.044	-0.561	-0.808	-0.102	-0.336	-0.354	0.009	0.059	0.058	0.000	
27-Jul-07	-0.276	0.022	0.013	0.005	0.004	0.001	0.008	0.009	-0.429	-0.389	0.041	0.118	0.055	0.078	-0.036	-0.134	-0.200	-0.276	-0.516	-0.490	-0.012	0.076	0.072	0.024	
1-Aug-07	-0.236	0.068	0.038	0.023	0.023	0.016	0.053	0.012	-0.414	-0.297	0.240	0.140	0.078	0.139	-0.121	-0.188	-0.219	-0.387	-0.347	-0.296	-0.035	0.140	0.149	0.128	
9-Aug-07	-0.454	-0.024	-0.014	-0.002	-0.009	-0.006	0.004	0.042	0.541	0.543	-0.008	0.019	0.027	0.054	-0.011	-0.068	-0.192	-0.246	-0.450	-0.545	0.008	0.045	-0.017	-0.016	
10-Aug-07	-0.289	0.010	-0.014	-0.003	-0.006	-0.010	-0.052	-0.098	-0.244	-0.297	0.090	0.148	0.041	0.073	-0.229	-0.263	-0.360	-0.255	-0.320	-0.336	0.182	0.092	0.142	0.030	
21-Aug-07	-0.015	0.020	0.000	0.021	0.005	0.021	-0.083	-0.177	-0.005	-0.028	-0.014	0.143	0.160	0.221	-0.318	-0.253	-0.248	-0.245	-0.041	-0.099	0.220	0.174	0.046	0.072	
22-Aug-07	0.118	0.164	0.102	0.100	0.087	0.043	0.023	-0.105	0.094	0.145	0.072	0.145	0.187	0.295	-0.183	-0.542	-0.188	-0.456	-0.575	0.013	0.173	0.212	0.109	0.276	
28-Aug-07	-0.011	-0.088	-0.016	-0.004	0.004	-0.003	-0.102	-0.202	0.025	-0.276	-0.138	0.134	0.159	0.187	-0.248	-0.327	-0.364	-0.351	-0.021	0.029	0.206	0.125	0.017	-0.047	
30-Aug-07	0.378	0.038	0.028	0.272	0.206	0.164	0.112	-0.011	-0.136	0.076	0.388	0.038	0.210	0.262	0.318	-0.244	-0.486	-0.213	-0.846	-0.208	0.145	0.225	0.128	0.331	
31-Aug-07	0.317	0.209	0.145	0.083	-0.002	0.002	-0.202	-0.304	0.391	0.476	0.048	0.238	0.129	0.191	-0.268	-0.326	-0.228	-0.552	-0.248	-0.239	0.220	0.194	0.100	0.355	

Notes: The dark blue squares represent precooling, while the red squares represent changes in cooling load during curtailment period. The dark brown squares are the the three hours immediately following the curtailment events. Results are aggregated event day and hour level with respect to unit tons

Table2-A: Sample Design for EM&V Special Monitoring Group

Stratum	Type of Device	Total Tons From All Units	Multiple AC Units on Site (1=Yes)	Program Participants as of 06/11/2007	Design Sample Size
1	PCT	<4	0	483	93
2	PCT	<4	1	6	3
3	PCT	>=4	0	148	37
4	PCT	>=4	1	34	17
5	Switch	<4	0	1,404	77
6	Switch	<4	1	21	5
7	Switch	>=4	0	637	54
8	Switch	>=4	1	123	17
Total				2,856	303

Notes: Source-EMA,2008

Table3-A: Average Precooling, Load Impact Response & Rebound During Curtailment Days by all Units

Date	Average Daily Temperature (°F)	Average Change in Consumption (KWh/h)		
		Before Event	During Event	After Event
12-Jul	71	0.04	-0.02	0.02
17-Jul	73	-0.04	-0.07	0.06
23-Jul	81	0.32	-0.58	0.07
26-Jul	74	0.04	-0.13	0.11
27-Jul	76	0.03	-0.15	0.12
1-Aug	77	0.11	-0.58	0.17
9-Aug	75	0.06	-0.21	0.01
10-Aug	78	0.10	-0.32	0.24
21-Aug	84	0.22	-0.56	0.10
22-Aug	83	0.16	-0.37	0.18
28-Aug	84	0.14	-0.55	0.25
30-Aug	87	0.31	-0.51	0.11
31-Aug	88	0.22	-0.75	0.13
Mean		0.13	-0.41	0.12

Notes: The values in columns 3,4, and 5 were computed by averaging the difference between the predicted mean hourly demand and observed demand on curtailment days. The baseline model was estimated using hourly observations on non-curtailment days alone. Impacts are calculated for each 15-minute interval. The curtailment day loads are estimates is the same for each 15-minute interval in the hour. The impacts are calculated separately for each observed 15-minute average. The average precooling and rebound values were computed by averaging the differences between the observed and predicted cooling loads in the three hours prior to and after the curtailment event respectively.

Table 4-A: Net Environmental Impacts of SmartAC Program (SO₂)

Average Emission (Reductions)/Increases (metric tons of SO ₂ e)				
Date	Before Event	During Event	After Event	Net Effect
12-Jul	0.00	-1.39	0.18	-1.21
17-Jul	(0.60)	-0.11	0.21	-0.50
23-Jul	1.91	-5.14	0.38	-2.85
26-Jul	0.22	-0.61	0.49	0.10
27-Jul	0.16	-0.72	0.54	-0.01
1-Aug	0.60	-3.17	0.76	-1.80
9-Aug	0.33	-1.07	0.04	-0.70
10-Aug	0.55	-1.64	1.08	-0.01
21-Aug	1.20	-3.64	0.55	-1.89
22-Aug	0.88	-3.12	0.99	-1.26
28-Aug	0.82	-3.61	0.66	-2.13
30-Aug	1.49	-2.73	0.71	-0.52
31-Aug	1.11	-4.92	0.71	-3.10
Mean	0.67	-2.45	0.56	-1.22

Notes: The values in columns 3,4, and 5 were computed by multiplying the hourly changes in consumption with hourly marginal emissions intensities per KWh in the Western Interconnection using SO₂ lbs/MWh values in Kochen and Mansur 2012. The values in the last column is the sum of values in columns 3,4, and 5 respectively. There were 8,843 cooling units enrolled in the program. The average ton per unit was 3.3 and these were then converted to KW by multiplying with a factor 3.517

Table 5-A: Effect Sequence of Curtailment Days

Variables	Estimates
CUMMS+3Hrs before Event	0.009*** (0.0001)
CUMMS+2Hrs before Event	0.019*** (0.0004)
CUMMS+1Hr before Event	0.026*** (0.0001)
CUMMS*Curtilment Period	-0.014*** (0.0005)
CUMMS+1Hr After Event	0.107*** (0.0005)
CUMMS+2Hrs After Event	0.013*** (0.0006)
CUMMS+3Hrs After Event	0.009*** (0.0005)
15-Minute Fes	Yes

Notes: *,**,***-Stars indicate significance effect at 0.10,0.05 and 0.01 levels respectively. The dependent variable is the average 15-minute interval cooling demand (kW) with the same explanatory variables as those used in the main specification. However, the regressions were estimated using GLS estimator and Huber-White robust standard error. The estimates presented here from an aggregate model.

MANUSCRIPT 2

**Fuel Economy and Vehicle Safety: Selection on Moral Hazard and
Estimates of External Accident Costs**

By

Edson Okwelum, Jim Opaluch and Corey Lang

Department of Environmental and Natural Resource Economics

University of Rhode Island

To be Submitted for Publication to the *Journal of Research in Transportation Economics*

Fuel Economy and Vehicle Safety: Selection on Moral Hazard and Estimates of External Accident Costs

By

Edson Okwelum, Jim Opaluch and Corey Lang
Department of Environmental and Natural Resource Economics
University of Rhode Island

June 4, 2014

Abstract

A long held view is that fuel economy standards, designed to reduce emissions from vehicles have led to changes in fleet weight distribution leading to increased accident fatalities. It is sometimes argued that estimates of the cost of a heavier fleet on accident costs could be biased by selection on moral hazard where drivers of heavier vehicles get into deadlier accidents. However, previous empirical research has not accounted for the effect of selection on moral hazard in estimates of accident costs. This paper provides empirical estimates of the effect of changes in fleet weight distribution on vehicle safety due to regulatory constraints. We propose a two-step simulated maximum likelihood estimation strategy to identify and account for selection on moral hazard and unobserved heterogeneity. We find significance evidence of selection on moral hazard effects resulting in downward bias in estimates of the external costs of accident.

Keywords: *Moral Hazard, CAFE, Externality, Selection, SML estimation*

Introduction

In a bid to promote efficiency improvements and reduce competitive impacts across the vehicle fleet, the corporate average fuel economy standards (CAFE) set less stringent targets for larger/heavier vehicles. Two of the most important design parameters that help determine vehicle's fuel economy and CO₂ emissions are weight and performance. Between the late 1970s and 1988, dramatic improvements in fuel economy were accompanied by declines in vehicle weight. And in the two decades between mid-1980s and mid 2000s, fuel economy rose while average fleet weight increased to pre-CAFE levels due to technology innovations. Since 2005 however, improvement in automotive technology has improved fuel economy and performance while keeping weight constant and reducing emissions (EPA 2013). The result is a heterogeneous fleet with conflicting implications for safety.

In motor crashes, the larger weights of the heavier vehicles offer more protection to its occupants while imposing greater risks to occupants of the lighter vehicles. Therefore the social costs of increasing vehicle weight include safety gains by occupants of the heavier vehicle, the safety losses imposed on the struck vehicle, pedestrians, bicyclists, motorcyclists, etc. However, some of the increased risks for occupants who switch to lighter vehicles would be offset by the decreased risk for individuals already in those vehicles (NRC 1992). If the fleet is made up of predominantly lighter vehicles, then a greater percentage of crashes will be between lighter vehicles, resulting in lower fatality rates than if the crashes were between a heavy vehicle and lighter one. In addition, a lighter fleet will result in lower fatalities in collision between vehicles and pedestrians, bicyclists, motorcyclists, etc. In

contrast, if the fleet is predominantly heavy, then we would have higher fatality and injury rates. In addition, given that it is now less costly for drivers of heavier vehicles to be in accidents could induce them to change their risk profile and thus increase the number of accidents. This problem is best formulated as a classic prisoner's dilemma game where each individual reduces risk by increasing weight of her vehicle, but doing so reduces safety to others. The Nash equilibrium will be vehicles that are predominantly light and possibly reduce risk to occupants of other vehicles and lowering of gas mileage.

This paper examines the safety effects of changing fleet weight distribution by estimating how much injury risks would change if consumers purchased increasingly heavier, larger, and less efficient vehicles rather than lighter vehicles given regulatory constraints. The paper estimates both the risk of dying given a crash and the crash frequencies of different vehicles in the fleet. We use a unique data set that combines data from fatality analysis reporting system (FARS) and NASS¹⁸ General Estimates Systems. The estimation strategy used corrects for selection bias due to moral hazard problems. The two equations are estimated simultaneously using simulated maximum likelihood without the need for exclusion restrictions (Green 2003, 2007). This approach allows us to obtain estimates that are causal.

While there are several works looking at the effect of vehicle weight and size on traffic costs, past empirical estimates have ignored the issue of selection on moral hazard. Jacobsen 2013 finds that increasing the weight of a vehicle by 1000 pounds increases the number of fatalities in other vehicles by 45% while reducing own

¹⁸ National Automotive Sampling System

fatalities by 54%. Jacobsen adjusts for sample selection bias that may exist due to the lack of reliable data on non-fatal crashes by linking driver riskiness to increased fatalities from single vehicle crashes. Anderson and Auffhammer 2013 find a 40-50%¹⁹ increase of risk on others of increasing own weight by 1000 pounds. They make no specific assumption about selection bias, arguing that their unique data set and multiple identification approach preclude selectivity. Though they include several driver, vehicle and environmental risk factors.

Evans (2001) estimates both the protective and aggressive effects of increasing weight by 1000 pounds of 40% and 42% respectively. He uses differences in the number of occupants in the struck and striking vehicles to identify the effect of weight. This strategy helps avoid a host of selection issues, since it allows weight to vary holding all other attributes of the vehicle fixed. Kahane (2003) also finds estimates in the order of -30% to 70% for the protective effect of increasing vehicle weight and cautions that his estimates for the heavier vehicles may be biased due to driver selection.

However, an important potential weakness in these approaches is the existence of Peltzman-type moral hazard problems and endogeneity of vehicle choice. In the presence of Peltzman-type moral hazard problems, drivers in heavier vehicles will find it advantageous to change their behavior in ways that have opposite effect (Pederson 2003, Peltzman 197, Risa 1992, 1994). The change in their risk behavior results in

¹⁹ Anderson and Auffhammer notes that these are lower bounds, given that heavier vehicles are involved in more collisions.

their being involved in more fatal crashes²⁰. Given these, it is more likely that estimates of the causal effect of weight in literature are biased. In contrast, it is also possible that safety conscious drivers may buy heavier vehicles and drive more cautiously. Which of these effects dominates is an empirical question. There are two challenges with correcting for selection on moral hazard issues. The first issue is whether the incentive effect occurs at all. The second issue is whether the effect is large enough to completely offset the reduction in the probability of injury such that the number of injuries increases as the vehicle fleet becomes heavier and safer. Also from a public policy perspective if increasing the weight of a vehicle increases the probability of its being involved in collisions, then we might be interested in more than the impact of heavier vehicles conditional on a crash occurring.

We make three main contributions in this paper. First, we provide empirical evidence of effect of selection bias due to moral hazard in estimates of the effects of heavier vehicles on traffic injury rates. Combining results from the theoretical framework and empirical estimates, we develop simple qualitative tests to show evidence of selection bias due to drivers of heavier vehicles involvement in more fatal/serious injury accidents. We find that increasing the weight of any vehicle increases the probability of it being the heavier vehicle in a collision and the probability of it being involved in a fatal accident. The external costs translate to about 38 cents/gallon of gasoline. These values though larger, are comparable to values obtained by Jacobsen (2013) and Anderson and Auffhammer (2013). The second contribution of the paper is methodological. It features an application of bivariate

²⁰ White (2004) and Anderson (2008) estimate that light trucks are 13% to 45% more likely to experience multivehicle collisions than passenger cars.

random effects estimator in a count model (Riphahn *et al* 2003; Terza 2010; Green 2012). The third contribution relates to the use of the underlying estimates of the rates of traffic injury in policy simulations.

The paper is related to several lines of literature. First is the large volume of literature on impact of CAFE standards on traffic safety (NRC 1992, 2002). Much of the debate has centered on the relationship between weight, fuel economy, vehicle choice and safety (Li 2009; Crandall and Graham 1989; Noland 2004; 2005; Parry 2006; Small and Van Dender 2006; Khazzoom 2000). Another is the literature on individual choice of activity on self protection, averting behavior and moral hazard in traffic safety (Pedersen 2003; Blomquist 2004; Bolmquist, Miller and Levy 1996; Avery, Heyman and Zeckhauser 2001; Ashenfelter and Greenstone 2002; Chestnut 2001; Li 2012 ;). The third area is the literature on the protective and aggressive effects of vehicle weight on occupant safety (NRC 2002; Peltzman 1975; White 2004; Gayer 2004; Evans 2001; Anderson 2008; Hultkrantz and Lindberg 2011; Jacobsen 2013; Anderson and Auffhammer 2013). This paper is most closely related to the work of Jacobsen 2013, and Anderson and Auffhammer 2013. However, this work is different from theirs in several respects. Unlike Jacobsen we use data from both fatal and nonfatal accidents²¹. Our paper is also different from the two papers because we account for selection bias arising from moral hazard through the correlation of unobservables in the two dimensions of vehicle safety.

The rest of the paper is organized as follows. The next section describes the theoretical framework. Section III presents the data, while section IV describes the

²¹ Anderson and Auffhammer also use data on both fatal and nonfatal accidents, however, they do not account for selection on moral hazard.

specification. Section V presents the results. Section VI performs policy experiments, while section VII concludes.

Theoretical Background

The theoretical framework is an extension of the standard law and economics of the choice of care model that accounts for a policy variable (Liao and White 2004; Shavell 1979, 1982; Cummins et al 2001). Road transport is generally recognized to be characterized by several externalities –accident, air pollution, congestion and other externalities (Parry et al 2004, Elvik 1994, Newberry 1988). The severity of an accident is influenced by speed and rate by traffic density. The effects of these two factors tend to work in opposite directions, as increased traffic levels are often associated with a reduction in average speed. The external effects arise anywhere the risk faced by one driver depends on the actions of other drivers. The externality arises from the fact that even a careful driver could be injured in an accident that is caused by another driver who was careless. On the other hand, the actions of the careful driver will lead to reduction of risks imposed on others. Defining externality as above makes it easy to express accident externality as a function of the dimension of behavioral choice that is faced by the agent. Generally, when externalities are present, equilibrium behavior is inefficient and the overall level of actions taken by society to prevent accidents will be inadequate.

We assume that accidents are bilateral, drivers maximize expected utility with initial wealth w_i , are risk averse with respect to effort, and dislike effort or care (e_i). The accident probability is denoted by φ and it is a function of the vehicle miles

travelled- vmt and care level of each driver. Hence we write the probability of accident for driver i as

$$\varphi_i = \varphi_i(e_i, vmt)$$

with the probability of accident falling with care ($\partial\varphi/\partial e' < 0, \partial^2\varphi/\partial e^2 > 0$). There are two states of the world-the loss state with probability φ and the no-loss state with probability $1 - \varphi$. The utility function is written as

$$u_i(w_i, e_i)$$

The utility is twice differentiable. These *priori* assumptions about the functions are quite common with models of traffic safety. First it is assumed that utility is increasing and strictly concave function in wealth and that the drivers' marginal utility is strictly decreasing in her effort to be careful. Where $u_w > 0, u_e < 0, u_{ee} < 0$ and $u_{ww} < 0$ ²². The accident losses are assumed to be non-stochastic and consist of economic losses represented by l , (medical bills, lost earnings, general damages). There are other losses such as pain and suffering that are not fully covered and are represented by g . Insurance is available at actuarially fair premiums and drivers are assumed to be fully insured. Hence the premium is given as

$$\pi_i = \widehat{\varphi}\lambda_i l$$

The accident cost l is a function of several parameters which are differentiated among drivers such as severity of the accident, λ_i , accident environment, size of family, etc. The severity in turn depends on several vehicle attributes such as weight

²² There are no restrictions on u_{eu}

of both involved vehicles, (m_i, m_j) . Therefore we write

$$\lambda = (\lambda_i, \lambda_j) \text{ with } \lambda_i = (m_i, m_j), \lambda_j = (m_i, m_j), \frac{\partial \lambda_i}{\partial m_i} < 0, \frac{\partial^2 \lambda_i}{\partial m_i^2} > 0, \frac{\partial \lambda_j}{\partial m_j} > 0, \frac{\partial^2 \lambda_j}{\partial m_j^2} > 0,$$

Increasing own vehicle weight reduces injuries in this struck vehicle, but at the same time increases probability of injuries in other vehicles. Given the above assumptions, driver i maximizes the expected utility given below:

$$EU_i = [1 - \varphi(e_i, vmt)]w_i + \varphi(e_i, vmt)[(w_i - \lambda_i(g + l)) - p_m m - \pi_i(\cdot)] \quad (1)$$

with respect to weight of car, m^{23} , where p_m is the price per unit weight of car m .

The first order condition with respect to m gives the optimal choice of weight for each driver.

$$\frac{\partial EU_i}{\partial m_i} = -\varphi(e_i, vmt) \left[\frac{\partial \lambda_i}{\partial m_i}(l) + \frac{\partial \lambda_i}{\partial m_i}(g) \right] - p_m = 0 \quad (2)$$

The sum in brackets is the marginal benefit of additional unit of expenditure on weight. The last term is the marginal cost of incurring an additional expenditure on weight. The above FOC shows that an increase in weight will result in lower injury risk reflected in the lower losses. A risk adverse driver who maximizes utility will choose the optimal vehicle weight (or other attribute) such that the marginal benefit of having that extra weight just equals the increased cost of weight. This benefit could be in the form of extra protection from an accident and the cost includes things like additional expenditure on gasoline. If individual drivers choose m to maximize utility,

$EU(m, e, w, l)$, the sufficient condition for optimality is

²³ Another physical factor affecting injury likelihood is vehicle size, specifically the distance from the front of a vehicle to its occupant compartment. The longer this is, the lower the forces on the occupants, provided vehicle designers take advantage of the extra length. In theory the lighter weights of smaller cars could be offset by increasing the sizes of their front ends, keeping weight down by using materials like aluminum, plastic, or titanium. But this typically doesn't occur because such materials cost so much.

$$A \equiv \frac{\partial^2 EU_i(\widehat{m}, e, w, l)}{\partial m^2} < 0 \quad (3)$$

Differentiating equation 2 with respect to effort leads to clear conclusions about how a risk adverse driver's effort changes with the optimal weight of a vehicle and probability of accident. The results are intuitive behavioral responses. For example, an increase in weight reduces the marginal effect of the driver's effort on the cost of an accident, i.e. the safer the car, the lower the marginal benefit of effort and thus the lower the driver's optimal effort level leading to higher accident rates. This is an example of the rebound effect discussed in literature (Sorrel 2007). In effect, an increase in weight could lead to an increase in accident frequency given that the higher weight offers more protection. Anderson and Auffhammer (2013) argues that other than the behavioral response, the physical features of heavier vehicles could also induce them to be involved in more collisions.

General results from partial equilibrium suggest that the behavior represented by the above FOC is inefficient. To fix ideas, compare this to a social optimum case in which there are two types of drivers i and j . The first order condition leads to a situation in which the driver i chooses the level of policy variable that is higher than the social optimum such that her injury risk decreases and that of the second car increases. The degree of concavity of $u(w)$ and the convexity of the loss functions ensures existence of interior solutions. The FOC for this is shown below:

$$\frac{\partial(EU_i + EU_j)}{\partial m_i} = -\underbrace{\varphi \left[\frac{\partial \lambda_i}{\partial m_i} l + \frac{\partial \lambda_i}{\partial m_i} (\cdot) \right]}_{\text{internal benefit}} - p_m - \underbrace{\varphi \left[\frac{\partial \lambda_j}{\partial m_i} l + \frac{\partial \lambda_j}{\partial m_i} (\cdot) \right]}_{\text{Negative Externality}} = 0 \quad (4)$$

The first part of equation four is similar to the benefit part of equation 2. Equation 4 is the standard outcome with negative externality and the solution is Pareto optimal because the maximization of any social welfare function that is based on the utility function of individual parties is Pareto optimal (Silberberg 1990). The second part of the equation above is the negative externality that results from an increase in mass of one vehicle relative to the other vehicle. The external cost component includes all costs not taken into account by the driver himself. For example, loss of life and health inflict pain and suffering on victims' relatives; these are not taken into account by the driver (Jansson 1994).

If the risk is not fully internalized, the market will produce too large of the last part of equation in brackets. For each choice of m_j , there is a socially optimal choice $m_i^*(m_j)$ and vice versa. The overall social optimum will occur where the equations cross each other and we will have a Nash equilibrium such that $m_i > m_i^*$ and $m_j > m_j^*$. Without regulation of the fleet, the optimal action of i vehicle is to impose more risk on j and j 's best action is to take more averting behavior. The Nash equilibrium could lead to too much risky behavior by one driver and subsequently more precautions being taken by other drivers at the social optimum. Whether or not this produces less risk than the social optimum is an econometric question. Also, rather than being symmetric, the nature of the resulting externality is asymmetric. Generally, the situation is such that we have one person (m_i) imposing the risks and another taking averting behavior (m_j). The asymmetric nature of the equilibrium is significant in that, we can regulate the actions of m_i alone to restore the social optimum (Avery *et al* 2001).

To obtain a tractable form for the estimation equation, we make the following additional assumptions and simplifications. We can represent the external costs part of the accident as

$$\alpha F_{ij}, \quad (5)$$

where α is the cost per unit of accident (statistical value of life or statistical value of injury), and F_{ij} is the number of fatalities or serious injuries in the struck vehicle in a crash with vehicle j . F_{ij} is attributable to several causal factors which we will discuss later. We will also make assumptions about the general form of F_{ij} , the distribution of fatality or injury states. If X denote the causal factors determining F_{ij} , then we have

$$E[F_{ij}|X] = f(x) \quad (6)$$

where $f(x)$ is some regression function of the causal factors.

Data

The paper uses several datasets for the analysis. The main datasets are census of all fatal accidents (FARS) and the National Automotive Sampling System General Estimates

System (NASS-GES) covering the period 2008 through 2012. The paper also uses data on the number of registered vehicles in each state by type and model from Polk NVPP as well as number of miles driven by each vehicle class from the Federal Highway Performance Monitoring System. We also include information on state level demographic attributes, weather variables and other macroeconomic variables for the analysis.

Accident Data

The accident data are from several sources. FARS and NASS GES provide data on vehicles involved in crashes. Vehicles involved in crashes are grouped into three different classes based on severity of injury: fatal crashes, injury crashes, and property damage only (PDO) crashes. Approximately 10 million passenger vehicles per year were involved in police reported motor vehicle crashes between 2008 and 2012. This average figure reflects a crash involvement rate of over nineteen passenger vehicles per minute for the years covered in this paper. Since the GES observations are from a sampling survey, they therefore have a sampling weight. The combined data for this study contained 150,633 fatal involvements and 263,749 unweighted GES accidents, when weighted, results in an estimated 27,191,919 crash involvements.

A closer look at the data from FARS and GES reveals that light duty trucks (SUVs, Pickup Trucks and Minivans) crash involvement rate in fatal accidents was 1.82 (0.250²⁴) while for passenger cars was 1.41(0.18). Involvement rate of passenger vehicles in injury accidents over the sample period was 112 (9.981) while for light duty trucks were 102 (9.718). Looking only at accidents classified as PDO, passenger car rates over the same period was 255 (7.705) and the values for light duty trucks was 254 (14.35). It is also worth noting that over the same period, the number of registered SUVs and trucks increased by over 70%.

NASS GES contains detailed information on a representative sample of thousands of property damage/no injuries, serious injury, and fatal crashes. Field

²⁴ The involvement rate values reported are rates per 100 million vmt. The values in parentheses are the standard errors.

research teams located at Primary Sampling Units (PSU's) across the country study about 50,000 crashes a year involving passenger cars, light trucks, vans, and utility vehicles. The NASS GES relies on a select sample of police accident reports (PARs) from 400 police jurisdictions within 60 geographic areas (PSUs). Data from two-vehicle crashes in which occupants were either killed or seriously injured were extracted from the dataset. The dataset was further restricted to vehicles with model years later than 1979. Using the NASS vehicle body type classification, all vehicles in the dataset were grouped into four classes- passenger cars, vans, SUVs and pickup trucks. There were 417,510 observations satisfying the above criteria of which 86, 567 observations were single-vehicle crashes and 265,633²⁵ observations were two-vehicle crashes. The remaining 74,310 observations are crashes in which multiple vehicles were involved and were subsequently dropped. Since each two-vehicle crash contains two vehicles- the “own vehicle” and the “other vehicle”, care was taken not to double count observations by assigning a group identification number based on accident case number, year of accident and vehicle number in the NASS GES data. The model estimation was based on only those observations for which all the independent variables were not missing. We also dropped accidents in which school buses, emergency and police cars were involved.

Table 1 presents summary statistics of the accident data. Columns 1,2 and 3 report statistics for the entire two vehicle collision dataset. The mean vehicle weight is 3,388²⁶ pounds with average age of 2000. The number of fatalities per accident is 0.02612 with standard error of 6.5, while the number of serious injuries per accident is

²⁵ That comes to about 11,538 accidents after data cleaning

²⁶ The ratio of “other” vehicle weight to “own” vehicle weight is 1.01

0.746 with standard error of 36.7. In our sample, alcohol is involved in 6.1% of all two vehicle crashes. In all two-vehicle crashes, 92% of the drivers had some form of restraint such as seat belt.

Vehicle Miles Driven and Car Registration Data

Vehicle registration data are compiled by R. L. Polk & Company in the National Vehicle Population Profile (NVPP). State level data on vehicle miles travelled by vehicle class were obtained from the highway statistics series of the Office of Highway Policy information from the department of Transportation. The data are broken down by state, vehicle type (passenger car, vans, SUV, pickups and other pickup types), and type of road (rural or urban). Since the data contain separate information on SUVs, vans, pickups and passenger cars at the state levels, I am able to test whether crash frequencies for the different injury severity levels vary across vehicle class types. The data cover the period 2008 to 2012 for all states in the US. Closely related to the vehicle miles travelled data is the number of registered vehicles by vehicle type at the state level and the public road length by state in miles. These were also obtained from highway statistics series of the Department of Transportation.

Demographic Characteristics

Additional information on state level demographic characteristics like per capita alcohol consumption, unemployment rate, population data, and investment on road covering the period 2008 to 2012 were also obtained from several sources. The per capital alcohol consumption by state was obtained from surveillance report on 1977–2012 apparent per capita alcohol consumption in the United States published annually by National Institute on Alcohol Abuse and Alcoholism (NIAAA). The data

are based on alcoholic beverages sales data. The population data and unemployment rate were obtained from the statistical abstracts of the US published by the US Census Bureau.

Econometric Model

The number of fatally or seriously injured occupants, F_{ij} , in the struck vehicle is a count variable and nonnegative. Therefore we are inclined to use count models for our estimation. However, we are faced with two main challenges. First crash severity, a function of the weight of the two vehicles and their relative velocities at the point of impact is not observed. It reflects the influence of the mass of both vehicles and the behavior of both drivers²⁷. And there are several interrelated human factors, which are not easily quantifiable, but are however related to vehicle weight and crash severity. In effect, unobserved factors related to weight and crash severity may be correlated with the error terms thereby introducing endogeneity. Second, from the derived external cost of accident section 2, one obtains that drivers of heavier vehicles are induced to drive more dangerously given that it is now less costly for drivers to be in an accident. Drivers will expend fewer resources to avoid being in an accident, and thus the number of accidents will increase. This is the well known moral hazard effect discussed by Peltzman (1975). Selection on moral hazard implies that drivers of heavier vehicles are more likely to be involved in more fatal accidents in which occupants of other vehicle are killed or seriously injured. In effect, drivers of heavier vehicles could be induced to drive more dangerously because of several factors that

²⁷ Several approaches have been used to control for crash severity. See Toy and Hammit for the various authors researchers have used to control for crash severity.

are not clear²⁸. If drivers of heavier vehicles in our sample are on average more likely to be involved in more severe crashes, then selection bias would be a problem.

To account for unobserved driver effects, while also correcting for possible selection on moral hazard, we use a two-step procedure similar to that developed by Terza Nelsusan (1994). In the first stage, we estimate the probability of being involved in an accident. These estimates are then plugged into the second stage negative binomial regression estimates of the injury severity equation.

The selection model is

$$P[Acc_{its} = 1 | w_i] = \Phi(w_i \gamma + \mu_i) \quad (7a)$$

$$w' = [vmt, l_{ts}, X_{is}, YEAR2003 - YEAR2012]$$

where Φ is a standard normal distribution function. But we do not require it to be normal. Acc_{ist} is number of crashes in state s by vehicle type i at time t ²⁹, vmt is the vehicle miles travelled, l_{ts} ³⁰ is the traffic density at time t in state s , X_{its} represents all other covariates such as weight of the vehicles, per capita alcohol consumption, share of population using seat belts, several variables representing population distribution, rural and urban speed restrictions, measures for weather variables (number of days with snowfall greater than 1 and 5 inch, number of days with temperature lower than $0^\circ F$, number of days with precipitation greater than 5 inches), control for census division, and investment in road network.

²⁸ It is also possible that safety conscious drivers could pick heavier vehicles because they are safer, in part explaining the reduction in fatality or injury rates of such vehicles.

²⁹ It can also represent 1 if the vehicle was involved in a fatal crash and 0 otherwise, which is the well known Probit selection model. We tried both specifications and the results are similar.

³⁰ The traffic density is defined as the ratio of VMT over length of public road.

We specify the number of fatalities or serious injuries in each vehicle as a function of the driver characteristics, the characteristics of the two vehicles, the environment in which the accident took place, and other covariates of interest. The model is:

$$P(F_{ij}|m_1, m_2, \varepsilon_i) = \exp(\alpha + X_i' \beta + \sigma \varepsilon_i), \quad \varepsilon_i \sim N[0,1] \quad (7b)$$

where $X_i =$

$$(m_1, m_2, \text{restraint}_i, \text{Airbag}, \text{SPDL}, \text{NoOccupants}_i, \text{Others}_i)$$

The dependent (F_{ij}) variable is the number of occupant injuries or fatalities in the struck vehicle. m_1, m_2 denote the curb weight of the struck and striking vehicles respectively, Restraint is a dummy indicating whether the driver of the struck vehicle properly used a seat belt and Airbag is also a dummy indicating whether a driver-side air bag was deployed, SPDL denotes the speed limit on the accident location. *Others* include all other independent variables like location of accident (rural or urban road), dummies indicating if the driver is a male and greater than 65 years, dummy for gender of the driver, dummy if the driver is less than 21, dummies for if alcohol and drug was involved. A series of three indicator variables is used to capture the impact of crash configuration on the injury risk, year and state effects. ε_i the unobserved heterogeneity, is normally distributed with mean zero and standard deviation σ ; and in our framework this includes distance from a hospital, mental state of the driver that day, did the driver suffer heart condition, etc. Given the presence of the unobserved heterogeneity in equation 7b, Green (1998, 2008) and Cameron and Trivedi (1990,

1998, 2005) show that the right hand side of equation 7b is Negative Binomial (NB) distributed.

Writing the models in this form allows us to operationalize the selection problem through the correlation of the ε and μ error terms. The unobservables, ε and μ are assumed jointly normally distributed, i.e.

$$(\varepsilon_i, \mu_i) \sim N(0, 0, \sigma_\varepsilon^2, \sigma_\mu^2, \rho) \quad (7c)$$

Where σ_ε^2 and σ_μ^2 are the variances of the ε and μ respectively, ρ is the correlation between the error terms. Conditional on selection, the dependent variable is unlikely to be either Poisson or NB distributed which would allow estimation through usual methods. Drawing on the results of Terza (1995, 1998) and Green 2008, the likelihood function for the full model is equivalent to the joint density function for the observed data. The full set of parameters (σ_ε , σ_μ , β , γ and ρ) can be estimated through a single step simultaneously using full information maximum likelihood (FIML) or simulated maximum likelihood (SML) procedures without the need for exclusion restrictions (Ripham *et al* 2003, Green 2008). We proceed by forming the log-likelihood from the joint density of the selection model (6a) and injury severity (6b) as $P(F_{ij} | Acc)f(Acc)$. Following Green (1997, 2008), the log-likelihood function is estimated with the use of a SML.

Results

Model Performance

As a way of evaluating the preference of the selection model over the simple Poisson or NB specifications, we first estimated both models and then conduct

hypothesis tests based on the estimated correlation parameter, ρ , as well as a likelihood ratio test. Tables 2 and 4 present the regression estimates for the models without selection and with selection respectively. The results suggest that there appears to be selection effects in the estimates, as shown by different coefficients in the conditional model. A Wald test for the existence of selection effect against a null hypothesis that the correlation coefficient is zero, i.e. $H_0: \rho=0$ is rejected. The Wald test produces a test statistic of $(-100.024)^2 = 50.012$ (with a p value of 0.000), which is larger than the critical value of 1.96 for a standard normal, leading to the rejection of the null hypothesis. A likelihood ratio test was computed as the difference between the log-likelihood for the full model with selection and log-likelihood for the model without selection effect (when ρ equals zero). The result is

$$\lambda_{LR} = 2[-4,898.88 - (-260.802 + (-4,803.801))] = 331.446$$

and the hypothesis is also rejected.

This value leads us to reject the simple Poisson or NB model while accepting the bivariate model that allows significant correlations across the two dimensions of the vehicle safety. The test statistics confirm the high statistical significance of the structural parameters σ_v , σ_μ and ρ . The negative sign on the correlation coefficient indicates that the latent factors which decrease the severity of injury in two vehicle crashes increase the probability of being involved in an accident, a situation that is consistent with the selection bias. The model with correlated counts is deemed the most appropriate one to use in this situation and is the basis for the policy simulations in a later section.

Further, we also investigated the fit for the data generating process for the two count variables. Figures 1A and 1B show that the NB model is a better fit for the data generating processes. Figure 1A represents the mean-variance relationship for the accident frequency model while Figure 1B is the variance relationship for the severity model. The Poisson and Negative Binomial (NB) models both do a good job of predicting fatal accidents at low mean values. However, the Poisson model does not do a good job at mean values above 0.2 compared to the NB model. Figure 2 is plot of predicted probability of being involved in a fatal accident at different weight classes. The figure suggests that as the weight increases, the probability of being involved in a fatal accident increases. The following sections discuss the results from the first stage and the second stages as well as policy simulations.

Selection Model

Table 3 reports the first-stage regression results, i.e. regressing the whether or not a vehicle was involved in an accident in a state or not against the weight of the vehicle, vehicle miles travelled and other covariates. The linear probability estimates and the Probit estimates indicate that vehicles which are heavier by an average of 1000 pounds are 0.1-0.15 more likely to be involved in more fatal or serious injury accidents given their exposure, age, and driver attributes, etc. The coefficient of vehicle miles travelled and the share of light trucks are significant and positive suggesting that increasing the distance travelled increases the probability of being involved in an accident. The positive sign on the weight coefficients are in line with prior estimates of heavier vehicles being involved in more crashes elsewhere (Anderson 2008, White 2004, and Gayer 2004).

Fatalities in Two-Vehicle Crashes

Table 4³¹ presents results of the second stage regression estimates of bivariate model for number of fatalities in struck vehicle against covariates of interest and other confounding attributes. The key explanatory variables are the curb weights of the struck (m_1) and striking (m_2) vehicles. The two weight variables are in 1000 units so that their coefficients measure the factor change in expected number of fatalities in the struck vehicle when the weight of struck or striking vehicle changes by 1000 pounds in line with earlier studies. The coefficients on the struck and striking vehicles have the predicted signs. Other covariates include speed limit at the point where the accident occurred, age of the driver, number of occupants in the struck vehicle, dummies to indicate whether or not the driver was wearing seat belt, dummies to indicate the role drug or alcohol played in the accident. It also includes state by year fixed effects and model year fixed effects.

Columns 2 and 3 report the second stage results along with the model without selection for comparison in the first column. Our preferred model is column three, estimated through simulated maximum likelihood method. The results in Column 2, were obtained using a control function method and are in line with those obtained through the preferred method. Column 3 results indicate that increasing the weight of the striking vehicle by 1000 pounds increases the expected number of fatalities by a statistically significant factor of 1.643 ($t=7.1$). This is equivalent to an increase of 64.3% over our mean number of fatalities. Increasing the weight of the struck vehicle

³¹ Table 2 presents the results of the Negative Binomial regression estimates of the number of fatalities in the struck vehicle against covariates of interest. The values for the marginal effects are similar and consistent to those obtained by Anderson and Auffhammer (2013). The results indicate that increasing the weight of the striking vehicle increases the expected number of fatalities by about 40-45%.

reduces the expected number of fatalities in the struck vehicle by a mere 21.8%. These results, though similar to those obtained by Anderson and Auffhammer (2013) and Jacobsen (2013), with the results from the control function and SML approaches being larger. The first column results are similar to those obtained by Anderson and Auffhammer (2013) and Jacobsen (2013). The coefficient on curb weight of striking vehicle implies that a 1000 pound increase in its weight, increases the expected number of fatalities in the struck vehicle by a statistically significant factor of 1.42 with t value of 5.25. Across all the specifications, increasing the weight of the striking vehicle by 1000 pounds increases the expected number of fatalities in the struck vehicle by about 50-64%, while increasing the curb weight of the struck vehicle by the same amount results in a decrease in the expected number of fatalities by about 21-26%.

Serious Injury in Two-Vehicle collisions

Table 4 reports the results from second stage regression estimates as well as the regression without selection for comparison. The results reported are for only SML approach. The dependent variable in this case is the number of seriously injured occupants in the struck vehicle. Similarly we conduct Wald and likelihood ratio tests for the presence of selection effects and we reject the null hypothesis that ρ equals zero. Based on these tests, there does appear to be selection problem and the coefficients do change somewhat in the conditional model.

The predicted sign for the struck vehicles are negative again since increasing the weight of the struck vehicle will reduce the incidence of serious injuries in the vehicle. Similarly, the sign of weight for the striking vehicle is positive. Increasing the

weight of the striking by a 1000 pound will lead to an increase in the expected number of seriously injured occupants by a statistically significant factor of 1.23. This is equivalent to a 23% increase in the expected number of serious injuries in the struck vehicle. While increasing the weight of the struck vehicle will decrease the expected number of occupant injuries by 0.84 with a t-value of 8.28. This is equivalent to a decrease in the expected number of seriously injured occupants by about 18%.

Alternative Specifications

Table 5 presents the results of alternative specification in which we replace the weight of the two vehicles with a new explanatory variable that depicts difference in weights. The difference in weights is defined as ratio of weight of the struck vehicle to that of the striking vehicle. The coefficient of the ratio of weights is more difficult to interpret. The coefficient is negative and significant at the 1% level. The coefficient of mass ratio is -1.460 with a t value of 7.87.

Moral Hazard and Falsification Tests

We will discuss the moral hazard issues and falsification (Keuzenkamp and Magnus 1995; Buck and Lady 2005) together because the issue of identification of the parameters is strongly linked to the assemblage of data and the hypothesis to be tested: existence or absence of moral hazard. We employ a largely qualitative procedure to test for existence of moral hazard presented by the econometric model. The emphasis here is placed less on the quantitative measure, but rather on the sign partials of the risk factor (weight) in the two equations. This approach allows the hypothesis to be falsified regardless of identification. The approach to estimation of the model parameters provide us with a unique opportunity to in-principle test for behavioral

adaptation does take place. The intuition behind the test is that if weight has opposite signed partials in the two equations (probability of accident and probability of injury given an accident), it is reasonable to infer that behavioral adaptation will take place. Since weight reduces the probability of occupant injuries, then if behavioral adaptation occurs, then the accident frequency will increase. Therefore, the differently signed partial effects of weight of the struck vehicle in the two stages of the estimation procedure (probability of accident and probability of occupant injury given that an accident has occurred) is a manifestation of moral hazard effects.

This reasoning is in line with Fridstrom 1999 and Peltzman 1975, who hypothesized that whenever an accident risk factor has differently signed partial effects in the accident frequency and occupant injury severity equations, that risk factor is subject to risk compensation. These two tests provide weak support for our selection on moral hazard hypothesis.

Next we discuss results of the falsification tests. If drivers of heavier vehicles do not involve in more fatal accidents, then one should not see any additional effect of increased weight on injuries and fatalities on motorcyclists, bicyclist and pedestrian. If we do find a statistically significant coefficient, then this is an indication that drivers do indulge in Peltzman type behavioral response, by exposing non occupants of vehicles to higher risk. To do this, we estimate another model for crashes between a vehicle and pedestrian, bicyclist and compare the coefficients. The results for the model are reported in Table 7. The coefficient of increasing the weight of vehicle is highly significant with a t-value of *10.23*

Policy Simulations

The conceptual framework and econometric specification is able to predict qualitatively and quantitatively how the weight of a vehicle affects the fatalities and injuries in two vehicle crashes. Using the estimates from the model, one can simulate the responsiveness to a policy change that results in the change in weight distribution of the fleet. That is we can estimate the net effects of the safety gains to occupants of the heavier vehicle and safety losses that the increased weight imposes on the occupants of the struck vehicle. The model allows us to simulate how individual risks changes if they switch from driving a heavier vehicle to a lighter vehicle. For example, if a policy creates the incentives for households to switch to a heavier vehicle then weighted average risk of the heavier drivers will change because their probability of being involved in accidents would change to that of the heavier drivers. The intuition being that factors such age and number of people in a household can interact with vehicle choices.

Expected Externality Cost

The econometric results suggest that the expected accident cost is affected by both the probability of being the heavier vehicle in a collision and the expected weight difference in a two vehicle collision. Increasing the weight of vehicle increases both of these. To explore the causal effect of weight on fatalities and serious injuries as well as their distribution, we randomly draw a vehicle with weight m_i and estimate the marginal effect of adding 1000 pounds of weight to the striking vehicle. Then estimate the probability of it being the heavier /lighter vehicle. From equation 4, we see that the net costs are proportional to

$$\varphi \left[\frac{\partial \lambda_j}{\partial m_i} l + \frac{\partial \lambda_j}{\partial m_i} (\cdot) \right]$$

and is a function of the weight of the two vehicles. However, the results show that the net costs depend on whether the vehicle is the heavier or lighter one in a two vehicle collision. The probability of being the heavier or lighter vehicle is computed as the proportion of all the vehicles in the sample that are heavier or lighter than the randomly selected vehicle. The net cost is estimated as the expected cost of the vehicle being the lighter one in a collision and of it being the heavier vehicle (Hultkrantz and Lindberg 2011). The external cost is defined as:

$$External\ Cost = VSL * \left[\left(P(m_i < m_j) \left(\frac{\partial \lambda}{\partial m_i} (\cdot) \right) \right) \Big|_{m_i < m_j} \right] + \left[\left(p(m_i > m_j) \left(\frac{\partial \lambda}{\partial m_i} \right) \right) \Big|_{m_i > m_j} \right]$$

Using the actual weight data from accident reports, we compute the estimated change in fatalities and serious injuries given the change in the weight distribution of the US light duty vehicle fleet. The data on US fleet was obtained from sales data from Good Car & Bad Car tracking of sales data and the distribution of fleet weight from EPA's Light-Duty automotive technology fuel economy trends from 1975 to 2013. The marginal effect values are estimated through iteration using the actual values for each observation in the sample. The first iteration uses the first observation; the second iteration used the second observation, etc. The average of all the iterations was then taken. We performed this twice, first using the vehicle as the lighter vehicle and then again as the heavier vehicle. The values of the marginal causal effect of the striking vehicle are 0.0011 and 0.0015 respectively. Using the procedure disclosed earlier, we estimate the probability of accident if the vehicle was the lighter vehicle at 0.001564 and 0.0702 if the vehicle was the heavier vehicle. Using a statistical value of

life of \$5.5Million in 2008 dollars (DOT 2013³²), we obtain accident net costs of increasing the fleet by 500³³ pounds of \$40.26 billion related to fatalities and \$10.13 billion related injuries (statistical value of serious injury used is \$577,500³⁴). The total externalities due to serious injury and death are \$50.38 billion. These results are presented in Table 8.

External Costs due to Gasoline

We have estimated that the expected external cost of accident imposed on lighter vehicles is about \$44.6 billion. The heavier vehicles also consume more fuel due to their larger weights. If we take the 2008 model year vehicle, a 500 pound increase in fleet weight is about 10.35%. Klier (2010) finds that the elasticity of fuel economy with respect to weight is about 0.67, translating to 6.93% lower fuel economy estimates in these vehicles. And given the long run fuel economy estimates for VMT of about -0.25, the effect on heavier vehicles is about 6.45% increases in gasoline consumption. We also assume 2008 model year fuel economy of 23.9, annual VMT of 15,000 miles, the annual fuel consumption of 628 gallons, average 2008 gasoline price of \$2.5, a social cost of carbon of \$20/ton, annual emissions per gallon of gasoline of 8,887gCO₂/gallon. The additional costs due to increased gasoline consumption and emissions are \$1.14 billion and \$0.1 billion respectively for the whole US fleet.

³² http://www.dot.gov/sites/dot.gov/files/docs/VSL%20Guidance_2013.pdf

³³ From data on US fleet trend, there has been a 500 pound increase US fleet over the same period to 2005.

³⁴ The statistical value of serious injury is obtained by multiplying the Fraction of VSL for a serious injury (0.105)

Policy Approaches to External Accident Costs

Several policy approaches have been implemented to reduce the accident costs. One approach is to tax drivers based on odometer readings, or through some other technological systems that are attached to vehicles. An optimal policy would involve per mile fees that differ on the basis of driver risk and vehicle attributes. Some other policies could also be gasoline tax, etc.

Vehicle miles traveled (VMT) fees

A VMT or per mile fee charges drivers based on the distance they drive.

Unlike the gasoline tax, the vmt fee is a true usage based fee because a driver is charged on actual use of the roadway. VMT fees generate more and stable revenue than a gasoline tax because under a gasoline tax, more fuel efficient vehicles will consume less gas for every distance covered than less fuel efficient vehicles. Under VMT fees, technology improvements might not affect the distribution of fleet weight or improve fleet fuel economy because it charges all vehicles the same flat rate. The fees could be uniform or differentiated based on vehicle class. A uniform vmt charge scheme is not efficient because it tends to levy too low a charge on vehicles with high external costs (leads in low welfare gains) and levy too high a charge on vehicles with low external costs (high welfare gains). One way to overcome this is to design a differentiated vmt fee.

Under a differentiated mileage charge, a per-mile charge based on vehicle type is levied on each vehicle class. The resulting effects of charge in mileage is a function of own price-effect and cross-price effect from increases in other vehicle classes. The potential welfare gains from a differential mileage charge is lower than a gasoline tax because some of the reduced demand from one vehicle class will be offset from the

increased demand from another class (Parry 2003). Preliminary evidence from mileage fee initiatives in the United States and other countries recommend several ways to gather mileage data and charge fees. Significant privacy-related concerns from the public and cost challenges have been raised in applying mileage fees to passenger vehicles. Other challenges include how drivers can contest charges, substantial start-up and administrative costs.

Gasoline Tax

The government can also impose a per unit tax on the amount of gasoline consumed by drivers. The intuition is that increased price will reduce the amount of driving which in turn leads to aggregate reduction in external costs. How external costs changes with improvements in fuel economy due to improved technology is not well understood. For example, downweighting could improve fuel economy but may make occupants less safe; although drivers might internalize this risk when making purchase decisions.

Cost of administration of a gasoline tax is very low compared to mileage based charge or fee. However, the costs associated with a gasoline tax are higher than an equivalently scaled mileage tax because of the larger substitutions effect from the gasoline tax. These costs are represented by reductions in consumer surplus under the aggregate demand curve for gasoline. Also, under the gasoline tax system, the per-mile charge differs across vehicles in proportion to their mpg ratings. Because of the positive correlation between gasoline per mile (from weight mpg relationship) and external costs of accidents, we expect gasoline tax reduce more external costs than an equivalently scaled uniform vmt fee system (Parry 2003).

Weight Based Tax through a “Pay at the Pump” Tax

Since the externality costs vary based on the weight/type of the vehicle, one way to internalize the externality is through a differentiated weight based tax. The tax could be levied *ex ante* or *ex post* as a liability. One approach for implementing the weight based tax *ex ante* could be through increased premium on liability insurance. Since several states require drivers to obtain liability insurance, it would be easy to increase the existing premiums to account for weight difference. Because insurance already charge higher premiums for riskier drivers, it would be possible to charge heavier vehicles more because they are more likely to be involved in more accidents and impose more risks on others. This weight based premium could be implemented in such a way that it interacts with factors such as driving record through a multiplicative factor as opposed to a fixed dollar increase related to weight. Charging heavier vehicles higher premiums based on their weight could encourage drivers of such vehicles to adjust their risk levels. The downside to the weight based tax is that some states allow drivers to carry no insurance at all and others to only carry up to \$10,000 coverage (with most common coverage amount being \$25,000). Given the existence of such limited liability insurance contracts, there are few incentives for insurance companies to set higher insurance premiums for heavier vehicles. Another issue is the large percentage of uninsured drivers in some states³⁵.

In place of gasoline taxes, or fixed mileage charges, drivers could be charged a “Pay at the Pump” tax that accounts for both the weight of the vehicle and fuel consumption. The “Pay at the Pump” tax (Khazoom1997) involves bundling the additional insurance

³⁵ As much as 30% of uninsured drivers exist in some states (White 2004)

premium with gasoline purchases. “Pay at the Pump” levies a surcharge that reflects the externality per mile for each driver. Adding the surcharge will increase the gasoline price, which in turn may lead to some reduction in travel demand based on elasticity of driving to gasoline price changes. The “Pay at the Pump” policy ensures that the VMT will equal the optimal level and remove the distortions in resource allocation. Because the marginal cost facing the driver incorporates the full marginal exposure to risk. The “Pay at the Pump” surcharge that internalizes the total externality amounts to 34 c/gallons (\$46.13 Billion/137.5 billion gallons of gasoline). The advantages of the “Pay at the Pump” tax are that it ultimately reduces the externality cost by removing uninsured motorists from the road. In addition, in the short term it provides incentives for drivers to switch to more efficient vehicles because the resulting increase in gasoline price due to the surcharge induces driver to switch to more fuel efficient vehicles will pay less than her marginal exposure risks. The converse is true for a low fuel efficient vehicle. Over time, the increase gasoline price (arising from surcharge) induces a demand for newer, lighter, more efficient vehicles by owners of heavier less efficient vehicles.

Conclusion

Following the introduction of weight indexed fuel efficiency standards which have resulted in an increase in the weight of the vehicle fleet; this study examines the effect on societal safety of consumers demand for heavier vehicles. We present empirical estimates of the effect of vehicle heavier fleet on accident costs that adequately accounts for selection bias due to moral hazard issues. We capture unobserved driver behavior as correlated outcomes between the two dimensions of

vehicle safety. We find that increasing the weight of heavier vehicles increases the expected of occupant fatalities by 50-64% while decreasing expected number of fatalities by 21-26%. Based on our estimates, we find that not accounting for selection bias due to moral hazard problems results in estimates of the effect of weight externality that are downwardly biased by as much as 14%.

We first develop a theoretical framework model of individual driver choice for vehicle weight and risk aversion behavior that leads to Nash equilibrium with a predominantly heavier fleet. Then we employ data on fatal and non-fatal injuries to estimate the external costs of additional weight in two vehicle crashes. Though difficult to estimate empirically, we apply an estimator with a lognormally distributed random effect for bivariate count processes to account for selection bias due to moral hazard. In addition, we use two different testing schemes to confirm the existence of selection bias due to moral hazard problems.

First, we find that increasing the weight of a vehicle will increase the probability being involved in an accident. This is because, increasing the weight of a vehicle will increase the probability of it being the heavier vehicle in an accident, offering better protection, so the driver will reduce care level³⁶. We also find that factors which reduce the probability of having severe injuries in two vehicle crashes induce them to be involved in more collisions weakly confirming the presence of Peltzman type behavioral responses. When we apply the 2008 statistical values of life values to our estimates, we find that the external costs in 2008 dollars are \$51.84 billion. These results are obtained under the assumption of certain risk behavior by

³⁶ Also, the increased probability of heavier vehicles being involved in accidents could be explained by their physical characteristics and the potential response of drivers of heavier vehicles.

drivers and the results could vary under different assumptions about the risk attitudes of drivers. These costs are borne by the society in the form of medical bills, lost earnings, social security payments, etc.

These values have serious policy implications. In theory, an efficient fleet could be achieved through a carefully calibrated tax based on careful cost-benefit analysis. However, such a tax policy may pose practical problems of implementation in terms of public acceptance and monitoring. Given that the societal costs of traffic accidents vary by weight, we also find that implementing a weight based tax such as a “Pay as You Pump” surcharge of the 38c/gallon could fully internalize the costs. A policy such as the proposed one will also lead to the adoption of more fuel efficient vehicles over time and reduction in the weight of the fleet.

REFERENCES

- Anderson, Michael. "Safety for Whom? The Effects of Light Trucks on Traffic Fatalities." *Journal of Health Economics* 27, no. 4 (July 2008): 973–989. doi:10.1016/j.jhealeco.2008.02.001.
- Anderson, Michael, and Maximilian Auffhammer. *Pounds That Kill: The External Costs of Vehicle Weight*. Working Paper. National Bureau of Economic Research, June 2011. <http://www.nber.org/papers/w17170>.
- Anderson, Soren T., Ian WH Parry, James M. Sallee, and Carolyn Fischer. "Automobile Fuel Economy Standards: Impacts, Efficiency, and Alternatives." *Review of Environmental Economics and Policy* 5, no. 1 (2011): 89–108.
- Ashenfelter, Orley, and Michael Greenstone. *Using Mandated Speed Limits to Measure the Value of a Statistical Life*. Working Paper. National Bureau of Economic Research, August 2002. <http://www.nber.org/papers/w9094>.
- Avery, Christopher, S. Jody Heymann, and Richard Zeckhauser. "Risks to Selves, Risks to Others." *American Economic Review* 85, no. 2 (1995): 61–66.
- Blomquist, Glenn C. "Self-Protection and Averting Behavior, Values of Statistical Lives, and Benefit Cost Analysis of Environmental Policy." *Review of Economics of the Household* 2, no. 1 (2004): 89–110.
- . "Values of Risk Reduction Implied by Motorist Use of Protection Equipment: New Evidence from Different Populations," 1996.
- Cameron, A. Colin, and Per Johansson. "Count Data Regression Using Series Expansions: With Applications." *Journal of Applied Econometrics* 12, no. 3 (1997): 203–223. doi:10.1002/(SICI)1099-1255(199705)12:3<203::AID-JAE446>3.0.CO;2-2.
- Crandall, Robert W., and John D. Graham. "Effect of Fuel Economy Standards on Automobile Safety, The." *JL & Econ.* 32 (1989): 97.
- Dahl, Carol, and Thomas Sterner. "Analyzing Gasoline Demand Elasticities: A Survey." *Energy Economics* 13, no. 3 (July 1991): 203–210. doi:10.1016/0140-9883(91)90021-Q.
- Duncan, Chandler S., Asad J. Khattak, and Forrest M. Council. "Applying the Ordered Probit Model to Injury Severity in Truck-Passenger Car Rear-End Collisions." *Transportation Research Record: Journal of the Transportation Research Board* 1635, no. 1 (1998): 63–71.
- Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards*. Accessed March 13, 2014. http://www.nap.edu/catalog.php?record_id=10172.
- Elvik, R. "The External Costs of Traffic Injury: Definition, Estimation, and Possibilities for Internalization." *Accident; Analysis and Prevention* 26, no. 6 (December 1994): 719–732.
- Evans, L. "Causal Influence of Car Mass and Size on Driver Fatality Risk." *American Journal of Public Health* 91, no. 7 (July 2001): 1076–1081.
- Frondel, Manuel, Jörg Peters, and Colin Vance. *Identifying the Rebound - Evidence from a German Household Panel*. Ruhr Economic Paper. Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Ruhr-Universität Bochum, Universität Dortmund, Universität Duisburg-Essen, 2007. <http://ideas.repec.org/p/rwi/repape/0032.html>.
- Gayer, Ted. "The Fatality Risks of Sport-Utility Vehicles, Vans, and Pickups Relative to Cars." *Journal of Risk and Uncertainty* 28, no. 2 (2004): 103–133.

- Greene, David L. "Feebates, Footprints and Highway Safety." *Transportation Research Part D: Transport and Environment* 14, no. 6 (2009): 375–384.
- Greene, William. *A General Approach to Incorporating Selectivity in a Model*. Working Paper. New York University, Leonard N. Stern School of Business, Department of Economics, 2006. <http://ideas.repec.org/p/ste/nystbu/06-10.html>.
- . *Correlation in Bivariate Poisson Regression Model*. Working Paper. New York University, Leonard N. Stern School of Business, Department of Economics, 2007. <http://ideas.repec.org/p/ste/nystbu/07-15.html>.
- . *Functional Form and Heterogeneity in Models for Count Data*. Working Paper. New York University, Leonard N. Stern School of Business, Department of Economics, 2007. <http://ideas.repec.org/p/ste/nystbu/07-10.html>.
- Gurmu, Shiferaw, Paul Rilstone, and Steven Stern. "Semiparametric Estimation of Count Regression Models." *Journal of Econometrics* 88, no. 1 (January 1999): 123–150. doi:10.1016/S0304-4076(98)00026-8.
- Hultkrantz, Lars, and Gunnar Lindberg. *Accident Cost, Speed and Vehicle Mass Externalities, and Insurance*. International Transport Forum Discussion Paper. OECD Publishing, 2011. <http://ideas.repec.org/p/oec/itfaab/2011-26-en.html>.
- Jacobsen, Mark R. "Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior." *American Economic Journal: Applied Economics* 5, no. 3 (July 2013): 1–26. doi:10.1257/app.5.3.1.
- Jacobsen, Mark R. "Fuel Economy, Car Class Mix, and Safety." *American Economic Review* 101, no. 3 (2011): 105.
- Klier, Thomas, and Joshua Linn. "New-Vehicle Characteristics and the Cost of the Corporate Average Fuel Economy Standard." *The RAND Journal of Economics* 43, no. 1 (2012): 186–213.
- Keuzenpamp, Hugo and Jan Magnus, "On Tests and Significance in Econometrics," *Journal of Econometrics*, 67, (1995): 5-25.
- Buck, Andrew and George Lady, "Falsifying Economic Models," *Economic Modelling*, 22, (2005): 777-810.
- Li, Shanjun. "Traffic Safety and Vehicle Choice: Quantifying the Effects of the 'arms Race' on American Roads." *Journal of Applied Econometrics* 27, no. 1 (2012): 34–62. doi:10.1002/jae.1161.
- Newbery, David M. "Road Damage Externalities and Road User Charges." *Econometrica* 56, no. 2 (March 1, 1988): 295–316. doi:10.2307/1911073.
- Noland, Robert B. "Fuel Economy and Traffic Fatalities: Multivariate Analysis of International Data." *Energy Policy* 33, no. 17 (2005): 2183–2190.
- . "Motor Vehicle Fuel Efficiency and Traffic Fatalities." *The Energy Journal* 25, no. 4 (January 1, 2004): 1–22.
- Park, Byung-Jung, and Dominique Lord. "Application of Finite Mixture Models for Vehicle Crash Data Analysis." *Accident Analysis & Prevention* 41, no. 4 (July 2009): 683–691. doi:10.1016/j.aap.2009.03.007.
- Parry, Ian W. H., Margaret Walls, and Winston Harrington. *Automobile Externalities and Policies*. Discussion Paper. Resources For the Future, 2007. <http://ideas.repec.org/p/rff/dpaper/dp-06-26.html>.
- Peltzman, Sam. "The Effects of Automobile Safety Regulation." *Journal of Political Economy* 83, no. 4 (August 1, 1975): 677–725.

- Reiss, Peter C., and Matthew W. White. "Household Electricity Demand, Revisited." *The Review of Economic Studies* 72, no. 3 (July 1, 2005): 853–883. doi:10.1111/0034-6527.00354.
- Small, Kenneth A., and Kurt Van Dender. "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect." *The Energy Journal* Volume 28, no. Number 1 (2007): 25–52.
- Terza, Joseph. "Parametric Nonlinear Regression with Endogenous Switching." *Econometric Reviews* 28, no. 6 (2009): 555–80.
- Terza, Joseph V. "Estimating Count Data Models with Endogenous Switching: Sample Selection and Endogenous Treatment Effects." *Journal of Econometrics* 84, no. 1 (May 1998): 129–54. doi:10.1016/S0304-4076(97)00082-1.
- Toy, Edmond L., and James K. Hammitt. "Safety Impacts of SUVs, Vans, and Pickup Trucks in Two-Vehicle Crashes." *Risk Analysis* 23, no. 4 (August 1, 2003): 641–50. doi:10.1111/1539-6924.00343.
- White, Michelle J. "The 'Arms Race' on American Roads: The Effect of Sport Utility Vehicles and Pickup Trucks on Traffic Safety." *Journal of Law and Economics* 47, no. 2 (2004): 333–55.
- Whitefoot, Kate S., and Steven J. Skerlos. "Design Incentives to Increase Vehicle Size Created from the US Footprint-Based Fuel Economy Standards." *Energy Policy* 41 (2012): 402–411.
- Wilhite, Harold, and Loren Lutzenhiser. "Social Loading and Sustainable Consumption." *Advances in Consumer Research* 26 (1999): 281–287.

Figures and Tables

Figure 1-A: Mean-Variance Relationship for Number of Fatal Crashes in Months by State

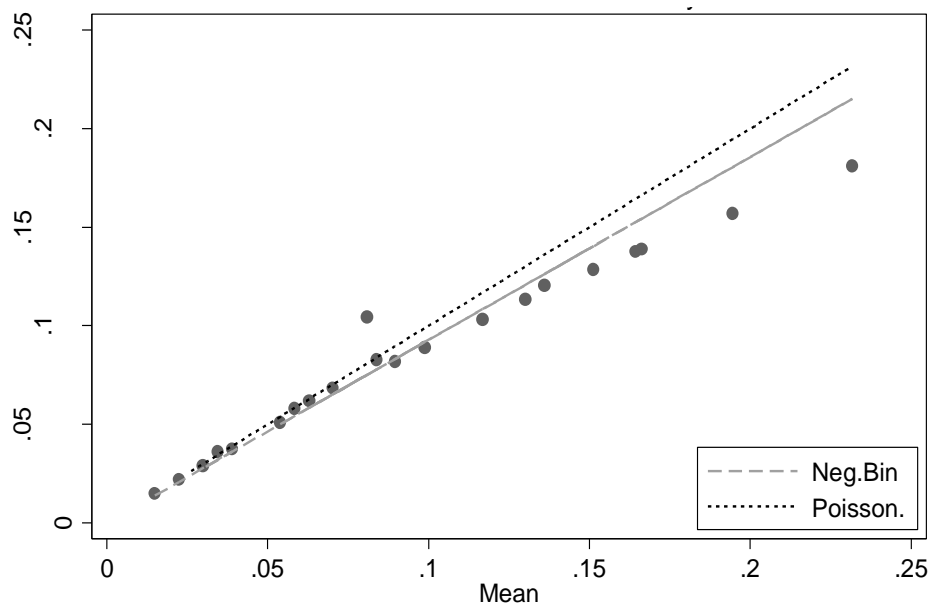
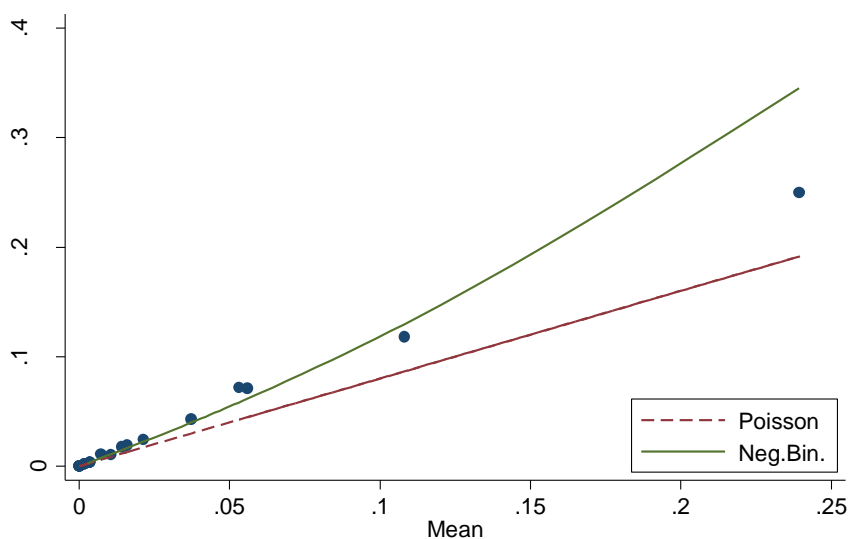


Figure 1-B: Mean Variance Relation for Number of Fatalities in Struck Vehicle



Notes: Figure 1A and 1B above are plots of variance against mean for two different models with exactly the same covariates. The figure was plotted by first creating 20 groups based on Poisson and negative binomial model and then computing the mean and variance for each group, and then finally plotting the mean-variance relationship.

Figure 2: Predicted and Observed probabilities of Fatal Accidents

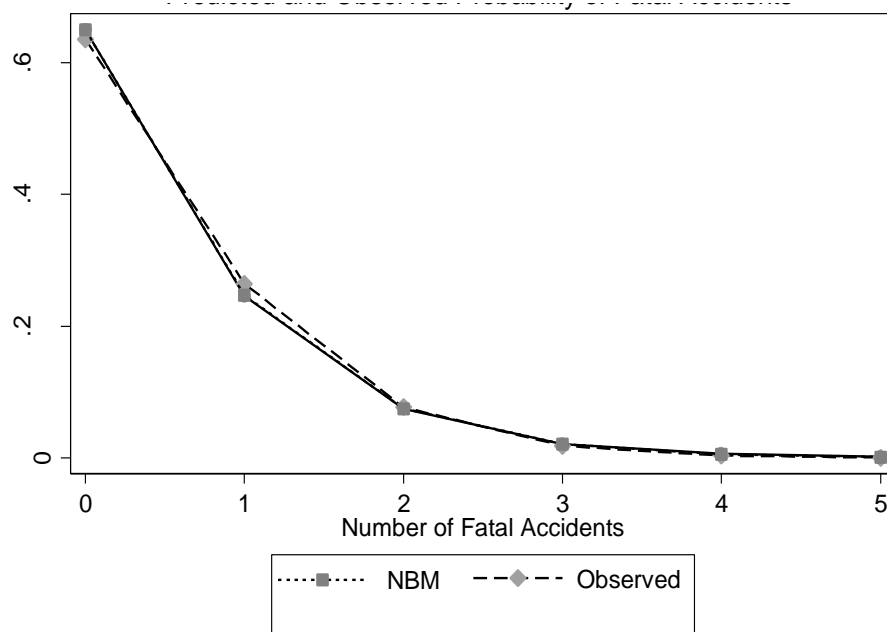


Figure 3: Predicted probability of Fatality in Struck Vehicle against Striking Vehicle Weight

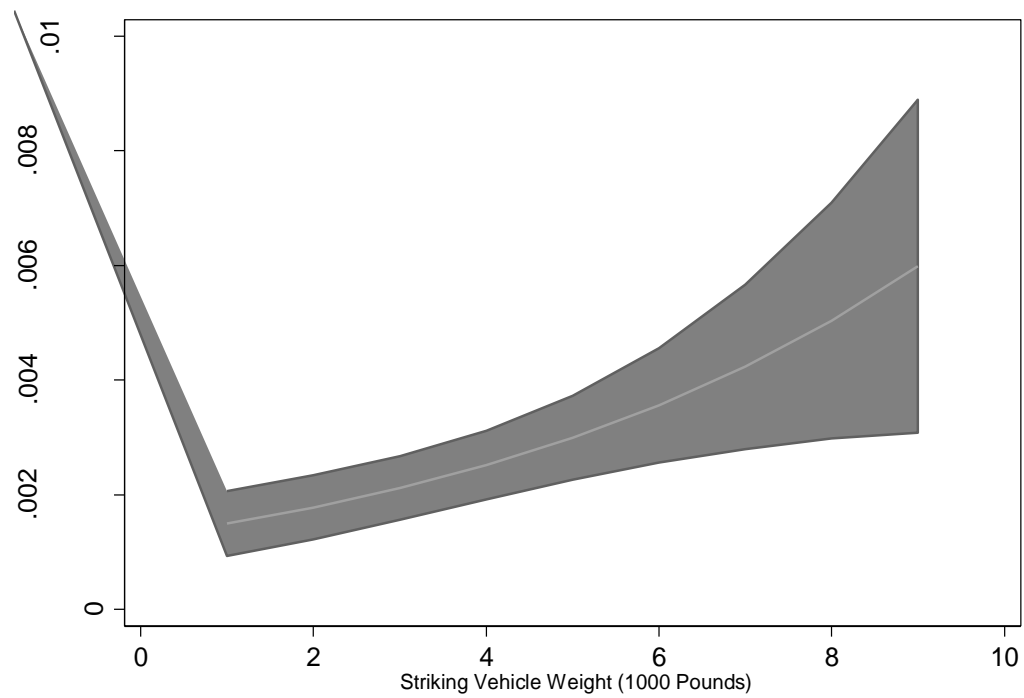


Table 1: Summary Statistics for Two Vehicle Collision

	Full Sample ^a			Regression Sample ^b		
	Sample Size	Mean	Std. Dev.	Sample Size	Mean	Std. Dev.
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Fatalities	265,633	0.18%	6.5	248,572	0.26	6.5
Number of Seriously Injured Occupants	265,633	5.36%	36.8	248,572	7.64%	36.7
Age	252,127	39.17	15.99	248,572	39.20	18.93
Sex	261,174	0.44	0.50	248,572	0.46	0.49
Alcohol Involved	265,633	6.01%	0.36	248,572	10.97%	0.31
Restraint Used	265,633	92.06%	1.55	248,572	93.27%	3.10
Number of Occupants	258,759	1.41	0.67	248,572	1.47	0.87
Model Year	265,633	2002	5.72	248,572	2000	5.12
Weight (lbs)	262,110	3388	551	248,572	3473	630.29
Light Trucks Share	258,130	29.9%	36.53	248,572	30.6%	36.9

Notes: ^aThe full sample contains observations that have information on the injury outcome of drivers' involved in crashes between 2008 and 2012 and the vehicle was manufactured on or after 1980.

^bThe regression sample contains observations from the full sample that does not have a missing value for any of the independent variables.

Table 2: Regression Estimates of Number Occupant Fatalities- No Selection

	Dependent Variables = Number of Fatalities in Struck vehicle		
	NBREG	NBREG	OLS
Weight of Struck Vehicle	-0.307*** (0.085) [-0.0009]	-0.337*** (0.070) [-0.0009]	-0.001*** (0.000)
Weight of Striking Vehicle	0.336*** (0.047) [0.0010]	0.338*** (0.065) [0.0011]	0.001*** (0.000)
Effect of 1000 Pounds increase in striking vehicle weight	40%	42%	40%
Driver Attributes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes
county Dummies	Yes	Yes	Yes
Model Year Dummies	No	Yes	Yes
Observations	248,572	248,572	248,572

Notes: Heteroskedasticity-Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The stars indicate the level of significance. Marginal Effects (evaluated at the representative values of the explanatory variables) for the negative binomial model are provided in brackets. We only report for the key independent variables: the weight of the vehicles involved in the crash. The weight effects of the striking vehicle are computed on the basis of a 1000 pound increase in weight over all observations in the model.

Table 3: First-Stage Regression Results

	Dependent Variables = P(Accident=1)	
	OLS	Probit
Natural Log of Total Vehicle Miles Travelled ^a	0.0011*** (0.0001)	0.0158*** (0.0772)
Percentage of Share of Light Duty Truck VMT	0.0206* (0.0103)	0.1719*** (0.0087)
Weight of Involved Cars per 1000 Pounds	0.0010** (0.0004)	0.0232** (0.0099)
Includes State Dummies	Yes	Yes
Includes Year Dummies	Yes	Yes
Year by State Dummies	Yes	Yes
Other Demographic Controls	Yes	Yes
log likelihood		-4,147
# Observations	333,763	333,763

Notes: Heteroskedasticity-Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The stars indicate the level of significance. The dependent variable is an indicator variable equal to one if the vehicle was involved in an accident in which there was serious injury or fatality.

a. The value reported is the difference between the coefficient of natural log of 100 Million vehicle miles travelled per state and that of traffic density.

Table 4: Second Stage Regression Results

Dependent Variables = Number of Fatalities in Struck Vehicle			
	No-Selection	CF	SML
Weight of Struck Vehicle	-0.307*** (0.085) [-0.0009]	-0.276*** (0.081) [-0.0008]	-0.246*** (0.089) [-0.006]
Weight of Striking Vehicle	0.336*** (0.064) [0.0011]	0.460*** (0.074) [0.0015]	0.497*** (0.070) [0.0017]
ρ		-0.344 (0.010)	-0.395 (0.011)
Effect of 1000 Pounds increase in striking vehicle weight	42%	57.9%	64.37%
Effect of 1000 Pounds increase in struck vehicle weight	-26.4%	-23.9%	-21.8%
#Observations	240,105	240,103	240,013

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The stars indicate the level of significance. Estimates are for two-vehicle collisions from 2008 to 2012. The results displayed are estimated after accounting for selection. The key explanatory variable is the weight of the vehicle. Column CF uses the control function methodology to correct for the selection while the SML column is the preferred model and is described in the paper.

Table 5: Second Stage Regression Estimates for Serious Injury Model

Dependent Variable = Number of Seriously Injured Occupants in Struck Vehicle		
	No Selection	Selection
Weight of Struck Vehicle	-0.171*** (0.016) [-0.096]	-0.165*** (0.0192) [0.028]
Weight of Striking Vehicle	0.163*** (0.012) [0.012]	0.202*** (0.0513) [0.0317]
ρ	0	0.5622 (0.0190)
Effect of 1000 Pounds increase in striking vehicle weight	19%	23%
Effect of 1000 Pounds increase in Struck vehicle weight	17%	
# Observations	248,857	248,003

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The stars indicate the level of significance. The marginal effect of a 1000 pound increase in weight of striking vehicle are in brackets. Estimates are for two-vehicle collisions from 2003 to 2012.

Table 6: Alternative Specification

Dependent Variable = Number of Fatalities in Struck Vehicle (7)	
Mass Ratio Struck Vehicle to Striking Vehicle	-1.460*** (0.187)
Driver Attributes	Yes
Weather Attributes	Yes
Air Bags	Yes
Dummy for Alcohol Use	Yes
Year Dummies	Yes
City Dummies	Yes
Model Year Dummies	Yes
#Observations	247,222

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8 Falsification Test

Dependent Variable = Number of Pedestrian/Motorcyclists/Bicyclists Fatalities	
Weight of Vehicle	0.170*** (0.0165)
Driver Attributes	Yes
Weather Attributes	Yes
Air Bags	Yes
Dummy for Alcohol Use	Yes
Year Dummies	Yes
City Dummies	Yes
Model Year Dummies	Yes
#Observations	10,910

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Stars indicate level of significance

Table 8: Accident-Related Externalities

	Total External Costs/Benefits (Billions)	External Costs per (\$/gallon)
<u>Traffic Related External Costs</u>		
Fatalities Only	40.26	0.29
Serious Injury Only	10.13	0.07
Fatalities + Serious Injuries	50.38	0.37
<u>Gasoline Related Externalities</u>		
Fuel Savings	1.36	
Reduced Emissions	0.10	
Fuel Savings + Reduced Emissions	1.46	
<u>Total Externalities</u>	<u>51.84</u>	<u>0.38</u>

Notes: The externality cost of traffic accident is estimated under the assumptions that increasing the mass of a vehicle increases the probability to be the heavier vehicle in a collision and the expected mass difference given a collision with a lighter vehicle. The expected externality is computed as the sum of the expected cost given a collision with a heavier vehicle and expected cost given a collision with a lighter vehicle. The externality computation uses the Department of Transport's guidance on statistical value of life (VSL) value of \$5.5 Million and statistical injury of \$577,500 (2005) updated to 2008 dollar values.

MANUSCRIPT 3
Heterogeneity and Sorting in Consumer Valuation of Energy Efficiency
Investments: Empirical Evidence from US Automobile Market

By

Edson Okwelum and Jim Opaluch
Department of Environmental and Natural Resource Economics
University of Rhode Island

To be submitted to *Journal of Econometric Applications*

Heterogeneity and Sorting in Consumer Valuation of Energy Efficiency Investments: Empirical Evidence from US Automobile Market

Edson Okwelum and Jim Opaluch
Department of Environmental and Natural Resource Economics
University of Rhode Island

July 31, 2014

Abstract

If consumers are heterogeneous in their preference for vehicle attributes and rates of time preference, they will sort into different vehicles based on fuel economy. Bento *et al*(2012) suggests that estimates of MWTP for future fuel costs can be distorted if individuals choose vehicles according to idiosyncratic tastes and rates of time preferences. This paper describes how sorting resulting from individual unobserved heterogeneity and time preference biases estimates of MWTP for future gasoline costs and implement estimation strategies that correct such bias.

Using data from consumer expenditure survey, we recover MWTP estimates that do not show any evidence of consumer undervaluation of MWTP for future gasoline costs for the average consumer. We recover each consumer's valuation of fuel economy as well as the distribution of estimates of value of fuel economy in the population using a model that integrates vehicle choice with driving intensity. Given consumer's expectations of future gasoline price and tastes, they first choose a vehicle and then conditional on vehicle stock; dynamically adjust their driving habits in response to gasoline prices. We use an estimation strategy that accounts for both the likely correlations in error terms between the discrete and continuous margins as well as cross-equation restrictions imposed by the integrated behavioral model.

We find that consumers are heterogeneous in the valuation of fuel economy and uncover three segments based on their valuation of fuel economy. One group, those with environmentally driven preference pays significantly more than \$1 per dollar gas savings and uses their vehicles frugally (about 18%). A second group (68%) only looks it as a purely financial decision and rationally values fuel economy. This group roughly pays \$1 per dollar gas savings. A third group, which one can consider to be myopic, are willing to pay less than \$1 per dollar savings.

Our results suggest that gasoline taxes would be a more efficient instrument to increase fleet wide fuel economy.

Keywords: *Fuel Economy, Heterogeneity, Sorting, Gasoline Prices, MWTP*

*Email: (Okwelum) edson_okwelum@my.uri.edu. I graciously thank Jim Opaluch and John Burkett, at the University of Rhode Island for providing useful insights and comments about modeling heterogeneity and endogeneity. All errors are mine alone.

Introduction

In the last few decades, the U.S. government has introduced several taxes and subsidies on the purchase of new vehicles that depend on the performance of vehicle fuel economy. The main driver for taxation of fuel economy is to correct externalities that result from fuel consumption. Cost-benefits analyses of these regulations require precise estimates of the consumer marginal willingness to pay (MWTP) for reduced discounted future fuel costs (Parry et al 2007). While the literature on how consumer value fuel economy is well established (Dreyfus and Viscusi 1995, Hausman 1979 Busse et al 2013, Alcott and Wozny 2013), a review of the literature on consumer valuation of fuel economy is inconclusive (Greene 2010). Some of the studies find that consumers overvalue fuel economy while others find undervaluation of fuel economy.

Several reasons have been given for the apparent undervaluation of fuel economy. (Helfand and Wolverton 2009).³⁷ The consumer heterogeneity argument posits that if consumers have heterogeneous preferences, they will sort into different groups (Deleire *et al* 2013, Bayer *et al* 2011). In a differentiated product market such as automobiles, we expect that consumers with similar preferences for certain attributes will make similar choices if they have similar preferences and face similar constraints. In terms of fuel economy, consumers with high MWTP for discounted future gasoline costs will sort into fuel efficient vehicles while those with low MWTP will sort into inefficient vehicles. This sorting bias has implications for the ability of econometric models to recover the value of fuel economy and other attributes.

³⁷ One hypothesis is that consumers' decision on vehicle type is based solely on financial considerations. However, consumers might have other fundamental values for fuel economy, including concern for the environment.

However, the sorting process also reveals information about consumer heterogeneity that can be used to evaluate how consumers value fuel economy.

In this paper, we contribute to the literature on *Energy Paradox*³⁸ by providing empirical evidence that unobserved consumer heterogeneity could result in different consumers sorting into different vehicle types. This implies that the coefficient on fuel economy also reflects values on other attributes that are associated with fuel economy. Unlike existing literature in this topic, this paper accounts for sorting bias due to unobserved heterogeneity by using an “equilibrium sorting”³⁹ model that uses a mixture of distributions to characterize unobserved and observed heterogeneity among households (Train and Winston 2007). The “equilibrium sorting” model uses the properties of market equilibria, as well as information on the behavior of economic agents, to infer structural parameters that characterize agent heterogeneity. We pair the random coefficient discrete choice model with a supply side model in which firms compete in Bertrand Nash fashion with differentiated products; where price depends on both the elasticity of demand and costs of production. With the random utility model, we allow annual vehicle miles travelled, time preferences and expectations of gasoline prices to vary across consumers. We treat the discounted operating costs and vehicle costs as random variables. This allows us to obtain a distribution of households’ preferences for fuel economy across the population.

³⁸ Energy Paradox is defined as the disconnect between net present value estimates of energy conserving cost savings and what consumers actually pay on energy conservation (Metcalf and Hasset 1999, Jaffe et al 2001).

³⁹ The paper uses a straight forward extension of the framework used by Bayer and Timmins (2003) to examine sorting in consumer choice of vehicle type.

Recent works on how consumers value fuel economy tend to use monthly within vehicle variation over time in gasoline prices to identify consumer tradeoffs between fuel costs and vehicle costs. The question is that in the presence of sorting bias due to observed and unobserved consumer heterogeneity, whether this is an accurate way to identify tradeoffs between vehicle costs and fuel economy. Additionally, firms respond to gasoline prices in the short term by adjusting vehicle costs to match sales. A key aspect of this is that prices are negotiated by dealers, and will depend upon inventories. For example, if gas guzzlers are not selling, dealers will offer price discounts. When demand is high for vehicles with high mileage, discounts are lower, and dealers could charge a premium over the retail price. For example, MacManus (2005) finds that the shift to higher fuel efficiency vehicles brought about by rising gasoline prices were obscured by price cuts disproportionately aimed at gas guzzlers. Panel data and individual fixed effects provide potential solutions when very large data sets spanning long period are available. Chintanuga (2001) has shown that it is important to control for both price endogeneity and heterogeneity to avoid potential biases in demand side parameter estimates.

Therefore, we identify consumers weighting of discounted gasoline costs by using the combination of functional form assumptions supported with information on vehicle demand and elasticity extracted from observed vehicle transaction prices. We use data on household vehicle holdings from the automobile extract of the consumer expenditure survey covering the period 2002-2005. This dataset contains information on the vehicle purchases and driving habits of a random sample of US households. We have actual data on the number of miles travelled by each household and vehicle

type⁴⁰. The individual household data allows us to match vehicles with the household demographic characteristics of those who own them and estimate how different segments of the population value fuel economy. We exploit the unique attributes of the CEX dataset (it contains information on vehicle miles travelled, when consumers trade-in older vehicles and expenditure on fuel costs) to estimate a richer random coefficient model in which we are able to identify heterogeneity in the weighting of discounted future gasoline costs by different consumers. In addition, we also account for volatility in gasoline price by allowing households to continually update their expectations of gasoline price. In future time periods, households dynamically adjust their driving habits in response to gasoline price.

The consumer heterogeneity we find is substantial and significant. The heterogeneity arises from huge differences in the amount of miles travelled by consumers and heterogeneity in expectation of gasoline prices and time preferences. We find that a substantial portion (61%) of upper the 95% of households in our sample correctly value fuel economy as they are willing to pay \$0.99 to reduce obtain a \$1.00 discounted future gasoline costs over the lifetime of the vehicle. Another group, those with environmentally driven preference pays significantly more than \$1 per dollar gas savings and uses their vehicles frugally (about 28% of the upper 95% overvalue fuel economy as they are willing to pay an average of \$1.57). The lower 5% of our sample are willing to pay only \$0.50 for a \$1 reduction in future fuel costs. These results are in contrast to those reported elsewhere by Alcott and Wozny (2009)

⁴⁰ This is important because consumers will most likely adjust their mileage driven of different vehicles based on gasoline prices. At high prices, consumers are more likely to put more miles on the Prius than a Suburban.

but support results from the simulation study by Bento *et al* 2012 who find that not accounting for sorting bias and consumer heterogeneity could lead to a conclusion that consumers undervalue fuel economy. In that light, our results might not necessarily imply that consumers overvalue/undervalue fuel economy, but can be interpreted as being heterogeneous in their valuation of fuel economy.

How consumers weigh temporal effects of future fuel costs have important policy implications. This is because the nature of the temporal weighting helps in determining if market failure exists and also helps indicate whether policy prescriptions that affect initial vehicle costs such as gas guzzlers tax will reduce fuel consumption at lower costs than gasoline tax. In addition, results from such “equilibrium sorting” models can be used to develop theoretically consistent predictions for the welfare implications of future policy changes on fuel economy and gasoline taxes.

In the last few decades, the literature on how to estimate consumer’s quality-adjusted prices of automobiles have grown. Most of the literature in this area has mostly involved the relationship between fuel economy and fuel price and price-vehicle safety tradeoffs. However, the fuel economy–price tradeoff has received the most attention (Hausman 1979, Mannering and Winston 1985, Zauberman 2003, Greene 2010). And it is this part of the literature that this paper contributes. Estimates of how much consumers are willing to pay to reduce fuel costs vary greatly (Greene

2010). Of the 25 studies reviewed by Greene, about half find that consumers undervalue fuel economy while just five report values suggesting overvaluation.⁴¹

Recent studies, however, appear to agree that households either fully value or modestly undervalue energy costs when making car purchase decisions. Alcott and Wozny (forthcoming) and Busse *et al* 2013, report implied discount rates closer to borrowing rates and conclude that these are evidence of lack of consumer myopia. Alcott (2013) finds slight undervaluation in fuel economy. Sallee *et al* (2009) use micro data on used vehicle prices, odometer readings and gasoline prices to estimate the relationship between fuel costs and vehicle prices. In the same vein, Sawhill (2008) using ARIMA models of fuel price expectations, does not find any evidence that consumer's underweight future operating costs. Using simulated data, Bento *et al.* (2012) show that consumer heterogeneity can result in sorting leading to downward bias in estimates of the WTP for fuel economy savings.

This work is closely related to the work of Bento *et al* (2012), Sawhill (2008), Alcott and Wozny (2009) Bayer *et al* (2011). However, this paper extends the literature in several respects. Our work is different from Sawhill's in several respects⁴². The main difference between this work and these papers is on the identification strategy; explicitly accounting for heterogeneity and sorting; and including information about the supply side. First, we use individual household data why Sawhill used aggregate data. In addition, while we both control for price endogeneity (Sawhill uses BLP's contraction mapping method and does not include information about pricing behavior). Unlike Bento *et al* (2012), this work is an

⁴¹ From \$0.01 (Berry *et al* 1995) to \$1.00 (Dreyfus and Viscusi 1995).

⁴² Both papers use random coefficient logit models.

empirical study rather than a simulation and such provides empirical evidence of the effect of consumer heterogeneity in estimates of the consumer trade-off of fuel economy and vehicle costs.

This paper is structured as follows. Section two discusses the equilibrium model composed of a demand side and supply side model. Section three, present the data used for the analysis. Section Four discusses the identification and estimation strategy as well as the challenges in estimating the specification. This is followed by the results and discussion and finally, we conclude in section six.

Methodology and Conceptual Framework

The model is consistent with a structural model of equilibrium of a heterogeneous product competition. The approach is based on earlier models by Goldberg (1995) and Amil Petrin (2002), but with significant differences. The method aggregates individual preferences into market level demand system, and combines assumptions on cost functions and pricing behavior to generate equilibrium prices and quantities. The estimation strategy is divided into two different steps. In the first step, the paper estimates household-level demand functions and then aggregates these individual functions to construct estimated firm demand curves. In the second step, we use the estimated demand curves to solve firms' first order conditions under the assumptions of Bertrand –Nash competition. The demand side of the model is based on a random utility function of consumer vehicle choice following Berry, Levinsohn and Pakes (1995; henceforth, BLP). The demand specification is paired with a supply side.

Consumer Preferences for fuel economy

This paper treats the demand for a new vehicle as an intertemporal choice problem in which consumers' trade-off future fuel savings and vehicle price. Each household derives utility from both vehicle ownership and utilization. We derive and estimate a model of automobile demand that accounts for households' vehicle quantity, vehicle type (make, model and vintage) and the continuous choice of utilization.

Household i maximize utility U_{ijt} from choosing vehicle j or not (the outside good) in each choice occasions $t=1, \dots, T$. The household's joint utility from quantity and vehicle type j is additive and can be expressed with a random indirect utility given by:

$$U_{ijt} = V_{ijt} + \varepsilon_{ij} \quad (1)$$

$$V_{ijt} = -\alpha_i p_j + \gamma_{it} G_{ijt} + \beta_i x_{ij} + \sigma v_i \quad (2)$$

where α_i , β_i and γ_i are individual specific taste parameters. p_j is transaction price for vehicle j . We allow preferences over price to vary with income so that α_i represents the individual specific marginal utility of income. Specifying vehicle price this way (with income) ensures that the indirect utility function from which we obtain the demand for discrete alternatives satisfy Roy's Identity. This implies that within any household vehicle ownership level, the marginal rate of substitution between gasoline costs and vehicle costs is income independent (Mannering and Winston 1985). x_j , is the vector of observed vehicle attributes interacted with demographic variables (includes product intercept term). A random variable (σv_i) is included to account for correlation in unobserved utility over the choice alternatives (Brownstone and Train

1999).⁴³ By specifying the x_{ij} appropriately, you can construct any pattern of covariance across alternatives. G_{ijt} is the expected future fuel cost over the vehicle lifetime is represented by:

$$G_{ijt} = E_t \left[\sum_{s=0}^{T_j} (1+r_i)^{-s} g_{t+s} m_{ij,t+s} 1/mpg_j \right] \quad (3)$$

where T is the vehicle expected lifetime for model j , r is household specific discount rate, g_{t+s} is the expected gasoline price at time $t+s$, $m_{ij,t+s}$ is the annual vehicle miles travelled in year $t+s$ by vehicle j for individual i . mpg_j is the fuel consumption rate in miles per gallon. In this model, heterogeneity derives from different consumers having discount rates and annual miles travelled. The model therefore explicitly accounts for consumer heterogeneity in their valuation of fuel economy. $\varepsilon_{ij} = \eta \xi_j + \epsilon_{ij}$. η is a vector of parameters associated controls (Kim and Petrin 2010), ϵ_{ij} is iid extreme value. ξ is the average utility of consumers from unobserved attributes of product j independent and idiosyncratic utility an individual obtains from the vehicle (however it could be correlated with the unobserved product attributes) and it is iid extreme.

The terms in $\eta \xi_j$ represent unobserved utility component that induces correlation as well as substitution in vehicle to overcome the independent of irrelevant alternative problem (IIA) (Train 2003). The implication of the expectation term is that when consumers' tradeoff of future fuel savings against vehicle prices, they consider how fuel efficient the vehicle is, the number of miles to drive, what the price of gasoline will be in the future and other vehicle characteristics such as horsepower, environmental factors, etc.

⁴³ The coefficient σ , is the standard deviation of the random coefficient reflecting the degree of correlation among the choice alternatives and v_i is the iid standard normal deviate.

The specification of the expectation term allows mileage to vary with customer and gasoline prices. Conventionally testing for whether consumers undervalue fuel economy is a test of the null hypothesis of $\alpha=\gamma^{44}$ (Alcott and Wozny 2009, Anderson *et al* 2013; Bento *et al* 2012). A major assumption that these papers and others have made is that gasoline prices follow a random walk model, that is consumers' future gasoline prices expectation at time $t+s$, (g_{t+s}) is equivalent to the current gasoline price. This assumption allows the expectation term to be reduced to

$$\frac{1+r}{r} \left(1 - (1+r)^{-T}\right) \frac{m_{ij} * g_t}{mpg_j} \quad (4)$$

While most authors assume that consumers' expectations of future gasoline prices follow a random walk model, some authors have used different assumptions about future expectations of gasoline price. Alcott and Wozny 2009 assume that future expected gasoline prices is a function of the future price of crude in the commodity futures market and a factor to account for refining costs, taxes and distribution costs. However, oil price futures have been known to be very poor predictors of future price changes in crude markets (Hamilton 2009, 1983; Menzie and Coibion 2013). Sawhill (2008) and Khan (1986) both fit ARIMA models to retail gasoline prices to obtain consumers expectations of future gasoline price.

Vehicle price is affected by omitted variable bias for several reasons (Train 2003) and as such the ξ_j term does not have a zero mean. For example, the econometrician may have information on fuel efficiency, weight, length, horsepower, color, etc, but may not have information on other attributes such as prestige, style,

⁴⁴ Alcot and Wozny (2009) assume that consumers all consumers are homogeneous in their ability to trade-off future fuel savings for current vehicle prices. The model specified in 2 allows some heterogeneity in consumers' ability to tradeoff vehicle price for future fuel costs at different rates.

comfort, etc. These attributes affect price. The existence of the unobserved attributes implies that vehicles sort into groups or market segments, and that models within each segment compete more closely with one another than with models belonging to other market segments. This causes the price, to be correlated with the unobserved attributes, ξ_j . One method to overcome this endogeneity is to move ξ_j into the component of the utility that is observed by introducing an alternative specific constant for each product j (Petrin and Train 2006, Train 2009, BLP 1995). However, due to the difficulty of estimating the contraction required for the BLP approach and the fact that some vehicle models could have zero shares because the model is not offered for sale in some periods; this research will adopt the control function approach to deal with potential endogeneity of vehicle price (Kim and Petrin 2010, Klein and Vella 2010).⁴⁵

The probability that household i choose alternative j on choice occasion t , conditional on the model structural parameters is given by

$$\Pr_{it}(j) = \frac{e^{V_{ij}}}{\sum e^{V_{ij}}} \quad (5)$$

Averaging P_{ij} over the N consumers in the sample (Train and Winston 2007), we can obtain the market share for each product. We then express the market share for each product as:

$$S_j = \sum_i P_{ij} / N \quad (6)$$

Conditional on the household choosing alternative j , the utilization/demand for vehicle miles travelled is obtained from equation 2 by applying Roy's Identity. If we

represent household i actual utilization of vehicle j at period t as M_{ijt} , and \tilde{M}_{ijt} as the observed utilization (because of measurement error and the fact that preferences are not actually observed), then the utilization equation is $\tilde{M}_{ijt} = M_{ijt} + \vartheta_{ijt}$, where ϑ_{ijt} is the error term. The same variables are used in estimating both the choice and utilization equations.

Supply Model and Equilibrium Prices

This section develops the supply side of the model in which vehicle manufacturers set prices to maximize profits, given price of its competitors. Vehicles are differentiated, that is two very similar vehicles from different manufacturers will be priced differently. The supply side is similar to the supply side model of BLP (1995) and Train (2003). The intuition behind the supply side model is that given the aggregated demand curves and market shares, one can solve for the product specific marginal costs by solving the equilibrium first order conditions. This approach allows us to solve for new equilibrium prices given counterfactuals and assumptions about consumer behavior. We assume that firms, $f = 1, \dots, F$, compete in a Bertrand Nash fashion under a differentiated product. The manufacturer sets prices $p_j = (p_{1j}, \dots, p_{Tj})$, let $\xi = (\xi_1, \dots, \xi_j)'$ denote vector of unobserved product attributes. In the short term, firms only change prices. Firm f 's profit function is given by:

$$\text{Max} \Pi_f(p, Z, \xi) = \sum_{j=1} p_{jt} - mc(X, Z)_j S_j(p) N \quad (7)$$

where p_j is as defined earlier, N is the number of US households, S_j is the predicted market share obtained by summing individual consumers weighted sum of vehicle selection probabilities, mc is the unit variable cost of product j . Similarly as in

equation 2, X contains the consumer expectations about future fuel costs and Z represents exogenous demand shifters.

Following the procedure developed by Villas-Boas (2007)⁴⁶, we solve equation 7 for the first order conditions with respect to price. From the first order conditions in a Bertrand Game, automobile prices depend on the marginal costs of a product. The marginal cost in turn depends on both observed product attributes and unobserved product attributes, input prices (rent, wages, and other cost shifters). Some of the cost shifters could be observed by the econometrician.

$$mc_j = W(X_j, c_j, \eta) + \mu_j \quad (8)$$

where $W(\cdot)$ is a parameter function. μ_j is a random error, normal, and depends on both observed and unobserved attributes. Solving for the first order conditions, one obtains equations that satisfy the price-cost mark-ups $p_j - mc_j$ ⁴⁷ and market share for each product that satisfies the function given below. The vector of mark-ups only depends on the parameters of the demand equations and equilibrium price vector.

$$\frac{\partial \Pi_f(p, X, \xi)}{\partial p_j} = S_j(p, X; \xi) + \sum_j (p_j - mc_j) \frac{\partial S_j(p, X; \xi)}{\partial p_j} = 0 \quad (9)$$

$$\mathbf{mc} = \mathbf{p} - \nabla(\mathbf{p}, \mathbf{X}; \xi)^{-1} \mathbf{S}(\mathbf{p}, \mathbf{X}; \xi) \quad (10)$$

where ∇ is a $J \times J$ matrix whose terms are given by:

$$\Delta_{j\eta} = \begin{cases} -\frac{\partial S_j}{\partial p_j} & \text{if } j \text{ and } \eta \text{ are produced by same firm} \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

We can easily use 11 for counterfactual policy simulation to solve for new equilibrium price vectors.

⁴⁶ Also Villas-Boas and Winer 1999

⁴⁷ The Bertrand Model without product differentiation implies price equals marginal cost. Because we are assuming product differentiation, this is not the case here.

Data

Data

The data sets used for the analysis were obtained from several sources. The primary data used for this analysis is the automobile extract from the Consumer Expenditure Survey. The Consumer Expenditure Survey (CEX), a rotating panel of US household purchasing pattern over 5 quarters conducted by the Bureau of Labor Statistics, is a stratified random sample. Every quarter, 25% of the households in the sample are replaced with new households. I make use of data from 2002-2005. Each quarter about 4500-5000 distinct individuals are interviewed on their expenditure; 75% of those interviewed are re-interviewed and a new set of households replaces the other 25%. The automobile extract provides detailed information on all cars owned by households at the beginning of the survey year, on any cars purchased between the households first and last year as well as on any disposal of vehicles by the households. Table 1 provides information on demographics⁴⁸ of the estimation sample, and it is consistent with U.S population's socioeconomic data. Since the CEX are weighted random sample of representative U.S. households, we expect the data to be representative of aggregate vehicle purchase pattern of all U.S.⁴⁹ This table also shows significant variations in household characteristics across the vehicle classes. For example, larger households especially with more kids prefer SUVs and larger vehicles. Wealthier households (as measured by total yearly expenditures) also possess larger

⁴⁸ Among the wide range of household demographic information the survey contains includes size of household, age of household members, sex, employment status, type of residence, whether or not the household is located in a rural or urban area, education level, household income

⁴⁹ However, Goldberg (1998) suggests that the CEX underestimate the total sales of new vehicles due to the existence of fleet sales.

vehicles. Also, households with more workers or income earners and those with male heads are inclined to have SUVs.

Each year about 400 households in the CEX sample buy new vehicles (that is about 10% of sampled households purchase new vehicles) from approximately 200 different make and models. Therefore some models only appear infrequently in our estimation sample while others may not appear at all. We exclude a small number of households due to missing information on key variables such as date of purchase and other attributes. We are then left with 1,506 households who purchased new vehicles between 2002 and 2005.

The information on automobiles includes the make, model, vintage and purchase price of each car, financing, disposal of old vehicles, and a large set of vehicle characteristics. Most importantly, the CEX includes the mileage of each car owned by the household during each quarter. I will supplement these datasets with data on vehicle characteristics from EPA fuel economy test data and Automotive News Market Data Book. The former provides information on fuel economy measured in miles per gallon⁵⁰. While the latter includes information on size, performance, and standard options of various models. Data on annual observations of certain microeconomic and macroeconomic variables also included as well as gasoline prices. Summary statistics of the attributes of the estimation model is provided in Table 2. Each consumer's choice set consists of 198 makes and models of new vehicles.

Information on gasoline prices (incl. state and local taxes) is taken from the U.S. Energy Information Administration (EIA). EIA collects weekly retail gasoline

⁵⁰ The mpg values are combined EPA mpg values: a weighted average of City and Highway MPG values that is calculated by weighting the City value by 55% and the Highway value by 45%.

prices for all formulations (conventional, oxygenated and reformulated gasoline prices), which are recorded by State. We match individual households in the automobile purchase dataset to state-level gasoline prices based on the respondents state of origin. A big advantage of focusing on the 2002-2005 period is that it includes the sharp decline of gas prices at the end of 2002-beginning 2003 so that there is ample variation in the data to identify the consumer responses to lower gasoline prices as well as higher gas prices in late 2005.

Finally, we also collect information on interest rates on new and old car loans from the Federal Reserve and Consumer Price Index for all goods from the Bureau of Labor Statistics.

Discounted Future Gasoline Price Expectations

Gasoline prices are generally difficult to predict. When consumers make vehicle purchase decisions, how they form expectations for future gasoline prices is not well known. The most common assumption past works have made is that consumers treat the gasoline price process as a random walk, with any price shock considered permanent with subsequent adjustments in demand (Biesebroeck and leuven 2010) . However if the price shocks decay rapidly, then estimated elasticity will be biased because price shock will have less of an effect on demand and measured price elasticities will be lower.

Consistent with prior literature, we will make the assumption that current gasoline prices are best predictors of the expected future gasoline prices, i.e. gasoline prices are follow a random walk process with time trend. Support for the random walk hypothesis comes from Anderson *et al* (2012) who concluded that the random walk

process is a good reflection of consumer expectations of future gasoline prices.

Additional support is provided by the result of a dickey fuller test in which we fail to reject the null hypothesis that average monthly gasoline prices exhibits a unit root.

Figure 1 is a time series plot of the natural log of the real price of gasoline prices between January 1994 to December 2012.

A look at this figure casts some doubt about the propriety of using a random walk model to represent consumer expectations of future gasoline prices. It looks more like a random walk with trend or ARIMA model. As a robustness check, we specify an alternative model of consumer expectation of future gasoline prices that follows an ARIMA process. Monthly data on real gasoline prices from 1994-2005 indicates an ARIMA (1, 1, 1) in the log of real gasoline prices. To construct the expectations series for each year with the ARIMA model, we use only lagged data to generate sets of forecasts that use information from the customers to update their most recent forecasts. Figure 1b is a plot of the forecasts from the model using this procedure superimposed over forecast that just uses one period ahead.

Another important assumption we make in constructing consumer expectations of future gasoline costs is related to consumer demand for annual vehicle mileage conditional on buying a new vehicle. We assume that consumer demand for annual vehicle miles travelled is heterogeneous and varies with gasoline price. While the short run elasticity of vehicle usage with respect to gasoline is low -0.15 (Hughes et al 2008, Gillingham 2011); in the long run, the average elasticity values hover around -0.7. That is a 10 percent rise in gasoline prices will reduce driving by about 7 percent through people switching to cars with better mileage and driving less. Most of the

existing literature recognizes that consumers show significant variation in their demand for annual mileage and we incorporate the consumers' heterogeneity for fuel mileage by using customer specific mileage reported in consumer expenditure survey.

Incorporating Heterogeneity and Sorting

The wide variation in estimates of consumers' valuation of fuel economy suggests that heterogeneity is still an issue. Take for example, the role of unobserved individual heterogeneity. One such heterogeneity is consumer driving habits. Green drivers may buy more fuel efficient vehicles and use them sparingly (i.e there is positive correlation in vehicle choice draws across drivers), then estimates of the consumer MWTP for fuel economy will be unnecessarily too high. It is also possible that the correlation will be in the opposite direction. That is, heavy drivers will sort into more fuel efficient vehicles in order to reduce cost per mile of driving.

A second problem arises if there is unobserved heterogeneity in individuals expectations of gasoline prices. That is expectations might vary over consumers and time. For example, consumers might expect gasoline prices to rise (especially if purchase decisions are myopic), and choose to purchase more fuel efficient vehicles thus lowering cost per mile of driving. When prices are high and rising, it might be perfectly rational to expect future prices will be higher compared to expectations that are formed when prices are low and stable. If we use average expectations of the consumer to estimate consumer demand for fuel economy instead, that estimate will be biased.

A third problem arises when individual drivers have heterogeneous rates of time preferences. In particular, consumers who put less value in future costs (have

present biased preferences) will underweight gasoline costs relative capital costs. This is because gasoline costs occur in the future. Therefore, consumers that exhibit present bias preferences will sort into less fuel efficient vehicles, biasing downward estimates of consumer valuation of fuel economy. Although panel data and individual fixed-effects offer one solution to preference based sorting, other researchers have also used other methods. Our estimation approach allows us to account for the individual unobserved heterogeneity. We adopt the general form of heterogeneity-random parameter representation. However, we re-parameterize the

$$\alpha_{it,k} = \alpha_k + z'_i \theta_k + \sigma v_i$$

$$\gamma_{it,k} = \gamma_k + z'_i \theta_k + \sigma v_i$$

mean estimates of the random parameters (vehicle capital costs and gasoline prices) to establish heterogeneity (Hensher and Green 2003). Where $z'_i \theta_k + \sigma v_i$ is the mean of the distribution, z_i is a vector of consumer specific attributes, $k = 1, \dots, K$ is simulated to find the values that maximize the simulated log likelihood (Green 2012). The time-invariant error terms reflects habit-persistence. The presence of a standard deviation of the two parameters accommodates the presence of preference heterogeneity in the sampled population.

Estimation and Identification Strategy

Econometric Model and Estimation Strategy

Our estimation approach is driven by several factors. First, is that the durable goods nature of automobiles implies that we have to account for both consumers' taste for fuel efficiency and intensity of utilization in estimating elasticity. Therefore,

households make a two-step decision. They first choose the number and type of vehicles to own based on driving habits and expectations of future gasoline prices. And in the second step, conditional on their vehicle bundles, they choose utilization intensity which determines fuel consumption. The second decision influences decisions about vehicle type choice in the first step. For example, a driver whose demand for mileage is greater than the average driver should have a larger elastic demand for fuel efficient vehicles and lower elasticity in the demand for gasoline price. Part of this stems from the fact that the MWTP for fuel economy is an increasing function of gasoline price (Van Biesebroeck and Leuven 2010).

Second, the error term of the different dimension of vehicle choice and utilization are correlated. This correlation could be in either direction. For example, if the unobserved taste parameter which induces some households to choose fuel efficient vehicles also induces them to drive less, then the error terms will be positively correlated. We can also imagine a situation in which the unobserved taste parameters induce heavy drivers to increase utilization, resulting in the error terms being negatively correlated.

Third, about 31% of the households in our sample have two vehicles and 14% owning multiple vehicles. This lends to “portfolio effects” as a possible interpretation for households owning more than one vehicle. That is the mix of vehicles owned by households satisfies different functions. We follow Dube(2004) and Hendel(1999) in accounting for households ownership of portfolio of vehicles in a mixed logit with repeated choice occasions framework. On each choice occasion, households make discrete decision of vehicle type and amount of utilization conditional on type.

Because we do not observe each choice occasion in our data set, we are unable to precisely characterize the actual situations in which households consume each alternative. However, we do observe the mile driven by each vehicle type, and this allows us to model the distribution of consumption occasions and make inference about the context in which they are consumed. One can relate different choice occasions as representing different activities undertaken by the household that requires driving such as school runs, shopping, weekend family gateways, etc. Intuitively, household vehicle portfolio and annual mileage indicates heterogeneity in tastes for various choice occasions. We explain the vehicle quantity of each household in terms of demographic attributes (number of people in the household that work), vehicle attributes, fuel costs, travel costs, etc. Intuitively, we expect households to switch from a gas guzzler to a more fuel efficient vehicle if gasoline prices rise and vice-versa.

Identification

Our identification strategy is based on the fact that firms set prices on the basis of demand for their products and that of the competitors. The prices (transaction prices) are the result of negotiations with customers, and auto dealers will provide a higher discount on cars that are selling slowly. The observed prices contain information about price elasticities, which if correctly extracted can be used to estimate the demand parameters. Therefore, we need to look at pricing behavior that can arise in markets and show how the specification of this behavior can be combined with the demand estimation to identify consumer demand for fuel economy. The intuition is that auto firms also respond to fuel price changes in the short run by adjusting relative prices to match sales. And in the medium run, firms can tweak small

design changes in months to improve fuel economy by orders of 0.1 to .2 mpg (Busse *et al* 2013, Sallee and Slemrod 2010). An exogenous change in fuel prices thus triggers endogenous changes in consumers' decision environment.

Given that we have a model for consumer expectations of future fuel costs, intuitively we can estimate the logit demand model and test the null hypothesis that the marginal utility of saving a dollar on the cost of a vehicle is equal to the marginal utility of saving a dollar on the discounted expected operating costs over the vehicle life span:

$$\frac{\partial U_i / \partial p_j}{\partial U_i / \partial E []} = \frac{\alpha_i \gamma_i}{\gamma_i} = \gamma_i = 1 \quad (6)$$

If firms compete under Nash Bertrand, when prices depend on elasticity of demand and marginal cost, the pricing equation is:

$$p_j + \left(\frac{p_j}{e_{ij}} \right) = mc_j$$

$$p_j = - \left(\frac{p_j}{e_{ij}} \right) + W(X_j, \phi) + \mu_j \quad (7)$$

We assume that ξ and μ are jointly normal; the conditional distribution of ξ_j given μ_j is also normal. The unobserved component of the demand equation, ξ_j , now enters the pricing equation through the elasticity. As such we can rewrite the pricing equation as (Kim and Petrin 2010):

$$p_j = X_j(x_j, \phi) + \mu_j \quad (8)$$

where x_j is a vector of all of the observed exogenous variables that affect marginal cost and the elasticity, $X_j(\cdot)$ is a parametric function of these variables, and u_j is the unobserved deviations of price around this function (Train 2010, Kim and Petrin 2010). We can estimate the pricing equation and retain the residuals. This is done by constructing estimates of the expected price for each model conditional on all exogenous factors⁵¹.

Estimation proceeds in two steps. In the first step, we recover μ^* , (the residuals from the pricing equation) by inverting the pricing function. The product specific unobserved factor, ζ , enters the reduced form pricing equation as an argument. The μ^* incorporates information about the μ_j and unobserved component of the elasticity based mark-up (Train 2002). In the second step, having specified the pricing function, we then estimate conditional choice model treating the controls as additional regressors. What is left is the standard errors. We need to correct them using the approach in Newey and McFadden (1994).

Because we have very few observations on the same nameplate over time and moreover, vehicle attributes change, new models are added, some models are discontinued, we cannot condition on a full set of fixed effects. Rather, we place restrictions across the vehicles. The only restriction we place is that the parameters of the pricing function are the same across all products and years.

Results

Parameter Estimates of Demand Model

⁵¹ We assume that μ and x are not correlated.

Point estimates and standard errors from the specifications are presented in Tables 4 and 5. Table 6 translates these results to elasticity estimates. The two variables of interest are capital cost of vehicle (α) and discounted annual operating costs (γ)⁵². Given that there is no specific direction in existing literature for other variables to include; we include weight, horsepower, size (wheelbase and track width), drive train, dummies if vehicle is an SUV, minivan, full sized van or pickup truck, and indicators that reflect make of a given vehicle. The indicators capture extent of household preference for a particular make of vehicle. We also control for engine size, but the values are not reported because they are not of any particular interests for our purposes. The demographic variables enter through interactions with vehicle attributes. The parameters vector has a multivariate normal distribution. We divide the coefficient on price by the income of consumer to allow the elasticity to vary with income. This way, households with high income are less sensitive to prices than those with low income.

Column 1 of Table 4 reports the logit specification without accounting for heterogeneity or including information from firms pricing behavior. All other coefficients are statistically significant. The coefficient on price is negative as expected, but that on operating cost is positive which is counter intuitive. In Column 2, we report random coefficient logit model with heterogeneity but do not include information on firms pricing behavior nor account for price endogeneity. The

⁵² Fuel economy enters the consumers' utility function as annual operating costs. This is determined as real price of gasoline per gallon divided by mpg of fuel, multiplied by the average annual vehicle miles. This formulation accounts for the wide variations in the price of gasoline in different locations and in annual miles travelled by different households. Importantly, it implies an inverse relationship between increasing fuel economy and utility since fuel economy has an inverse relationship with fuel consumption, while fuel expenditure have a linear relationship with consumption.

coefficient of discounted annual operating costs switches signs. In addition, price and all other variables have the right signs and are statistically significant. The values are in line with estimates elsewhere⁵³ (Train and Winston 2003, Petrin 2002, Kim and Petrin 2010, Sawhill 2008, and BLP 1994).

Column 3 of Table 4 is our preferred model. In addition to accounting for consumer heterogeneity and price endogeneity, this model also includes information in firms pricing behavior. We expect the price elasticity to be higher than the result from the random coefficient logit model (Train 2012, Kim and Petrin 2010). Correcting for endogeneity and sorting improves the elasticity of both the capital cost and gasoline costs significantly. The mean price coefficient is 3.581 and significant. The standard deviation of a normally distributed price coefficient is 0.54.⁵⁴ We recover mean price coefficient with respect to discounted operating costs of 6.78⁵⁵ with a t-value of 14.74. We also find standard deviation of the normally distributed coefficient of discounted yearly operating costs of 2.085 which implies that 95% of the drivers have mean coefficients ranging from 2.69 to 10.87.

We also report ratios of the coefficients of the two variables of interest: $\frac{\gamma}{\alpha}$, with a value of 1.88. A Wald test of the null hypothesis that the coefficients on price and discounted operating costs are equal is rejected at 5% level ($W=26.48 < \chi^2 < 3.65$). We find that 95% of the consumers in our sample have $\frac{\gamma}{\alpha}$ values ranging from 0.99 to 1.88. This large spread signifies heterogeneity in consumer response to fuel economy. In

⁵³ The coefficient on vehicle cost.

⁵⁴ This implies that 95% of the price coefficient over all models varies between 2.6 to 4.56.

⁵⁵ Estimates of operating costs we report are not comparable to those reported elsewhere since they represent different measures of fuel efficiency.

this vein, we uncover three segments of discounted operating costs: 28% of the consumers significantly overvalue fuel economy while 61% less so or are rational. That is 61% of consumers in our sample make vehicle purchase decisions based on financial considerations only. The rest, 11%, seems to underweight future gasoline costs when making decisions. We profile the segments in terms of their average demographic characteristics. On average, rational households are larger and tend to have above average income, have higher number of school age children, men older than 45 years.

These are in line with the predictions from Bento *et al* (2012) and the effect of sorting and heterogeneity on parameter coefficient estimates (Bayer *et al* 2011). These results reinforce their arguments that ignoring consumer heterogeneity leads to undervaluation of energy savings by consumers. This is because consumer heterogeneity results in sorting which results in downwards bias effect. These results are in contrast to Alcott and Wozny (2009) who finds that report that consumers undervalue fuel economy.

Cost Side Parameter Estimates

Table 5 reports the results of the regression of log of marginal costs on the vehicle characteristics. All the regressors enter with significant coefficients and have the expected signs. That is, it costs more to improve the fuel efficiency of a vehicle, build a bigger car, add additional luxury and comfort, power and acceleration, as well as increase weight. Next, we turn to the main question of this research by looking at the implications of the estimates of accounting for pricing behavior of firms in consumer valuation of fuel economy.

Do Consumers Undervalue Fuel Economy

Table 6 presents estimates of several ratios of $\frac{\gamma}{\alpha}$ computed over a range of discount rates and vehicle lifetime values. These values can be interpreted as the consumer willingness to pay for reduced future discounted costs. Expressing the results in this form makes it easy and the interpretation is intuitive. However, these ratios can be easily translated into a form that allows the values to be comparable to results reported elsewhere. We achieve this by dividing the ratio $\frac{\gamma}{\alpha}$ by $\sum_{s=0}^T \left(\frac{1}{1+r}\right)^s$. Using data from NHTSA vehicle survivability and travel mileage schedules (NHTSA 2006), we assume the average vehicle has a useful life (T) of 15 years. To check sensitivity, we try different values for T 10, 15, 20 and 25. Also, we assume discount rates ranging from 5-9%. Most published work have used discount rates of either 6 or 7%. The values in Column 3 are comparable to those reported by Alcott and Wozny (2009) who used a discount rate of 6% for their base model.

In Figure 3, we plot the distribution of WTP to reduce future operating costs of vehicles by households in our sample over different vehicle life times and interest rates. The last two columns of Table 6 report our results. We estimate that for every \$1 dollar saved in the future, consumers take no more than \$0.99 into account. We find evidence of small overvaluation and not one instance of undervaluation. These results are in stark contrast to results reported by Helfand and Wolverson 2010; Greene 2010; Alcott and Wozny 2009. We find that accounting for heterogeneity and sorting by consumers, tends to remove any significant evidence of consumer undervaluation of future fuel savings. The last two rows present lower 5% and upper 95% distribution of consumer valuation of fuel economy. Though the lower 5% consumers only place

values of \$0.51 of a dollar savings, it does not signify energy paradox by this population in the strict sense. This is because, there is a huge spread in the amount of annual mileage by consumers across the population and this leads to huge valuation in fuel economy. For example, if they drive a few standard deviations outside the mean, they might choose the appropriate vehicle. The major take away from this is that different drivers sort into different vehicles based on their taste and demographic attributes. Those consumers who place higher values on fuel economy will sort into more fuel efficient vehicles and those that place less value of future fuel savings sort into large vehicles. Any demand estimation strategy that does not account for heterogeneity in driving habits and taste preferences will probably arrive at the wrong conclusion.

Conclusions

How consumers value fuel economy is one of the most researched questions in to the automobile and energy sectors. The estimates are highly variable with some papers concluding that consumers grossly undervalue fuel economy, while others have reached the exact opposite situation. The high variability of the estimates signifies that probably, heterogeneity in consumer might play a role in reconciling the different results obtained. Despite the fact that individual unobservable heterogeneity and sorting have been found in consumer vehicle choice and driving behavior, their effect on estimates of consumer fuel economy valuation has not been presented empirically. This paper presents such empirical evidence of the effect of individual unobserved heterogeneity and sorting in the estimates of fuel economy valuation by consumers. We demonstrate that sorting and unobserved heterogeneity has the potential to bias

estimates of consumers' ability to optimally tradeoff capital and operating costs when making durable choice decisions.

To restate our findings, we find no evidence to support the argument that consumers systematically underweight the cost of future events in real market settings. However, we find significant evidence that different consumers sort into different groups such that rational consumers accurately tradeoff future operating costs and vehicle prices. Conservatively, about 61% of the population is accurately do not underweighting or overweighting operating costs when purchasing new vehicles. This research has several limitations which provide a useful roadmap for future studies. First, we do not include used vehicles in our model. There is no evidence that consumers value new vehicles and old vehicles in a similar version or that different consumers drive their vehicles over their useful life as developed in this paper. In fact, evidence from textbook market exists that forward and myopic consumers place different values on the resale price of durable goods. Ideally, the model could be extended to allow for dynamic vehicle replacement as well creating a fully dynamic framework. Second, the model developed in this paper makes strong assumptions about the functional form of the distribution of consumers in the population for ease of estimation. By restricting consumer taste parameters for price and operating costs to be normally distributed, we are necessarily ruling out more flexible distributions of consumer behavior which may be present in the population.

REFERENCES

- Ackerberg, Daniel A., and Marc Rysman. "Unobserved Product Differentiation in Discrete-Choice Models: Estimating Price Elasticities and Welfare Effects." *RAND Journal of Economics* 36, no. 4 (2005): 771–88.
- Allcott, Hunt. "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market." *American Economic Journal: Economic Policy* 5, no. 3 (August 2013): 30–66. doi:10.1257/pol.5.3.30.
- Allcott, Hunt, and Nathan Wozny. *Gasoline Prices, Fuel Economy, and the Energy Paradox*. Working Paper. National Bureau of Economic Research, November 2012. <http://www.nber.org/papers/w18583>.
- Amil Petrin and Kyoo-il Kim, Control Function Corrections for Omitted Attributes in Differentiated Product Models, 2010, Working Paper.
- Anderson, Soren T., Ryan Kellogg, and James M. Sallee. "What Do Consumers Believe about Future Gasoline Prices?" *Journal of Environmental Economics and Management* 66, no. 3 (2013): 383–403.
- Athey, Susan, and Guido W. Imbens. "Discrete Choice Models with Multiple Unobserved Choice Characteristics*." *International Economic Review* 48, no. 4 (November 1, 2007): 1159–92. doi:10.1111/j.1468-2354.2007.00458.x.
- Bento, Antonio M., Shanjun Li, and Kevin Roth. "Is There an Energy Paradox in Fuel Economy? A Note on the Role of Consumer Heterogeneity and Sorting Bias." *Economics Letters* 115, no. 1 (2012): 44–48.
- Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile Prices in Market Equilibrium." *Econometrica* 63, no. 4 (1995): 841–90.
- Berry, Steven T. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics* 25, no. 2 (1994): 242–62.
- Brownstone, David, David S. Bunch, and Kenneth Train. "Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-Fuel Vehicles." *Transportation Research Part B: Methodological* 34, no. 5 (June 2000): 315–38. doi:10.1016/S0191-2615(99)00031-4.
- Busse, Meghan R, Christopher R Knittel, and Florian Zettelmeyer. "Are Consumers Myopic? Evidence from New and Used Car Purchases." *American Economic Review* 103, no. 1 (February 2013): 220–56. doi:10.1257/aer.103.1.220.
- Chinn, Menzie D., and Olivier Coibion. "The Predictive Content of Commodity Futures." *Journal of Futures Markets* 34, no. 7 (July 1, 2014): 607–36. doi:10.1002/fut.21615.
- DeLeire, Thomas, Shakeeb Khan, and Christopher Timmins. "Roy Model Sorting and Nonrandom Selection in the Valuation of a Statistical Life*." *International Economic Review* 54, no. 1 (February 1, 2013): 279–306. doi:10.1111/j.1468-2354.2012.00733.x.

- Dreyfus, Mark K., and W. Kip Viscusi. "Rates of Time Preference and Consumer Valuations of Automobile Safety and Fuel Efficiency." *Journal of Law and Economics* 38, no. 1 (1995): 79–105.
- Gillingham, K. Forthcoming. Identifying the Elasticity of Driving: Evidence from a Gasoline Price Shock in California. *Regional Science and Urban Economics*
- Goldberg, Pinelopi Koujianou. "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry." *Econometrica* 63, no. 4 (1995): 891–951.
- Greene, D. 2010. How Consumers Value Fuel Economy: A Literature Review. Technical Report EPA-420-R-10-008.
- Hamilton, James D. *Understanding Crude Oil Prices*. NBER Working Paper. National Bureau of Economic Research, Inc, 2008.
<http://ideas.repec.org/p/nbr/nberwo/14492.html>.
- Hassett, Kevin A., and Gilbert E. Metcalf. "Energy Conservation Investment: Do Consumers Discount the Future Correctly?" *Energy Policy* 21, no. 6 (June 1993): 710–16. doi:10.1016/0301-4215(93)90294-P.
- Hausman, Jerry A. "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables." *Bell Journal of Economics* 10, no. 1 (1979): 33–54.
- Hughes, Jonathan E., Christopher R. Knittel, and Daniel Sperling. "Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand." *The Energy Journal* 29, no. 1 (2008): 113–34.
- Kahn, James A. "Gasoline Prices and the Used Automobile Market: A Rational Expectations Asset Price Approach." *The Quarterly Journal of Economics* 101, no. 2 (May 1, 1986): 323–39. doi:10.2307/1891118.
- Klier, Thomas, and Joshua Linn. "The Price of Gasoline and New Vehicle Fuel Economy: Evidence from Monthly Sales Data." *American Economic Journal: Economic Policy* 2, no. 3 (August 2010): 134–53. doi:10.1257/pol.2.3.134.
- Mannering, Fred, and Clifford Winston. "A Dynamic Empirical Analysis of Household Vehicle Ownership and Utilization." *RAND Journal of Economics* 16, no. 2 (1985): 215–36.
- McManus, Walter. *The Effects of Higher Gasoline Prices on U.S. Light Vehicle Sales, Prices, and Variable Profit by Segment and Manufacturer Group, 2001 and 2004*. Industrial Organization. EconWPA, 2005.
<http://ideas.repec.org/p/wpa/wuwpio/0505010.html>.
- . *The Link between Gasoline Prices and Vehicle Sales: economic Theory Trumps Conventional Detroit Wisdom*. MPRA Paper. University Library of Munich, Germany, 2007. <http://ideas.repec.org/p/pramprapa/3463.html>.
- Parry, Ian W. H, Margaret Walls, and Winston Harrington. "Automobile Externalities and Policies." *Journal of Economic Literature* 45, no. 2 (June 2007): 373–99. doi:10.1257/jel.45.2.373.

- Petrin, Amil. *Quantifying the Benefits of New Products: The Case of the Minivan*. NBER Working Paper. National Bureau of Economic Research, Inc, 2001.
<http://ideas.repec.org/p/nbr/nberwo/8227.html>.
- Petrin, Amil, and Kenneth Train. *Control Function Corrections for Unobserved Factors in Differentiated Product Models*. 2006 Meeting Paper. Society for Economic Dynamics, 2006. <http://ideas.repec.org/p/red/sed006/856.html>.
- Sallee, James M., and Joel Slemrod. "Car Notches: Strategic Automaker Responses to Fuel Economy Policy." *Journal of Public Economics* 96, no. 11 (2012): 981–99.
- Sallee, J.M., S. E. West and W. Fan, 2010. "The Effect of Gasoline Prices on the Demand for Fuel Economy in Used Vehicles: Empirical Evidence and Policy Implications", Draft version 1.1 of January 19, 2010.
- Sawhill, J.W., 2008. Are Capital and Operating Costs Weighted Equally in Durable Goods Purchases? Evidence from the U.S. Automobile Market, discussion paper, Department of Economics, University of California at Berkeley, Berkeley, California, April.
- Train, Kenneth. *Discrete Choice Methods with Simulation*. Online economics textbooks. SUNY-Oswego, Department of Economics, 2003.
<http://ideas.repec.org/b/oet/tbooks/emetr2.html>.
- Train, Kenneth E., and Clifford Winston. *Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers*. Working paper. Regulation2point0, 2006.
<http://ideas.repec.org/p/reg/wpaper/331.html>.
- Van Biesebroeck, Johannes. *The Demand for and the Supply of Fuel Efficiency in Models of Industrial Organisation*. OECD/ITF Joint Transport Research Centre Discussion Papers. Paris: Organisation for Economic Co-operation and Development, January 1, 2010. <http://www.oecd-ilibrary.org/content/workingpaper/5kmjp66c9vjc-en>.
- Verboven, Frank. "Quality-Based Price Discrimination and Tax Incidence: Evidence from Gasoline and Diesel Cars." *The RAND Journal of Economics* 33, no. 2 (2002): 275.
 doi:10.2307/3087434.
- Villas-Boas, J. Miguel, and Russell S. Winer. "Endogeneity in Brand Choice Models." *Management Science* 45, no. 10 (October 1, 1999): 1324–38.
 doi:10.1287/mnsc.45.10.1324.
- Villas-Boas(2007) A Note on Limited versus Full Information Estimation in Non-Linear Models, Working Paper, Haas School of Business, University of California, Berkeley.

Figure 1: Plot of Natural Log Real Gasoline Price

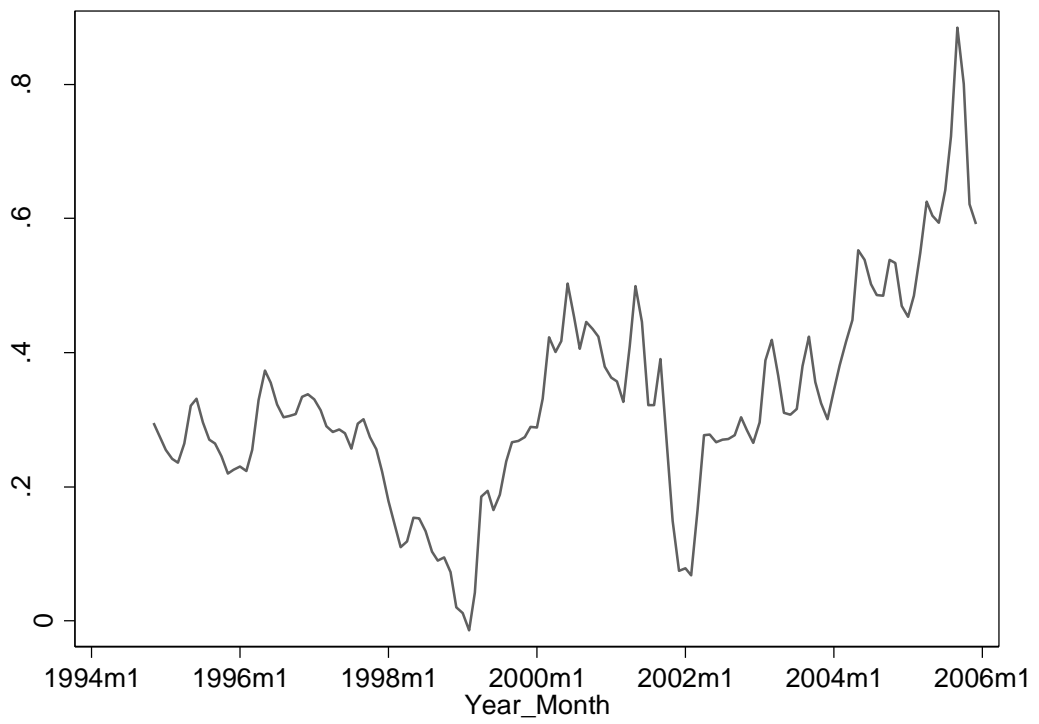


Figure 2: Plot of Natural log of Real Gasoline Price and Forecasts

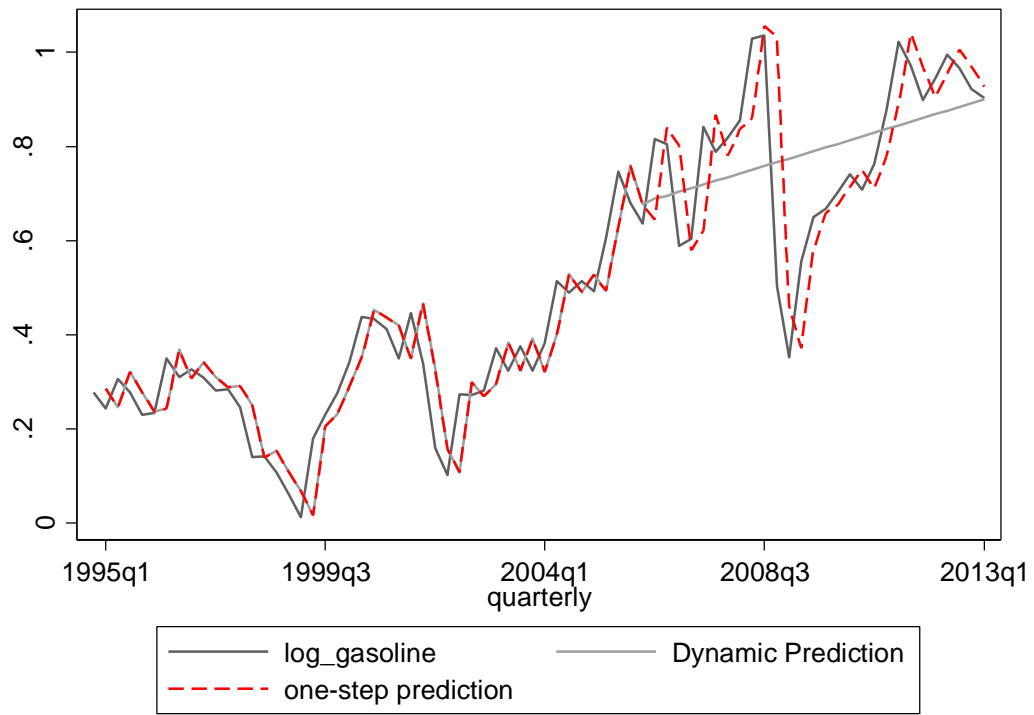


Figure 3: Distribution of Willingness to Pay to Reduce Operating Costs

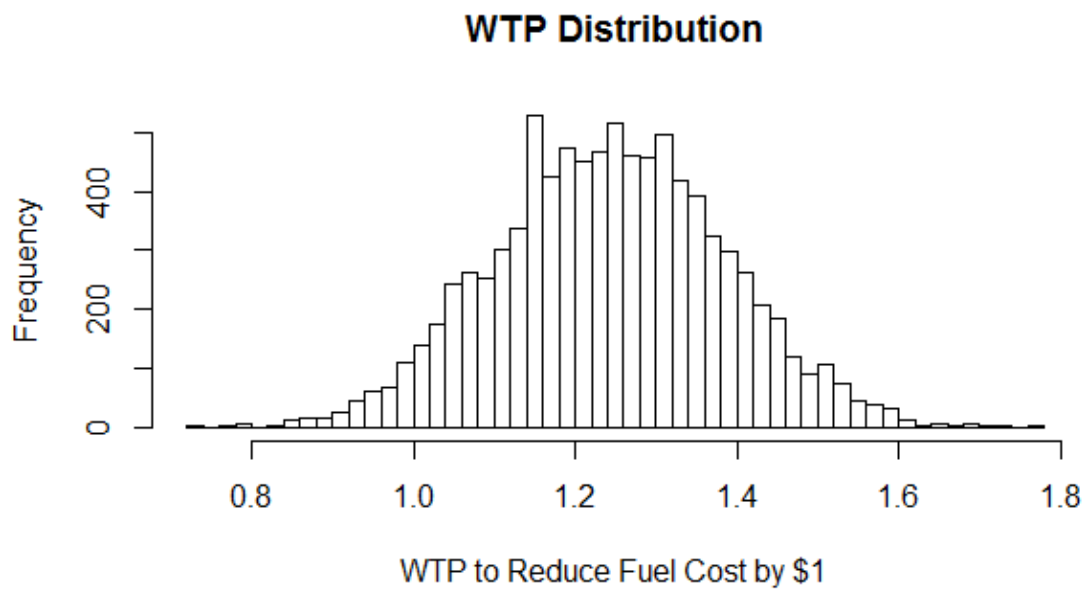


Table 1: Description of Sample Households who Bought New Vehicles

Socioeconomic Characteristics	
Variable	Sample Value
Average Household Income	\$68,908
Average Age	50
Average Household Size	2.61
Percentage Male	57
Percentage with Child 1-6	15
Percentage who live in large MSAs	57

Large MSAs are described as being within MSAs with a population greater than 1 Million people.

Table 2: Summary Statistics of Sample Vehicle Characteristics

Automobile Characteristics		
Variable	Mean Value	Standard Deviation
Annual Vehicle Miles Travelled (Miles)	11,861	17930
Real Gasoline Price (\$/gallon)	1.75	0.51
Vehicle Price	\$22,716	\$10,006
Length (Inches)	192.1	16.5
Wheelbase (Inches)	112.9	11.4
HP (Pounds)	194	52
Curb Weight (Pound)	3479	695
Mpg	20.16	1.45

Table 3: ARIMA (1,1,1) Model Annual US Retail Gas Price Data 1994-2005

Variable	Estimate	Standard Error
<u>Log of Real Gasoline Price</u>		
Constant	0.5208	0.126
ARMA		
AR1	0.8593	0.0691
MA1	0.4671	0.1182
Variance of Residuals (σ^2)	0.0102	0.0074

Table 4: Parameter Estimates of Consumer Demand for Fuel Economy

Demand Side Parameters	No Correction (1)	Random Coefficient-RC (2)	RC with Correction (3)
Mean(β's)			
Constant	-10.600 (0.360)	-7.079 (0.918)	-8.093 (1.009)
Vehicle Cost/Income (\$'000)- α	-0.76 (0.121)	-2.13 (0.029)	-3.58 (0.023)
Operating Costs/Income- γ	0.165 (0.080)	-4.412 (0.289)	-6.78 (0.460)
Vehicle Size (Wheelbase * Length)	3.46 (1.830)	2.63 (0.023)	2.89 (0.031)
HP/Curb Weight (<i>tons</i>)	-0.516 (0.020)	0.6 (0.042)	0.608 (0.063)
Curb Weight (<i>tons</i>)	0.213 (-0.019)	0.341 (0.003)	0.361 (0.001)
Std. Deviations ($\sigma\beta$'s)			
Vehicle Cost (\$'000)		0.54 (0.460)	1.08 (0.349)
HP/Curb Weight (<i>tons</i>)		4.05 (0.020)	5.866 (2.022)
Operating costs		2.066 (0.001)	2.085 (0.295)
γ/α	-0.22	2.08	1.84
# Observations	1,470	1,470	1,470

Notes: Standard errors in parentheses.

Table 5: Cost Side Parameter Estimates

Variables	Parameter Estimates	Standard Error
Constant	1.025	(0.194)
ln HP/Curb Weight (tons)	0.477	(0.003)
ln Curb Weight (tons)	1.03	(1.230)
ln(MPG)	0.423	(0.029)
ln(Size)	-0.048	(0.063)

Table 6: Robustness of Consumer Valuation of Fuel Economy

	Ratio $\Upsilon/\alpha=1.88$		
	(1)	(2)	(3)
	T=10	T=15	T=25
Discount Rate=6%	1.540	1.170	0.890
Discount Rate=7%	1.620	1.250	0.980
Discount Rate=8%	1.770	1.460	1.150
Discount Rate=9%	1.690	1.330	1.060
Upper 95%	0.8-1.88		
Lower 5%	0.51		

CONCLUSION

Policy makers have employed various policy instruments in an attempt to achieve at least two policy goals. One goal is to reduce peak demand, obviating the need for a large excess generating capacity which is rarely used. A second policy goal is to reduce overall electricity consumption, thereby conserving fossil fuels and reducing pollution. However, behavioral response interacts with these policies in a complex manner sometimes resulting in unintended consequences. Therefore an understanding of behavioral response is essential in designing effective energy conservation policies.

This dissertation provides answers to several questions related to unanticipated ways in which consumer behavior has affected the outcomes of energy policies. First, I provide evidence from a field experiment on a demand response program that investments in energy efficiency results in significant reduction of anticipated energy consumption while at the same time reducing resultant emissions. Such reductions offer utilities avenues for managing summer peak load without the necessity to build costly peaking power plants. However, due to strategic behavior and income substitution effects, households utilize these savings in different ways. I provide evidence of direct rebound and precooling effects that an in the order of 41% of saved energy, monetary and environmental benefits. While the DR program was designed to create a flatter load shape for the utility to help reduce operating needs and complexity during peak periods, the benefits of this program are in part offset by behavioral response by consumers through precooling and rebound. That is the program results in reducing peak load and emissions reduction, but not as much as one would predict

because of behavioral response. This implies that behavioral response needs to be taken into account if we are to have realistic forecasts of potential project outcomes. It also implies that policies might be more effective if they are designed to create incentives for behavioral response that are consistent with the goals of the policy, so that behavioral response reinforces the program. This is really the notion of incentive compatibility.

Second, due to the changes in consumers' consumption patterns that occur as result of strategic behavior, utilities incur additional production costs which are then passed through to consumers. The rebound and precooling results in an increase in energy supply services, with a corresponding increase in effective prices, the size of which depends on the underlying cost structure. Therefore utilities in response to increased load during off-peak periods call more expensive generation units to meet increased load. This results in an increase in the cost of conservation which may have negative effect on the overall social welfare.

Third, one of the key barriers to achieving energy efficiency is the so-called "energy paradox," the idea that consumers undervalue cost savings from investments in energy efficiency. The undervaluation of future fuel costs in the automobile market is a great example for demonstrating such market failure. However, it has always been challenging to accurately quantify the degree of the undervaluation, if it exists. Consumer heterogeneity and sorting are important factors that must be considered in order when quantifying the level of consumer undervaluation of investments in energy efficiency. If consumers tend to undervalue long-term cost savings associated with

efficient equipment, then this will reduce the effectiveness of some policies relative to others (e.g., carbon tax versus gas guzzler tax). I find that accounting for sorting bias and consumer heterogeneity explains a large fraction of the purported myopia found in literature and conclude that myopic behavior may be a less significant problem; and also that it may be confined to a subset of the population.

Fourth, CAFE standards are designed to improve gas mileage. However, CAFE as structured is less stringent for larger vehicles therefore may be behind the explosion in the demand for light duty trucks (SUVs, Minivans, Pickup-trucks, etc). This explosion in the light truck share of the automobile fleet has introduced “prisoner's dilemma” situation where drivers buy heavy vehicles for better self-protection in collisions and increasing risks to occupants of other drivers. These externalities have serious policy implications. In theory, an efficient fleet could be achieved through a carefully calibrated tax based on careful cost-benefit analysis. However, such a tax policy may pose practical problems of implementation in terms of public acceptance and monitoring. A policy such as the proposed one will also lead to the adoption of more fuel efficient vehicles over time and reduction in the weight of the fleet.

Finally, estimates of emissions reduction impacts of energy conservation programs show that they result in reducing emissions but as one might expect because behavioral response will at least in part offset potential emission reductions from efficiency-based policies. This suggests we can improve the performance policies by integrating incentive-based programs such as carbon taxes or other tools policies that

increase energy prices. This dissertation reviews these policies, and concludes that while these policies do contribute to their major objectives, the benefits are at least in part offset by unintended consequences due to behavioral response by consumers. The main take away from this dissertation is that an understanding of behavioral response is essential in designing effective policies.