UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Essays in Environmental Economics

Permalink

https://escholarship.org/uc/item/8c01z6q4

Author

Liao, Yanjun

Publication Date

2019

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays in Environmental Economics

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Yanjun Liao

Committee in charge:

Professor Mark Jacobsen, Chair
Professor Julie Cullen
Professor Joshua Graff Zivin
Professor Gordon McCord
Professor David Victor

2019

Copyright
Yanjun Liao, 2019
All rights reserved.

The dissertation of Yanjun Liao is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California San Diego

2019

TABLE OF CONTENTS

Signature Page . . . iii

Table of Contents . . . iv

List of Figures . . . vii

List of Tables . . . viii

Acknowledgements . . . x

Vita . . . xi

Abstract of the Dissertation . . . xii

Chapter 1 Weather and the Decision to Go Solar: Evidence on Costly Cancellations . . . 1
- 1.1 Introduction . . . 2
- 1.2 Weather in the Decision to Go Solar . . . 6
  - 1.2.1 An Illustrative Model on Projection Bias . . . 6
  - 1.2.2 Channels: Solar Production vs. Increasing Rate . . . 7
- 1.3 Policy Background and Data . . . 9
  - 1.3.1 The California Solar Initiative (CSI) . . . 9
  - 1.3.2 Purchase and Cancellation . . . 12
  - 1.3.3 Solar and Weather Data . . . 14
  - 1.3.4 Summary Indices for Solar Production and Energy Demand 15
  - 1.3.5 Economic Controls and Demographic Characteristics . . . 17
- 1.4 Econometric Framework . . . 17
- 1.5 Results . . . 19
  - 1.5.1 Do Cancellations Respond to Weather? . . . 21
  - 1.5.2 Do Sign-ups Respond to Weather? . . . 25
  - 1.5.3 Response to Weather Updates . . . 26
  - 1.5.4 Heterogeneous Effects . . . 29
- 1.6 Mechanisms . . . 35
- 1.7 Conclusion . . . 37

Chapter 2 How Hurricanes Sweep Up Housing Markets: Evidence from Florida . . . 40
- 2.1 Introduction . . . 41
- 2.2 Data and Background . . . 46
  - 2.2.1 Florida Housing Market . . . 47
  - 2.2.2 Hurricane History and Exposure . . . 49
  - 2.2.3 Home Mortgage Disclosure Act . . . 52
- 2.3 Econometric Framework . . . 53
  - 2.3.1 Housing Price Model . . . 53

2.3.2 Transaction Probability Model . . . 57
2.3.3 Variation in Hurricane Exposure . . . 59
2.4 Results . . . 60
2.4.1 Post-Hurricane Price Dynamics . . . 61
2.4.2 Post-Hurricane Dynamics of Transaction Probability . . . 63
2.4.3 Mechanisms for Price and Transaction Probability Dynamics 66
2.4.4 Implications for Population Turnover . . . 71
2.4.5 Heterogeneous Effects . . . 75
2.5 Discussions and Conclusion . . . 76

Chapter 3 Extreme Weather and the Politics of Climate Change: a Study of Campaign Contributions and Elections . . . 79
3.1 Introduction . . . 80
3.2 Data . . . 86
3.2.1 Database on Ideology, Money in Politics, and Elections . . . 86
3.2.2 ActBlue . . . 87
3.2.3 League of Conservation Voters Scorecard . . . 89
3.2.4 Weather Shocks . . . 91
3.2.5 Natural Disasters . . . 92
3.3 Empirical Framework . . . 93
3.3.1 Short-Run Weather Impacts . . . 93
3.3.2 Medium-Run Natural Disaster Impacts . . . 96
3.3.3 Medium-Run Weather Impacts . . . 98
3.4 Results . . . 99
3.4.1 Short-Run Weather Impacts . . . 99
3.4.2 Medium-Run Natural Disaster Impacts . . . 109
3.4.3 Medium-Run Weather Impacts . . . 116
3.4.4 Mechanisms and Limitations . . . 118
3.5 Conclusion . . . 120

Bibliography . . . 123

Appendix A Chapter 1 Appendix . . . 135
A.1 Additional Figures . . . 135
A.2 Additional Tables . . . 140
A.3 Kleissl (2013) PV Performance Model . . . 155
A.4 Late Cancellation Reporting . . . 156

Appendix B Chapter 2 Appendix . . . 161
B.1 Additional Figures . . . 161
B.2 Additional Tables . . . 169
B.3 Determining Hurricane Exposure . . . 178
B.3.1 Imputing Maximal Reach Radius of 96 Knots Wind Speed . 178

B.3.2 Hurricane Track Interpolation and Exposure Calculation . . 179
B.4 Zillow-HMDA Matching Procedure 181
B.5 A Model on Income and Location Valuation 182

Appendix C Chapter 3 Appendix 186
C.1 Additional Figures 186
C.2 Additional Tables 188
C.3 Additional Specifications in the Short-Run Analysis 196

LIST OF FIGURES

Figure 1.1: Benefit from Solar PV System 8
Figure 1.2: Timeline of a CSI application 10
Figure 1.3: Distribution of Complete and Cancellation Duration 12

Figure 2.1: Florida Housing Market Sales and Composition 48
Figure 2.2: Exposure to Hurricanes by Census Tracts, 1992-2017 50
Figure 2.3: Severe Exposure by Census Tracts, 1992-2017 51
Figure 2.4: Hurricane Effects on House Prices – Full Sample 62
Figure 2.5: Hurricane Effects on House Prices – Repeat Sales 62
Figure 2.6: Hurricane Effects on Transaction Probability by House Type 64
Figure 2.8: Changes in Characteristics of Transacted Homes – Effective Age 68
Figure 2.9: Changes in Characteristics of Transacted Homes 70
Figure 2.11: Hurricane Effects on House Prices – HMDA Sample 72
Figure 2.12: Hurricane Effects on Buyer Income – HMDA Sample 72
Figure 2.13: Heterogeneous Effects by Hurricane Intensity – Full Sample 76

Figure 3.1: LCV Score Distribution by Party Affiliation 90
Figure 3.2: Variations of Estimates Across Quarters in Election Cycle 108

Figure A.1: CSI Incentive Step Design 135
Figure A.2: Monthly Peak Demand in California, 2006-2011 136
Figure A.3: Support for Linear Specification 137
Figure A.4: Harvesting Effects in Sign-ups 138
Figure A.5: Density of Days to Next Incentive Step Change 139
Figure A.6: Monthly Distribution of Reported Cancellations 158
Figure A.7: Monthly Distribution of Reported Cancellations by Late Status 158
Figure A.8: Full Distribution of Reported Cancellation Duration 159
Figure A.9: Distribution of Reported Cancellation Duration by Program Administrator . 159

Figure B.1: Florida Borrowers Market Price and Sales Share 161
Figure B.2: Demographics in Florida Counties 162
Figure B.4: Family and Housing Statistics in Florida Counties 163
Figure B.6: Hurricane Effects on House Prices – Borrower Sample 164
Figure B.7: Florida Population by County Aggregation 165
Figure B.8: Hurricane Effects on the Share of Transactions Involving a Mortgage . . . 166
Figure B.9: Heterogeneous Price Effects by Hurricane Intensity – HMDA 167
Figure B.10: Heterogeneous Income Effects by Hurricane Intensity – HMDA 168
Figure B.11: Hurricane Track Interpolation 180

Figure C.1: Example of an ActBlue Donation 186
Figure C.2: Distribution of Individual ActBlue Donation Amounts 187

LIST OF TABLES

Table 1.1: Summary Statistics . . . 20
Table 1.2: Cancellation and Solar Insolation . . . 23
Table 1.3: Channel: Solar Production vs. Energy Demand . . . 24
Table 1.4: Pre-period Weather and Purchase . . . 27
Table 1.5: Responses to Weather Updates . . . 28
Table 1.6: Cancellations of Non-residential Applications . . . 30
Table 1.7: Peer Effects in Cancellations and Response to Weather . . . 32
Table 1.8: Heterogeneous Effects by System Characteristics . . . 33

Table 3.1: Summary Statistics . . . 100
Table 3.2: Actblue Donation Responses to Short-run Temperature Shocks . . . 102
Table 3.3: Heterogeneous Effects by Incumbent Characteristics . . . 106
Table 3.4: The Effects of Natural Disasters on Amount Raised ($1,000) . . . 112
Table 3.5: The Effects of Natural Disasters on the Number of Donors . . . 114
Table 3.6: The Effects of Natural Disasters on Elections . . . 117

Table A.1: Variable Descriptions, Continued . . . 140
Table A.2: Cancellation and Solar Insolation (113 Days After Sign-up) . . . 142
Table A.3: Cancellation and Solar Insolation (60 Days Before Completion/Cancellation) 143
Table A.4: Robustness Check: Varying Controls . . . 144
Table A.5: Robustness Check: Zip-Code-Specific Time Trends . . . 145
Table A.6: Robustness Check: Logit Model . . . 146
Table A.7: Robustness Check: Varying Period Length . . . 147
Table A.8: Robustness Check: Different Subsamples . . . 148
Table A.9: Channel: Solar Production vs. Energy Demand (Alternative Time Frame) . 149
Table A.10: Attenuation Bias with Different Period Lengths . . . 149
Table A.11: Potential Selection by Pre-contract Weather (System Size) . . . 150
Table A.12: Responses to Weather Updates . . . 151
Table A.13: Heterogeneous Effects by Area Characteristics . . . 152
Table A.14: Does Current Weather Predict the Near Future? . . . 153
Table A.15: Weather and Complete Duration . . . 154
Table A.16: Top 30 Contractors by Number of Applications . . . 160

Table B.1: Summary Statistics for Housing Transactions . . . 169
Table B.2: Summary Statistics for Parcel-Year Panel . . . 170
Table B.3: Variations in Hurricane Indicators (Housing Price Model) . . . 171
Table B.4: Variations in Post-Hurricane Indicators (Transaction Probability Model) . . 172
Table B.5: Variations in Cat-3 Hurricane Indicators (Housing Price Model) . . . 173
Table B.6: The Impacts of Hurricanes on Prices and Transaction Probability (Event Study) 174
Table B.7: The Impact of Hurricanes on Housing Prices (Diff-in-Diff) . . . 175
Table B.8: The Impact of Hurricanes on Transaction Probability (Diff-in-Diff) . . . 176

Table B.9: Differential Capitalization 177
Table B.10: Wind Speed and Maximal Reach Radius Model 178

Table C.1: FEMA Disaster Declarations, 1990-2012 188
Table C.2: Predicting Total Democratic Donations Using ActBlue Donations 189
Table C.3: Positive and Negative Temperature Shocks on ActBlue Contributions 190
Table C.4: The Effect of Extreme Temperature Events on ActBlue Contributions 191
Table C.5: Heterogeneous Effects Across Quarters in Election Cycle 192
Table C.6: The Effects of Extreme Temperature on Amount Raised ($1,000) 193
Table C.7: The Effects of Extreme Temperature on the Number of Donors 194
Table C.8: The Effects of Extreme Temperature on Elections 195

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisors Mark Jacobsen, Julie Cullen, and Josh Graff Zivin. In each of their own ways, they have taught me how to think deeply about research questions and examine the evidence critically. Their unwavering support has been indispensable to me in times of difficulties and will always inspire me to aim higher. I am also grateful to Judd Boomhower, Richard Carson, Gordon McCord, Jeff Shrader, Charlie Sprenger, and David Victor, who have helped and advised me in varying ways. To my coauthors Yann Panassiè and Pablo Ruiz Junco, I would like to thank them for their great contributions to our joint projects, all the fun and not-so-fun-but-necessary conversations, their encouragements, and their friendship. I would also like to thank Becky Fraenkel and Sam Krumholz for their efforts in significantly lowering our common data struggle.

On a more personal level, I would also like to thank many friends here in San Diego including but not limited to Yaein Baek, Jonathan Desponds, Ying Feng, Onyi Lam, Ce Liu, Xueying Lu, Mariam Ordyan, Fanglin Sun, Gautam Reddy, Shihan Xie, Wenxin Xie, and Xu Zhang. I am most fortunate to have their company in what is known to be a lonely journey. I am also grateful to many other teachers and advisors who have influenced me since my childhood. In particular, I am deeply indebted to Bill Barron, who led me into the field of environmental economics and had been an amazing mentor to me until his passing. Most importantly, I am grateful to my parents for always encouraging me to pursue what I want, and for the enormous sacrifices they made for me.

Chapter 1, in full, is currently being prepared for submission for publication of the material. Yanjun Liao. The dissertation author was the primary author of this chapter.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Yanjun Liao; Yann Panassié. The dissertation author was a primary author of this chapter.

Chapter 3, in full, is currently being prepared for submission for publication of the material. Yanjun Liao; Pablo Ruiz Junco. The dissertation author was a primary author of this chapter.

VITA

2013	B. A. in Economics *(first class honor)*, University of Hong Kong
2013-2019	Teaching Assistant, University of California, San Diego
2019	Ph. D. in Economics, University of California, San Diego

FIELDS OF STUDY

Major Field: Environmental Economics, Public Economics

ABSTRACT OF THE DISSERTATION

Essays in Environmental Economics

by

Yanjun Liao

Doctor of Philosophy in Economics

University of California San Diego, 2019

Professor Mark Jacobsen, Chair

This dissertation contains three essays exploring how economic agents and markets respond to weather events and natural disasters.

Chapter 1 examines whether short-run weather fluctuations affect the decision to go solar by residential customers in California. The results show that residential customers whose sign-ups are followed by less sunny weather are more likely to cancel their contracts. This behavior is inconsistent with rationality, as a solar PV system is a long-term investment whose return is not affected by short-run weather fluctuations. I analyze this behavior further using a theoretical model of projection bias combined with engineering models of solar power production and energy demand.

Chapter 2 studies housing market responses to hurricanes using detailed data on Florida housing markets during 2000-2016. We identify the causal effects of hurricanes in a difference-in-differences framework exploiting the exogeneity of hurricane path and timing. The results show that the housing market is dominated by a negative supply shock lasting up to three years. While these effects are transitory, we show that they are associated with the arrival of home buyers whose income is higher, holding fixed the quality of homes transacted. This implies a lasting shift in the local economic profile towards higher income, along with potential gentrification.

Chapter 3 investigates whether and how voters' preferences on environmental policy are reflected in legislative elections. People's belief in climate change is known to be affected by experiencing extreme weather events. We study the electoral responses to these randomly occurring events in House of Representative elections. To focus on the channel of environmental preference, we estimate differential responses based on the environmental ideology of the incumbent congressperson conditional on party membership. We find evidence that climate-related extreme weather events motivate political giving to challengers of anti-environment incumbents, increase the overall competitiveness in these races, and lead to a lower probability of re-election for these incumbents.

Chapter 1

Weather and the Decision to Go Solar: Evidence on Costly Cancellations

Abstract

This paper studies the effect of short-run weather fluctuations on solar panel adoption in California. This decision appears to respond strongly to weather patterns associated with solar panel productivity: I find that customers whose sign-up for solar panels is followed by unfavorable weather are more likely to cancel their contracts. In contrast, non-residential customers are not subject to the same effect. Together, these results suggest that short-run weather conditions affect customers' valuation of solar panels. The most plausible mechanisms are psychological biases such as projection bias or a salience effect, leading the decision-maker to rely too heavily on transient conditions when predicting long-run utility. This paper is among the earliest to document evidence of behavioral anomalies in the solar market.

1.1 Introduction

Many important economic decisions require the consumer to estimate future utility. When purchasing a durable good, the consumer needs to trade off her future utility against the present cost of purchase. Standard economic theory assumes that these estimates are accurately and consistently made, but emerging evidence shows that psychological mechanisms can often interfere.[1] For example, people with projection bias tend to exaggerate the degree to which future state will resemble the current state. People may also be affected by the salience effect, where current conditions accentuate certain attributes of a good and lead to a change in valuation.[2]

In this paper, I study whether residential solar customers are affected by short-run weather in their decision to acquire a solar photovoltaic (PV) system. If the current weather is favorable for real-time solar power generation, we might see more sign-ups because the benefit of a solar PV system is more salient, while bad weather would have the opposite effect. Moreover, because solar panels are not immediately installed after sign-up, weather conditions during the waiting period may also affect the customer's valuation of the system and lead to cancellation. These implications, with a focus on cancellations, are empirically tested in this paper.

Using transaction-level administrative data from the California Solar Initiative (CSI), a statewide subsidy program, I find that residential customers are more likely to cancel their solar contracts when they experience worse weather after signing up. Specifically, a one-standard-deviation decrease in solar radiation is associated with a 7.1-9.8% increase in cancellations, which, in turn, account for 12.06% of all CSI applications. This effect is identified using variation within each season while controlling for individual system characteristics and monthly economic conditions. It is consistent with the aforementioned psychological mechanisms and robust across a variety of other specifications, including county-quarter-year fixed effects, month fixed

[1] See DellaVigna (2009) for a review.

[2] Empirical evidence on projection bias: Conlin, O'Donoghue, and Vogelsang (2007), Magistris and Gracia (2016), Busse, Pope, Pope, and Silva-Risso (2015), and Simonsohn (2010). Empirical evidence on salience effects: Bordalo, Gennaioli, and Shleifer (2013), Hastings and Shapiro (2013), and Kőszegi and Szeidl (2012).

effects, nonlinear models, and alternative weather controls.[3] Moreover, I find that non-residential customers are not subject to the same effect. This is likely because their decision-making process is less susceptible to psychological factors in general.

Solar PV investment has some unique features that make it difficult to reconcile the above results with rational mechanisms. First, solar panels typically last more than 25 years. Short-run weather fluctuations prior to installation do not affect the return to such a long-run investment. Second, solar companies commonly provide estimates of electricity savings based on micro-climate to potential buyers. Short-run weather fluctuations do not contain useful information in addition to those estimates. For these reasons, the results also cannot be explained by present bias unless the discount rate is very large and the weather is very persistent.

As an outcome of interest, the *cancellation* of a signed contract has a number of advantages. In terms of interpretation, a cancellation directly reflects that the customer decides ex post that the contract was a mistake. Projection bias can thus be directly tested using this decision margin, as discussed in the literature (Conlin, O'Donoghue, and Vogelsang, 2007; Busse, Pope, Pope, and Silva-Risso, 2015). Furthermore, the cancellation often entails a penalty which supports the interpretation of a change in valuation.[4] An obvious alternative is *purchase*: consumers with projection bias are more likely to purchase solar panels on sunnier days. It is tempting, therefore, to estimate whether there are more purchases on sunny days.[5] However, the implications of such a pattern of purchase are less clear than cancellation. For instance, a solar purchase might simply be due to increased awareness of solar power rather than higher valuation by the customer. Also, a sunny day might "harvest" solar purchases that would have occurred rationally, which implies minimal welfare impacts. In contrast, cancellations necessarily carry losses in overall welfare in the form of wasted effort by both the contractor and the customer in coordination, negotiation,

[3] By controlling for a rich set of fixed effects, I use the remaining random variations in weather to identify the causal effect on cancellations. This fixed effects framework is first used in Deschenes and Greenstone (2007) and later adopted in numerous papers in the climate-economy literature. See Dell, Jones, and Olken (2014) for a review.

[4] Section 1.3 provides more details on the cancellation process and the associated costs.

[5] A concurrent paper (Lamp, 2018) examines this outcome in Germany. As supplemental evidence, I also look at solar sign-ups and find results that are similar in direction and scale.

site inspection, and other contracting costs. Another key advantage of cancellation over purchase is data availability. Cancellation can be analyzed at the individual level, whereas purchase can only be examined at an aggregate level due to lack of information on all potential buyers.

I examine how current weather can impact a customer's decision by explicitly modeling her over-dependence on current conditions along two dimensions. The model allows weather to affect the customer's valuation by (1) a "solar production" channel that affects her prediction of future solar generation, and (2) an "increasing rate" channel where solar generation displaces electricity on different price tiers. Applying engineering models, I transform the weather variables into two indices that represent the two channels. I then estimate the effects of these indices jointly. The results show that cancellation responds most strongly to variations in short-run solar panel productivity. I also show evidence that negative updates in weather conditions relative to the pre-contract period play an important role, and the responses are particularly strong for customers who sign up following a period of above-average weather. These results lend further support to the psychological mechanisms.

Lastly, I explore heterogeneous responses along various dimensions. I first examine whether customers from neighborhoods with higher solar-market penetration respond differently to weather. I find that they are less likely to cancel, which is consistent with the peer effects identified in Bollinger and Gillingham (2012). However, they are not less responsive to weather. There are heterogeneous effects on other margins: the weather effect is larger for customers whose system is owned by a third party, whose neighborhood of residence is more urban, and whose sign-up was followed by an imminent decrease in the incentive rate.

This paper makes several contributions. It is closely related to a growing literature in behavioral economics that presents "smoking gun" evidence of projection bias and salience effects in the field (Busse, Pope, Pope, and Silva-Risso, 2015; Chang, Huang, and Wang, 2018; Conlin, O'Donoghue, and Vogelsang, 2007; Hastings and Shapiro, 2013; Lamp, 2018; Simonsohn, 2010). Rather than studying the effect of weather on the purchase decision as most previous studies do,

this paper highlights the importance of weather updates, especially negative ones, in changing the customers' valuation. Furthermore, because a solar PV system is a big-ticket item with an instrumental nature, the findings suggest that such biases are more pervasive and economically important than previously understood.

This paper also adds to the literature that documents the behavioral anomalies in energy efficiency investments. Consumers are often found to mis-optimize due to inattention, imperfect information, or myopia (Allcott, 2011; Allcott, 2013; Attari, DeKay, Davidson, and De Bruin, 2010; Busse, Knittel, and Zettelmeyer, 2013; Hausman, 1979; Ito, 2014; Sallee, 2014). The decision to invest in solar PV is, in many ways, similar to the decision for energy efficiency. Therefore, the behavioral mechanism identified in this paper is also relevant for any energy efficiency investment whose real-time productivity depends on transient conditions.

Such behavioral anomalies have profound policy implications. They provide the basis for paternalistic intervention in energy efficiency (Allcott, Mullainathan, and Taubinsky, 2014; Allcott, 2016). They can also be targeted in new behavioral programs to achieve much higher cost-effectiveness than traditional economic incentives (Allcott and Mullainathan, 2010; Gillingham, Keyes, and Palmer, 2018). These lessons also apply to solar policies. In the U.S., there are many solar subsidies at the federal, state, and local levels. However, prior evaluations show that these programs often cost more than the environmental and industrial benefits they generate.[6] It will be fruitful for future policy to incorporate insights from the behavioral sciences. As behavioral research on consumer choice regarding solar is limited, the current study is an early effort in that direction. The findings reveal opportunities to improve welfare by presenting potential customers with accurate information and cautioning them against over-dependence on short-run weather conditions, thereby reducing costly cancellations.

[6]Crago and Chernyakhovskiy (2017) and Hughes and Podolefsky (2015) both find that the cost of carbon mitigation in subsidy programs is much higher than the social cost of carbon. Bollinger and Gillingham (2014) find that non-appropriable learning-by-doing in the solar industry similarly do not justify the cost. Furthermore, Borenstein and Davis (2015) notes the distribution of benefits of such programs are highly regressive. More generally, see Baker, Fowlie, Lemoine, and Reynolds (2013) for a review on the cost and benefit of solar electricity.

1.2 Weather in the Decision to Go Solar

This section presents a simple model to illustrate how specific weather patterns might affect customers' valuation of solar panels. It is adapted from existing theory on projection bias, but the implications are applicable to other psychological channels where consumers are over-influenced by current conditions.

1.2.1 An Illustrative Model on Projection Bias

Projection bias refers to the tendency for people to exaggerate the degree to which the future state will resemble the current state. As modeled in Loewenstein, O'Donoghue, and Rabin (2003), this framework distinguishes between $u(G, s_1)$, the actual utility from using good G in period one (with state s_1), and $\tilde{u}(G, s_1|s_0)$, the "projected" utility based on the current period (with state s_0). A simple representation of over-influence by the current state is

$$\tilde{u}(G, s_1|s_0) = (1-\alpha)u(G, s_1) + \alpha u(G, s_0), \tag{1.1}$$

where α is between 0 and 1. When α is nonzero, the projected utility is different from the actual utility and biased towards the period-zero utility. At the current period, the projected total utility from purchasing a good that lasts from period 1 through T is given by

$$\begin{aligned} \tilde{U}(G|s_0) &\equiv \sum_{\tau=1}^{T} \delta^{\tau} \tilde{u}(G, s_\tau|s_0) \\ &= (1-\alpha)U(G) + \alpha\delta\Big(\frac{1-\delta^T}{1-\delta}\Big)u(G, s_0), \end{aligned} \tag{1.2}$$

where $U(G) = \sum_{\tau=1}^{T} \delta^{\tau} u(G, s_\tau)$ is the discounted sum of the stream of actual utilities. The second equality follows from substituting in Equation (1.1). Therefore, the current state biases the total projected utility.

In the context of solar adoption, I consider the relevant states as different weather con-

ditions. Households decide whether to go solar by weighing the projected total utility against a fixed cost. Therefore, although current weather does not affect the true utility ($U(G)$) at all, it affects the projected utility and hence the decision of the marginal customer. This has two testable implications:

(I1) Better weather leads to a higher probability of signing up for solar.

(I2) Worse weather after the sign-up leads to a higher probability of canceling the contract.

While we can test both implications, there are conceptual and practical differences between them. Section 1.3.2 provides a detailed discussion on these issues.

1.2.2 Channels: Solar Production vs. Increasing Rate

The return to a solar PV system in each period depends on the total electricity bill savings in that period, which depends on the amount of power it generates and the cost of the displaced electricity.[7] I assume that the current period utility equals to the reduction in electricity bill under net metering, which takes the following form:

$$u(G, s_0) = \int_{D(s_0)-E(s_0)}^{D(s_0)} P(v)dv. \tag{1.3}$$

$D(s_0)$ is gross electricity demand, $E(s_0)$ is solar power production, and $D(s_0) - E(s_0)$ is the net demand that eventually enters into the price schedule. $P(\cdot)$ is the marginal price schedule of electricity and is weakly increasing in electricity usage due to the increasing-block pricing scheme in California.[8] As illustrated in Figure 1.1, Equation (1.3) is essentially measuring the shaded area for given electricity demand D and solar power production E.

[7] Some solar customers might have "green" preference. The model accounts for such preference so long as "green" utility is proportional to solar power generation. Other forms of green preference are ignored for simplicity.

[8] See Borenstein (2012) for details on electricity pricing in California. This rate structure is a major option for electricity pricing. It is also adopted in other important solar markets such as China.

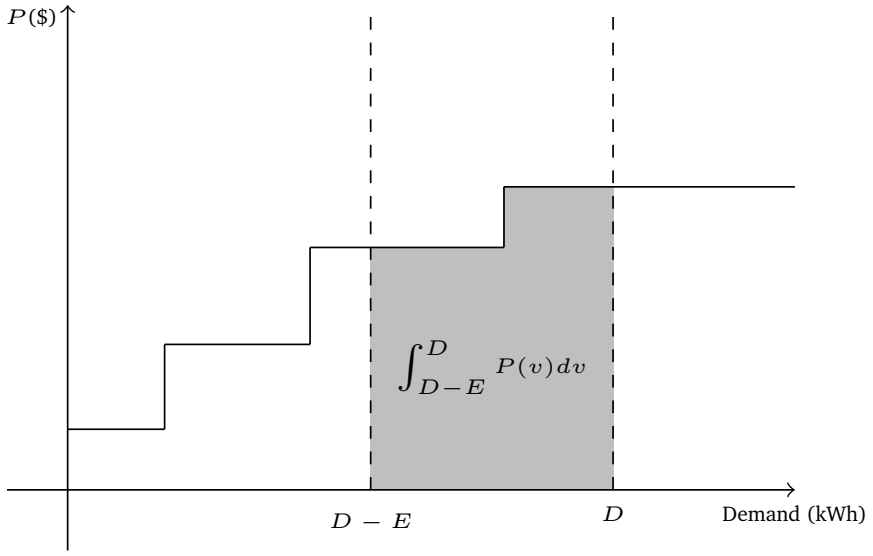


Figure 1.1: Benefit from Solar PV System

For simplicity, assume the scalar s_0 is a sufficient statistic that summarizes the state of the world in period zero. It is clear that s_0 affects the projected total utility through both $D(s_0)$ and $E(s_0)$. More formally, we can derive

$$\frac{d\tilde{U}(G|s_0)}{ds_0} = \alpha\delta\Big(\frac{1-\delta^T}{1-\delta}\Big)\Big[\underbrace{P|_{D(s_0)-E(s_0)} \cdot E'(s_0)}_{\textit{Solar Production Channel}} + \underbrace{\big(P|_{D(s_0)} - P|_{D(s_0)-E(s_0)}\big) \cdot D'(s_0)}_{\textit{Increasing Rate Channel}}\Big]. \quad (1.4)$$

Equation (1.4) highlights the two channels through which weather might affect the decision to go solar. The first term represents the solar production channel. When the current weather is unfavorable for solar power production, customers might underestimate the lifetime productivity of the solar PV system. The second term represents the increasing rate channel. During periods of greater energy demand, such as particularly hot or cold days, customers might overestimate their future energy demand. In turn, they will perceive greater savings from the solar panels under an increasing rate schedule.

Theoretically, the production channel will shut down when the solar PV system generates

sufficient power to offset all electricity consumption in a year.[9] On the other hand, the increasing rate channel will shut down when solar generation is not large enough to move the net demand to a lower price tier. In reality, however, these conditions are rare. It is worth noting that these two channels do not necessarily go in the same direction. For example, extreme heat reduces solar panel productivity but increases electricity demand for cooling. Another real-life complication is that customers have been shown in this context to respond to average rather than marginal price (Ito, 2014). We can account for this by replacing $P(\cdot)$ with a perceived price schedule that is constantly increasing. Under the new schedule, the qualitative result in Equation (1.4) remains valid.

1.3 Policy Background and Data

1.3.1 The California Solar Initiative (CSI)

The CSI is a solar rebate program in California for the customers of the three investor-owned utilities (IOUs) covering most of California. It funds solar on commercial, agricultural, government, and non-profit buildings as well as existing homes. Overseen by the California Public Utilities Commission (CPUC), the program had a total budget of $2.167 billion between 2007 and 2016 and a goal to install approximately 1,940 MW of new solar generation capacity. Over the years, California continues to be the leading solar market in the United States, accounting for about half of the nationwide installed capacity.

An important feature of the program is that the rebate rate is designed to decrease automatically based on the total volume of solar megawatts (MWs) with confirmed project reservations (see Figure A.1 in the Appendix). The exact timing of each step decrease is uncertain to solar market participants. This design is intended to encourage learning-by-doing and lower the cost of solar over time. During the sample period, the average cost has indeed been rapidly

[9] The billing cycle for California net metering customers is 12 months.

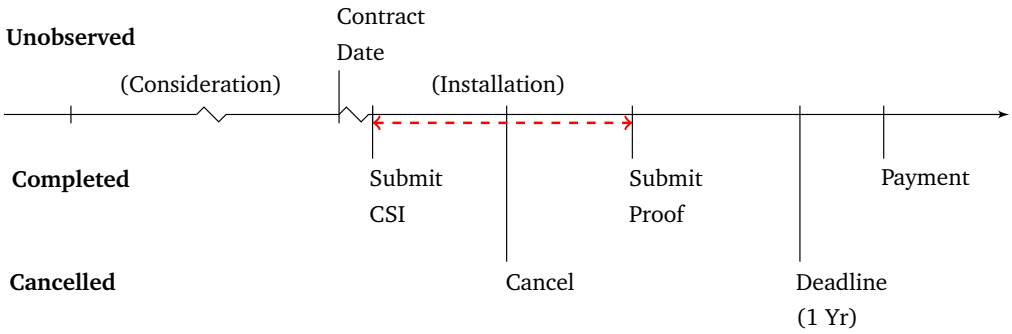


Figure 1.2: Timeline of a CSI application

decreasing.

The main data source in this paper is the administrative dataset from the CSI. It contains the universe of CSI applications, regardless of whether they are eventually completed. They represent about 80% of all solar projects in California before 2012 and much less afterward, as the incentive weakens. For each application, the dataset provides detailed information on system characteristics such as the CSI rating (generation capacity), total cost, incentive amount, the sectors for the host and the owner, and status of third-party ownership.[10] More importantly, it contains the zip code and relevant dates for each application, which can be matched to meteorological data at a fine geographical scale.

Figure 1.2 shows the typical timeline for a CSI application, where unobserved events are plotted above the line and observed ones below. The process begins with the household spending an unknown amount of time to consider going solar. In this process, they might be in contact with one or more solar contractors. Each contractor provides some relevant information such as estimated bill savings, design options, and quotes. The consideration period ends when the household signs a contract. This sign-up date is not observed. However, CSI records the date when a rebate application is electronically submitted, at which point a signed contract is required

[10]The sectors include residential, commercial, government or non-profit. The host is usually the customer, or where the system is physically located. Third-party ownership (TPO) means that the system is not owned by the host but a third party such as a solar developer or investor.

by program rules. The application might be submitted by either the homeowner or the contractor, but the subsidy is mainly captured by the homeowner (Pless and Benthem, forthcoming). Because a rebate is reserved for the homeowner at the current rate upon submission, the homeowner has the incentive to submit as soon as possible. This makes the submission date an effective proxy for the sign-up date.[11]

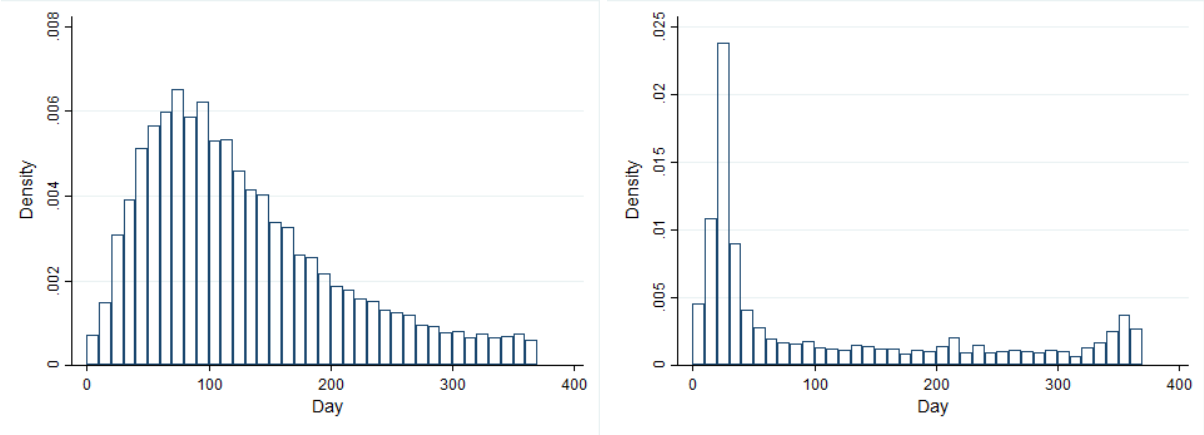
The majority of the applicants completed the installation and submitted the proof within one year. For these projects, I calculate the complete duration as the number of days between the initial application and the proof submission. On the other hand, 12.06% of the applications ended up being canceled.[12] Likewise, a cancellation duration can be calculated as the period between the initial application and the recorded cancellation date.

In Figure 1.3, I plot separate distributions for complete duration and cancellation duration. Because applications are subject to a one-year deadline, observations beyond one year are excluded. The distribution of complete duration is single-peaked, relatively dispersed and smooth with a 96.2% valid rate, which is what we would expect if most dates are reported accurately. On the other hand, the distribution of valid cancellation duration has a highly concentrated peak around a month, but it accounts for only 36.4% of all cancellations. In other words, the majority of canceled projects have a recorded cancellation date beyond one year.

When a project is canceled, there is no incentive for either the solar company or the customer to report to the CSI administrator on time. Thus, many cancellations might have gone unreported until the deadline, when the administrative process requires the project status. Indeed, when we look at the uncensored distribution of recorded cancellation duration (see Figure A.8 in Appendix A.4), there is strong evidence of bunching right after the deadline. Therefore, these dates are not likely to reflect true cancellation time frame accurately. In Appendix A.4, I investigate this reporting issue in more details and find that this is a pervasive issue across

[11] This is also confirmed by the author's communication with a CSI staffer.

[12] According to the author's interview with a major solar company, the cancellation rate observed in this dataset is lower than their own records. The primary cancellation reason for their customers is that further assessment lowers estimated capacity due to changes in parameters such as shading or azimuth.



(a) Complete Duration (within 1 year) (b) Cancellation Duration (within 1 year)

Figure 1.3: Distribution of Complete and Cancellation Duration

Notes: Panel (a) shows the distribution of completion duration, and panel (b) shows cancellation duration. Only observations before the 1-year installation deadline are included, which account for 96.2% of the 136,002 completed projects and 36.4% of the 18,653 canceled ones.

program administrators and contractors.

1.3.2 Purchase and Cancellation

The theory above generates testable implications of weather fluctuations on purchase (I1) and cancellation (I2). While both implications will be empirically tested, the focus will be on cancellation for a number of reasons.

First, cancellation can be analyzed at the individual level and its interpretation is clear. As first proposed by Conlin, O'Donoghue, and Vogelsang (2007) and later discussed by Busse, Pope, Pope, and Silva-Risso (2015), projection bias can be directly tested by finding evidence of consumers deciding *ex post* that a decision was a mistake. In contrast, the sign-up decision can only be examined in the aggregate. This is due to a data limitation: we only observe realized transactions but not the pool of potential buyers. As a result, instead of examining the binary purchase outcome, we test whether there are more purchases in a zip code when conditions become favorable. This is similar to prior research on health insurance take-up (Chang, Huang,

and Wang, 2018). However, it is difficult to interpret the effect as a change in customer valuation of solar panels, as this interpretation assumes that the pool of potential buyers does not change based on weather conditions. This assumption might fail if, for example, a sunny period increases the pool of potential buyers by reminding them of solar power. The test of (I1) does not rule out this possibility.

Second, cancellation reflects larger economic consequences than purchase. The California state law specifies a cool-down period of three business days for the customers to freely cancel their contracts. Outside of that period, a cancellation penalty can be charged legally, depending on the solar contractor and the progress of the installation.[13] Moreover, cancellation also means wasted time and effort by both the contractor and the customer. In contrast, a positive correlation between purchase and sunshine might simply reflect short-run postponement of sign-up with minimal welfare implications.

Because many of the recorded cancellation dates are invalid, we need to make alternative assumptions about the relevant time frame of the cancellation decision. I use two separate assumptions in the main analysis. First, I assume that the cancellation decision is made within the same time frame as installation. Thus, I use the median complete duration (113 days after the sign-up) as the relevant window for every project. Alternatively, the reported cancellation dates are more likely to be accurate if the duration is of a reasonable length (for example, less than 180 days). For this subset of observations, I assume the relevant time frame is the 30 or 60 days before the cancellation or completion date.

While both are imperfect proxies of the true decision time frame, the two assumptions complement each other in their strengths and weaknesses. In the first assumption, the 113-day period is consistent with the typical installation time frame specified in solar contracts. However, it is too long for some projects and too short for others. This introduces measurement errors in

[13] Anecdotes suggest that some companies impose no penalty if the work has not started, while others are notorious for their strict cancellation policy. The penalty could include the deposit, which is typically 500-1000 USD, a restocking fee if equipments have been ordered, and potentially more if the installation is in later stages. Some customers sued their contractors in order to exit from a contract, which incurs additional legal fees.

the relevant weather variables and attenuates the estimates. In the analysis, I test the robustness to alternative period lengths and whether the results are further attenuated as we move further away from the preferred length. The second assumption might capture the true decision time frame better. There are reasons to believe that more recent experiences of weather are more important for the decision, such as an increased urgency of the decision and recency bias. However, this assumption excludes the majority of cancellations in the data while the first allows us to examine all applications. Thus, the second assumption delivers stronger evidence while the first provides a more complete picture.

Besides cancellation, I also examine the purchase decision. In particular, I am concerned that favorable weather might lead some customers to overestimate the profitability of solar and sign contracts that will eventually get canceled. In that case, mean-reverting weather trends might introduce mechanical correlations between cancellation and post-contract weather. While this phenomenon is still consistent with projection bias, the main estimates on cancellation and the welfare consequences will be harder to interpret. In the empirical analysis, I address this issue by examining whether favorable weather is associated with more sign-ups or larger systems. I also explore how the dynamics of pre- and post-period weather jointly affect cancellations.

1.3.3 Solar and Weather Data

The weather data in this paper comes from two main sources. I obtain station-level data on daily maximum temperature and wind speed from the Global Surface Summary of the Day (GSOD) dataset at NOAA. I exclude weather stations with elevation above 1,500 meters, which leaves 191 stations operating in the sample period to be matched with 1,428 zip codes. Based on the center coordinates of each zip code, I find all the weather stations within 100 kilometers (62 miles) and calculate the mean of observations from these stations, weighted by inverse distance. These measures might not be precise for each application. This problem is less severe for temperature – which is more smoothly distributed across space – but more severe for wind

speed, which can be highly localized.

Surface solar radiation is a key weather element for solar power generation. I obtain daily observations of solar insolation from the Prediction of Worldwide Energy Resource (POWER) Project at NASA.[14] This data is derived using satellite observations and provided at a 1° latitude by 1° longitude resolution. For each zip code, I extract data using its center coordinates. As a zip code is much smaller than a grid cell, it is likely to lie entirely in the same grid cell as its center coordinates.

Finally, I convert daily observations of weather into summaries over the relevant period for each application: average solar insolation, average wind speed, number of days with maximum temperature below 40°F and the same for above 100°F.

1.3.4 Summary Indices for Solar Production and Energy Demand

The weather affects both energy demand and solar production, and both are related to the value of a solar PV system. To understand which factor plays a bigger role, I construct a proxy for each of the two factors and analyze their effects on cancellation simultaneously. These measures are calculated using data for solar insolation, wind speed, and temperature. As regressors, they allow for easier economic interpretations compared to the original weather elements.

Increasing rate (energy demand) channel. Weather affects energy demand mainly through heating (cooling) in cold (hot) days. A common practice is to calculate heating and cooling degree days (HDD and CDD):

$$HDD = 1\{Temp < 65\} \cdot (65 - Temp) \quad \text{and} \quad CDD = 1\{Temp > 65\} \cdot (Temp - 65).$$

Heat is a major driver of electricity use in California, as air conditioners and fans all run

[14] Solar insolation is the amount of electromagnetic energy (solar radiation) incident on the surface of the earth, measured in MJ/m^2 per day. See Stackhouse Jr, Westberg, Hoell, Chandler, and Zhang (2015) for methodology and accuracy of this measure.

on electricity. On the other hand, cold weather might be less relevant for solar in California because of its mild winters. Furthermore, electricity is not the primary choice for heating in Californian homes, accounting for only 8.5% of the heating demand (EIA, 2016). As a result, the time series of peak demand in the California ISO system is characterized by large spikes in the summer (see Appendix Figure A.2). This suggests that CDD is a good proxy for energy demand while HDD is not. I also define total degree days (TDD), the sum of CDD and HDD, as an alternative to further account for the smaller spikes in the winter. Finally, for each CSI application i, I calculate the daily CDD and TDD and then average them over the relevant period (length T):

$$\overline{CDD}_i = \frac{1}{T}\sum_{\tau=1}^{T} CDD_{i\tau} \quad \text{and} \quad \overline{TDD}_i = \frac{1}{T}\sum_{\tau=1}^{T} TDD_{i\tau}.$$

Solar production channel. Real-time solar generation is a nonlinear function of several weather elements. Solar energy generation is primarily affected by the intensity of solar radiation, but the conversion rate is also affected by heat negatively and wind positively. I use a PV performance model developed by Kleissl (2013).[15] This model simulates solar panel output for given inputs of weather and system characteristics. For each CSI application i, I generate a series of daily indices by plugging in daily weather variables:

$$ProdIndex_{i\tau} = F(Solar_{i\tau}, Temp_{i\tau}, WindSpeed_{i\tau}), \text{ where } \tau = 1, ..., T.$$

Similar as before, I calculate the average in the relevant period:

$$\overline{ProdIndex}_i = \frac{1}{T}\sum_{\tau=1}^{T} ProdIndex_{i\tau}.$$

Since I keep system characteristics the same across applications, the production index measures how favorable the current weather is for solar power generation.

[15] See Appendix A.3 for more information on the model.

1.3.5 Economic Controls and Demographic Characteristics

Variations in weather are usually orthogonal to most other variables, but it is less certain in this study because weather variables are averaged over a period. Therefore, I also control for economic conditions that might drive solar demand. I collected the following monthly economic variables: leading index and unemployment rate from the FRED database[16], the prime interest rate, and index of consumer sentiment and buying conditions for the West region from the Survey of Consumers. The CSI dataset does not provide individual demographics. Instead, I collect zip-code average demographics including race, education, income, housing cost, household size, etc. They are either from the 2010 Decennial Census or 5-year estimates from the 2011 American Community Survey (ACS). I use them to analyze whether responses to weather are heterogeneous across different communities.

1.4 Econometric Framework

The main analysis focuses on testing implication (I2), or how weather fluctuations in the post-contract period affect the probability of cancellation. The baseline estimating equation is a linear probability model[17]:

$$1(Cancel)_{izt} = \beta_1 Solar_{izt} + \beta_2 Weather_{izt} + \beta_3 System_i + \beta_4 Econ_{zt} + \delta_z + \delta_y + \delta_q + \varepsilon_{izt}, \quad (1.5)$$

where i denotes an application (project), z denotes the zip code, and t denotes the application date. $1(Cancel)_{izt}$ is an indicator of whether project i is eventually canceled. $Solar_{izt}$ is the average solar insolation in the relevant period for project i. $Weather_{izt}$ is a vector of temperature and wind

[16]The leading index for each state predicts the six-month growth rate of the state's coincident index, which in turn measures the current state of economic activity.

[17]Nonlinear models such as Probit or Logit are valid alternatives. However, a Probit model with zip code fixed effects might yield inconsistent estimates due to the incidental parameter problem (Greene, 2004). Results based on Logit models are reported in the Appendix, but marginal effects cannot be calculated because fixed effects are included.

speed variables to which households might also respond, corresponding to the same post-contract period. The coefficients on both solar and weather variables are of interest. $System_i$ is a vector of system characteristics in each application, including CSI rating (capacity), third-party ownership, and out-of-pocket unit cost. $Econ_{zt}$ are economic conditions in the month of date t, including the leading index, unemployment rate, prime interest rate, indices of consumer sentiment and buying conditions. δ_z, δ_y and δ_q denote zip code, year and quarter-in-year fixed effects, where the year and quarter are associated with the application date.

This specification incorporates a number of important considerations. First, it includes the relevant solar and weather variables simultaneously. Solar insolation is technically the most important factor in solar power generation and likely the most salient for customers. Wind speed and extreme temperature bins are also technically related to solar production and energy demand. Including them simultaneously addresses the concern of confounding weather patterns (Auffhammer, Hsiang, Schlenker, and Sobel, 2013). It also allows us to jointly examine the solar production channel and the increasing rate channel.

Secondly, this specification includes a rich set of fixed effects (FE) to account for non-weather drivers of cancellation that might be correlated with the weather. The zip code FEs control for cross-sectional correlations between weather amenities and neighborhood characteristics on a fine scale. The year FEs control for any statewide economic and policy shocks as well as climatic cycles such as the El Niño Southern Oscillation. The quarter-of-year FEs control for seasonal labor market conditions, which might lead to changes in demand for solar PV systems.

There is a trade-off in the choice of FEs: a richer set of FEs helps to rule out confounders, but it also eliminates meaningful variations. Adding more FEs might also exacerbate the attenuation bias due to measurement errors in the weather variables. To examine this trade-off in a transparent manner, I present four other specifications in the main analysis as refinements of the baseline. In my preferred specification, I control for seasonal economic dynamics specific to the county by quarter-by-county FEs, and any trend in solar PV installations – or common

economic and policy shocks to the entire state – by year FEs. Another specification includes county-quarter-year FEs, which addresses the concern that local policy or economic shocks might be correlated with the weather deviations. I also test the robustness of the results by adding zip-code-specific linear time trends where appropriate.

To further enhance identification, I control for economic conditions at the monthly level. I also include individual system characteristics as they might drive cancellation. In the later analysis of channels and other outcomes, I use the same set of FEs and controls because the threats to identification are similar.

Lastly, since there are overlaps in the weather stations matched with different zip codes, the standard errors are clustered by county in most tables. The number of clusters is 54.

1.5 Results

Table 1.1 presents the summary statistics. For the main analysis, I exclude applications whose host (customer) sector is non-residential, whose CSI rating is above 100 (less than 0.03%), and whose application submission starts after September 1, 2015. After cleaning, there are a total of 154,655 residential applications, of which 18,653 (12.06%) were eventually canceled. The average system cost is $35,377.04, which is more expensive than most home appliances. The average incentive amount is about one-tenth of system cost.

Canceled projects are larger and more expensive than completed ones. They are slightly more likely to be third-party-owned (TPO), that is, hosted by the customer but owned by the solar company. Canceled projects also appear to be located in zip codes with a slightly smaller fraction of whites and college graduates, and with lower median income.

In the formal analysis below, I first present a set of results and robustness checks on Equation (1.5). This establishes that residential customers indeed respond to weather fluctuations in the post-contract period. Then, I extend the analysis by further disentangling the solar production

Table 1.1: Summary Statistics

	All Applications		Completed	Canceled
Variable	Mean	SD	Mean	Mean
A. Application Characteristics				
CSI rating	4.86	3.02	4.82	5.22***
total cost ($)	35,377.04	23,829.66	35,090.3	37,467.7***
incentive amount ($)	3,551.24	4,838.65	3,526.86	3,729.01***
net unit cost ($)	6,897.71	3,179.06	6,901.94	6,866.86
third-party ownership (TPO)	0.500	0.500	0.495	0.541***
year	2011.11	1.83	2011.09	2011.27***
N	154,655		136,002	18,653
B. Weather				
solar insolation	18.96	5.85	18.98	18.81***
max temp (F)	74.50	9.40	74.53	74.24***
#days(max $\leq$ 40)	0.0160	0.214	0.0157	0.0177
#days(max $>$ 100)	3.24	9.50	3.26	3.14
wind speed (knots)	4.99	1.48	5.00	4.93***
#(precipitation$>$ 0)	23.43	17.46	23.43	23.45
production index	98.51	31.45	98.63	97.66***
cooling degree days	2.81	3.94	2.82	2.76*
heating degree days	5.92	4.74	5.91	5.97
total degree days	8.73	3.79	8.73	8.73
C. Economic Conditions				
leading index	1.41	1.42	1.39	1.55***
unemployment	10.02	1.78	10.01	10.15***
prime interest rate	3.56	1.04	3.57	3.46***
index of consumer sentiment	74.2	8.13	74.19	74.27
index of buying condition	128.07	13.12	128.07	128.17
D. Zip Code Demographics and Interconnection Characteristics				
white (%)	67.30	15.88	67.53	65.65***
bachelor degree (%)	36.15	17.87	36.29	35.12***
median income ($)	79,414.95	27,052.83	79,617.02	77,940.58***
monthly housing cost ($)	2,170.04	655.63	2,172.93	2,148.93***
mean household size	2.86	.460	2.85	2.88***
urban (%)	91.5	18.7	91.5	91.7*
current installed base	198.68	191.54	200.94	182.12***
installed penetration	16.61	15.06	16.77	15.38***

Notes: The weather variables in panel B are calculated over the 113-day period following the application submission. See Table A.1 for a description of the variables. Statistical significance of the differences between the completed and canceled subsamples are indicated in the last column.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

and increasing rate channel, and by examining the dynamics of purchase and cancellation in response to weather shocks before and after the contract. I also explore potential heterogeneous responses from residential vs. non-residential systems, and across a host of different system and neighborhood characteristics. Last but not least, I discuss the implications of the empirical results.

1.5.1 Do Cancellations Respond to Weather?

In this section, I estimate Equation (1.5) to examine whether solar customers respond to weather shocks. The primary variable of interest is the mean solar insolation in the relevant time frame after sign-up. For the same period, I also include the following weather variables: the number of days with maximum temperature below 40°F, the same for above 100°F, and average wind speed. Including them simultaneously addresses the concern of confounding weather patterns and provides a first sense of the relative importance of the solar production vs. the increasing rate channel.

As discussed in Section 1.3.2, I analyze two mutually-nonexclusive time frames: (a) 113 days after the sign-up and (b) 30 or 60 days before the completion/cancellation. Table 1.2 reports these results. Columns (1)-(2) present estimates under time frame (a), columns (3)-(4) show those under time frame (b) with a 60-day period, while (5)-(6) shorten the period to 30 days. In each pair of columns, the first column uses the baseline specification with zip code, year, and quarter-of-year fixed effects (FE), and the second allows the quarter-of-year FEs to differ across counties. As discussed in Section 1.4, the latter is my preferred specification because it is well-balanced between keeping meaningful variations and eliminating threats to identification.

The coefficient on solar insolation is negative and statistically significant throughout. This implies that a residential customer exposed to less sunshine after sign-up is more likely to cancel the contract. The baseline estimate suggests that a one-standard-deviation decrease in solar insolation is associated with a 7.1% increase in cancellations, while the cancellation rate is

12.06% among all applications.[18] The preferred estimate in column (2) shows a 9.8% increase. Despite using different variation in the data, the estimates in columns (3)-(6) are remarkably similar in scale to column (2). Other aspects of the weather also appear to play a role. Higher wind speed is positively associated with cancellation, and an additional hot day has the opposite effect. On the other hand, the effect of an additional cold day is not stable in either direction or scale. In California, such days are very rare. This might have been a major contributing factor to unstable results.

The key estimate on solar insolation is robust to three alternative fixed effects (FE) specifications: thethe inclusion of county-year FE, county-quarter-year FE, and month-of-year FEs (see Appendix Table A.2 for results under time frame (a), and Table A.3 for those under time frame (b)). It is also robust to varying the set of weather variables included in the model (Table A.4), and adding zip-code-specific annual or quarterly time trends (Table A.5). Finally, Table A.6 reports the results from Logit regressions with similar fixed effects. Although the fixed effects prevent valid calculations of marginal effects, the qualitative patterns are very similar and roughly proportional to the linear model.

I also look into the potential weaknesses of each time frame. The main concern for time frame (a) is that it might fail to capture the relevant period if the period length is not specified optimally. Table A.7 shortens the post-period length to 70 or 90 days. All of the results stand similar to before, as is expected from period lengths that are close to the optimal. On the other hand, time frame (b) relies on using a subsample with valid cancellation dates. In the main results, this is done by imposing a sample restriction to observations with a reported duration under 100 days. In Table A.8, I explore the effect of relaxing this restriction to include observations with duration up to 150 and 200 days. The qualitative pattern and statistical significance remain the same, but the scale of the estimates decreases. This is consistent with our expectation that these prolonged reports are less likely to be accurate. To briefly sum up, the results under time frame (a)

[18] Calculation: $\beta_1 \times SD(Inso)/D.V.Mean = 0.146 \times 5.85/12.06 = 7.1\%$.

Table 1.2: Cancellation and Solar Insolation

Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
solar insolation	−0.146	−0.203	−0.186	−0.192	−0.186	−0.205
	[0.055]**	[0.045]***	[0.045]***	[0.045]***	[0.048]***	[0.050]***
wind speed	0.262	0.799	0.815	0.754	0.617	0.593
	[0.241]	[0.271]***	[0.173]***	[0.162]***	[0.139]***	[0.143]***
#days(tmax $<$ 40)	1.055	1.406	−1.278	−1.572	−0.698	−0.691
	[0.383]***	[0.520]***	[0.489]***	[0.361]***	[0.450]	[0.490]
#days(tmax$\geq$100)	−0.020	−0.024	−0.256	−0.330	−0.149	−0.214
	[0.023]	[0.021]	[0.030]***	[0.065]***	[0.023]***	[0.048]***
Sample[a]						
Time Frame	First 113	First 113	Last 60	Last 60	Last 30	Last 30
Duration	All	All	0-100	0-100	0-100	0-100
D.V. Mean	12.06	12.06	6.70	6.70	6.70	6.70
R^2 (within)	0.033	0.035	0.072	0.076	0.071	0.074
N	154,519	154,518	62,069	62,062	62,069	62,062
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Notes: All coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are controlled for in all columns. Standard errors in squared brackets are by county. Full results and additional specifications are reported in Table A.2 and A.3. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.
[a]: Columns (1)-(2) use the full sample and a time frame of 113 days after the sign-up. Columns (3)-(4) restrict the sample to observations with a reported duration of less than 100 days and use a time frame of 60 days before the completion/cancellation. Columns (5)-(6) are similar but shortens the time frame to 30 days.

and those using the most reliable subsample under time frame (b) are consistent with each other.

Next, I further disentangle the two main channels outlined in Section 1.2.2. Specifically, I replace the solar and weather variables in Equation (1.5) with two indices corresponding to the solar production channel and the increasing rate channel, respectively. This specification directly estimates and compares the effects of the two channels following the logic of an encompassing test (Davidson and MacKinnon, 1993). The identifying assumption is that the other unobservable concerns are orthogonal to these two channels.

Table 1.3: Channel: Solar Production vs. Energy Demand

Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
$ProdIndex_{post}$	−0.761	−0.792	−0.650	−0.687	−0.830	−0.793
	[0.263]***	[0.259]***	[0.323]**	[0.362]*	[0.247]***	[0.252]***
CDD			−0.260	−0.223		
			[0.297]	[0.341]		
$TDD(=CDD+HDD)$					−0.126	−0.003
					[0.279]	[0.314]
D.V. Mean	12.06	12.06	12.06	12.06	12.06	12.06
R^2 (within)	0.034	0.035	0.034	0.035	0.034	0.035
N	154,519	154,518	154,519	154,518	154,519	154,518
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Notes: All indices are normalized. All coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

The construction of the two indices is detailed in Section 1.3.4. The index for solar production channel is generated from an engineering model. As expected, it is highly correlated with solar insolation ($\rho = 0.98$). As the increasing rate channel works through changes in energy demand, it is represented by either cooling degree days (CDD) or the sum of heating and cooling degree days (TDD). Both indices are normalized so that the estimates correspond to percentage point change in cancellation rate associated with an increase in the index by one standard deviation.

Table 1.3 reports these results under time frame (a). In columns (1)-(2) where the production index is the only regressor, its coefficients are negative and statistically significant. Columns (3)-(4) account for CDD, resulting in slightly smaller and less precise estimates on the production index. This might be driven by the correlations between the two indices. Nevertheless, they are still much larger in scale and more significant than the coefficients on CDD. The last

two columns switch to TDD to represent energy demand. The coefficients on TDD are even smaller, while those on the production index are similar to columns (1)-(2) in scale and statistical significance. Overall, these estimates suggest that the solar production aspect plays a greater role in the customers' decisions.

There are several possible explanations for the weak response through the increasing rate channel. For one, the channel does not apply to customers whose marginal electricity rate stays the same with or without the solar PV system. Moreover, customers tend to respond to average rather than marginal rate, which attenuates the rate increase. In fact, the increasing rate schedule might not even be salient in this decision. Lastly, for those customers who are primarily motivated by green preference, the billing aspect is simply unimportant.

These results are robust to using the the alternative time frame (b) (see Appendix Table A.9). Using the production index, I also examine the validity of using a linear specification. This is done by plotting the non-parametric relationship between cancellation and the index using a fitted local polynomial and a binned scatter plot (see Appendix Figure A.3). Both show a mostly linear relationship. Last but not least, I examine the attenuation bias regarding assumption (a) further. As discussed in Section 1.3.2, using a period length that is further away from the optimal point will result in larger attenuation bias and smaller estimates. I test this prediction by estimating the same model using alternative windows of 10, 30, 50, 70, and 90 days to construct the production index (see Appendix Table A.10). The estimates are increasing in the length of period and all except for the 90-day one are statistically different from the main (113-day) estimate. This pattern is consistent with the assumption that 113 days are close to the optimal length.

1.5.2 Do Sign-ups Respond to Weather?

This section tests implication (I1), that more favorable weather leads to an increase in purchase. This, in fact, is a recurring finding in the previous studies. I provide evidence in a

similar vein. Unlike cancellation, this outcome cannot be analyzed at the individual level as we do not observe all potential solar customers. The unit of analysis is zip-code-by-month instead. Moreover, since the relevant period for the sign-up decision is harder to infer, I show results with various pre-contract windows.

Table 1.4 shows these results with zip code, year, and month-in-year fixed effects. Columns (1)-(3) use the number of applications as the dependent variable, and columns (4)-(6) use its log. Three estimates are shown for each dependent variable, corresponding to an assumed window of 1, 2, and 3 months for calculating the production index.[19] The estimates are all positive and statistically significant for both level and log. When the index increases by one standard deviation, the level estimates suggest 0.6-0.9 additional applications and the log estimates suggest a 11.6-21.5% increase. I further test for harvesting effects by augmenting column (1) with one lead and six lags of the production index. Full harvesting implies that the positive effects of favorable weather should be followed by negative effects of similar magnitudes so that the net effect is zero. The results show that the positive effect peaks at a month later and reverses signs afterwards, but the overall effect is positive. This suggests that harvesting drives part of the dynamics in sign-up, but not entirely. These results align with Lamp (2018)'s main finding in Germany and complements the above findings on cancellation. On the intensive margin, favorable weather does not appear to induce systematic selections into larger system sizes (see Appendix Table A.11).

1.5.3 Response to Weather Updates

This section examines how the cancellation decision is affected by the dynamics of pre- and post-contract weather. Because cancellation reflects a change in valuation, it might be a

[19] The window starts from the current month. For example, column (2) uses the average production index in the current and previous month.

Table 1.4: Pre-period Weather and Purchase

(Zip-code-by-Month)	Number of Applications			log(Number of Applications)		
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-period Length	1Mo	2Mo	3Mo	1Mo	2Mo	3Mo
$ProdIndex_{pre}$	0.551	0.931	0.891	0.116	0.215	0.212
	[0.188]***	[0.194]***	[0.274]***	[0.034]***	[0.033]***	[0.047]***
D.V. Mean	3.36	3.36	3.36	-	-	-
R^2	0.410	0.410	0.410	0.481	0.482	0.481
N	45,875	45,858	45,768	45,875	45,858	45,768
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is the number of applications for columns (1)-(3), and its log for (4)-(6). The unit of analysis is zip-code-by-month. The length of the pre-period used to calculate the production index is indicated in the header. Monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

response to an update in post-period weather relative to the pre-period.[20] Therefore, analyzing the response to weather updates can provide insight into how customers are processing weather information at different points in time.

In Table 1.5, I examine whether the customers are responding to the change in post-period weather relative to the pre-period ("weather update" henceforth). I use a 30-day pre-period, a 113-day post-period, and the baseline FE specification in this analysis.[21] Column (1) estimates the relationship between cancellation and the pre-period production index. The coefficient is positive but insignificant. In column (2) I add weather update, which yields a negative and statistically significant coefficient. This suggests that conditional on the pre-period weather, customers respond to weather updates.

Column (3) tests whether a negative update has a different effect on cancellation than

[20]Two other studies have touched on this point. Conlin, O'Donoghue, and Vogelsang (2007) control for weather in both periods without explicitly examining updates. Chang, Huang, and Wang (2018) show evidence that the comparison is important.

[21]Using a 60-day pre-period yields qualitatively similar results (see Appendix Table A.12). Results under alternative FE specifications are also very similar and available upon request.

Table 1.5: Responses to Weather Updates

Cancel = 1	(1)	(2)	(3)	(4)	(5)
$ProdIndex_{pre}$	0.253	−0.472	−0.055		
	[0.235]	[0.313]	[0.287]		
$Update$		−0.802	−0.124	−0.257	0.368
		[0.269]***	[0.601]	[0.203]	[0.443]
$Indicator =$					
$1(Update < 0)$			−2.027		
			[0.404]***		
$1(ProdIndex_{pre} > \overline{Index_z})$				−0.383	
				[0.687]	
$1(ProdIndex_{post} < \overline{Index_z})$					0.437
					[0.351]
$Update \times Indicator$			−1.746	−0.556	−0.976
			[0.555]***	[0.314]*	[0.334]***
D.V. Mean	12.06	12.06	12.06	12.06	12.06
R^2	0.033	0.033	0.034	0.033	0.033
N	154,519	154,519	154,519	154,519	154,519
Fixed Effects					
Zip Code	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Quarter	Yes	Yes	Yes	Yes	Yes

Notes: $ProdIndex_{pre}$ is the normalized average production index for a 30-day pre-contract period. $Update$ is the average post-period index minus the pre-period one. $1(Update < 0)$ indicates whether $Update$ is negative. $1(ProdIndex_{pre} > \overline{Index_z})$ is an indicator of whether the pre-period index is higher than the zip-code average, and $1(ProdIndex_{post} < \overline{Index_z})$ is whether the post-period one is lower. All coefficients are multiplied by 100 for legibility. $Indicator$ in the interaction term in the last rows refers to the indicator variable that is included in the same regression. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

a positive one. The motivation is that a positive update might not be as informative given that these customers have already self-selected into a contract. I include an indicator of whether the update is negative and an interaction term of the indicator and the update variable, so that the latter will capture the differential effect of a negative update relative to a positive one. Indeed, the interaction term completely absorbs the effect of the $Update$ variable and becomes twice as large. This provides strong support for different information contents between positive and negative updates.

In the next two columns, I estimate a similar model but change the content of the indicator: it indicates whether the pre-period index is higher than zip-code average in column (4); and whether the post-period index is lower than zip-code average in column (5). Larger impacts are found on both margins. These results suggest that the customers fail to recognize the transient nature of weather shocks.

1.5.4 Heterogeneous Effects

This section explores whether responses are heterogeneous across a variety of customer and neighborhood characteristics. I classify each observation under one of two categories based on the characteristic of interest. I then estimate the differential responses by adding an interaction term of the production index and an indicator for one of the categories.

An informative margin of heterogeneity is residential vs. non-residential customers. The latter group – including small businesses, government agencies, and NGOs – are also eligible for the CSI rebate. Compared to residential customers, their decisions are more likely to involve formal profitability calculations and less prone to psychological effects. There are altogether 8,704 non-residential applications in the cleaned dataset, including 1,730 small systems ($<$10kW) and 6,974 large ones ($\geq$ 10kW). To apply for CSI rebate, the small systems follow the same 2-step procedure as residential projects while the large systems follow a 3-step process. Rather than a signed contract, CSI requires the payment of an application fee within 30 days of the initial

Table 1.6: Cancellations of Non-residential Applications

	All		Small Systems		Large Systems	
Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
$\beta_1 : ProdIndex_{post}$	−0.629	−0.669	−0.743	−0.779	−0.677	−0.714
	[0.260]**	[0.257]**	[0.265]***	[0.257]***	[0.255]**	[0.253]***
$\beta_2 : ProdIndex_{post}$	1.194	1.237	2.228	2.171	1.117	1.206
$\times 1(non\text{-}resid)$	[0.569]**	[0.567]**	[1.536]	[1.552]	[0.766]	[0.760]
$\beta_1 + \beta_2$	0.565	0.568	1.485	1.392	0.440	0.492
p-value: $\beta_1 + \beta_2 = 0$	0.38	0.36	0.34	0.38	0.59	0.54
R^2 (within)	0.057	0.058	0.040	0.041	0.055	0.057
N	163,059	163,058	155,947	155,946	161,336	161,335
Non-residential sample						
D.V. Mean	27.90	27.90	21.45	21.45	29.50	29.50
N	8,504	8,704	1,730	1,730	6,974	6,974
Fixed Effects						
Zip-type	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Notes: The regressions use pooled samples of residential and non-residential applications. Columns (1)-(2) include all non-residential applications, (3)-(4) include only small systems (<10kW), and (5)-(6) include only large systems ($\geq$ 10kW). The last two rows in the top panel show the value of $\beta_1 + \beta_2$, which is the weather effect for non-residential customers, and the corresponding p-value. The production index is normalized, and all coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are always controlled for. Zip-code-by-type FEs are included instead of zip code FEs. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

application for large systems. Overall, non-residential applications are much more likely to be canceled (27.90%), potentially due to greater complexity in the financing and installation process.

The regressions largely follow the baseline and preferred FE specifications but change the zip-code FEs into zip-code-by-type ones to control for finer cross-sectional variations. Columns (1)-(2) in Table 1.6 pool the residential sample with all 8,704 non-residential applications, while columns (3)-(4) include only small ones, and (5)-(6) include only large ones. The coefficient on *ProdIndex* captures the effect of weather on residential customers and is similar to the original estimates in Table 1.3. The interaction term has a positive coefficient in all columns. The sum of

these two coefficients suggests a slightly positive effect for the non-residential customers, which is the opposite of the effect on residential customers. The positive effect is even larger for small systems that are comparable in size to residential projects. Despite this puzzling observation, the F-test fails to reject having a null effect in all columns. Overall, these results suggest that non-residential customers do not fall under the same behavioral influence.

Next, I explore heterogeneous responses among residential applications in locations with different solar market penetration. As peer effects are found to play a role in the adoption of solar panels (Bollinger and Gillingham, 2012), they might also affect the cancellation decision. For instance, a solar customer might be less likely to cancel her contract if her neighborhood has more solar PV systems installed. That customer might also learn from her neighbors with older systems to recognize the transient nature of the fluctuations in solar power generation. These hypotheses are tested below.

Identification of peer effects in Bollinger and Gillingham (2012) relies on the fact that installation typically takes place with some delay after the sign-up, and the peer effects begin only after the installation. The increase in the current installed base (past sign-ups) is thus predetermined and uncorrelated with other contemporaneous shocks to the peer group, yielding a consistent estimate. Following this logic, I construct two measures of solar market penetration at the zip-code-by-month level and merge them with each application based on the sign-up month. They are predetermined under the same assumption. The first statistic is the number of all interconnected (installed) solar systems per thousand households, and the second is one for systems that are at least one year old. These calculations are based on the net energy metering (NEM) interconnection dataset, which contains all interconnected projects in California with information on zip code and the relevant dates.

Table 1.7 presents results of regressing cancellation on the production index, market penetration, and their interaction. Columns (1)-(2) are estimated using the current installed base and (3)-(4) are estimated using those older than one year. The coefficient on the production

Table 1.7: Peer Effects in Cancellations and Response to Weather

	Current Installed Base		Installed Base > 1 Year	
Cancel = 1	(1)	(2)	(3)	(4)
$ProdIndex_{post}$	−1.005	−0.716	−1.037	−.751
	[0.412]**	[0.367]*	[0.401]**	[0.351]**
$Penetration$	−0.096	−0.086	−0.124	−0.109
	[0.028]***	[0.027]***	[0.047]**	[0.046]**
$ProdIndex_{post} \times Penetration$	0.017	0.001	0.027	0.004
	[0.016]	[0.013]	[0.022]	[0.017]
D.V. Mean	12.06	12.06	12.06	12.06
R^2	0.054	0.032	0.054	0.032
N	150,859	151,029	150,859	151,029
Fixed Effects				
Zip Code		Yes		Yes
Zip Code-quarter	Yes		Yes	
County-quarter		Yes		Yes
Year	Yes	Yes	Yes	Yes

Notes: Market penetration is defined as the number of interconnected systems per thousand households in columns (1)-(2), and that of systems which have been interconnected for at least one year in columns (3)-(4). The production index is normalized. All coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

index is similar to before. As for market penetration, one more system per thousand households decreases cancellation probability by around 0.09 percentage point in columns (1)-(2), and 0.12 percentage point in (3)-(4). This finding complements the previous studies: peer effect not only motivates the households to sign up but also prevents them from regretting afterward. However, the coefficient on the interaction term remains small and statistically insignificant throughout, suggesting that having neighbors with older systems might not guard against the psychological effects of current weather.

I also explore whether responses differ with system and application characteristics (see Table 1.8). Column (1) shows that larger systems are more likely to be canceled, and (2) shows the same for those with a lower average cost. However, the small coefficients on the interaction

Table 1.8: Heterogeneous Effects by System Characteristics

Cancel = 1 Characteristic	(1) System Size Above Median	(2) Average Cost Above Median	(3) Third-Party Ownership	(4) Near Incentive Step Change
$\beta_1 : ProdIndex_{post}$	-0.753 $[0.309]^{**}$	-0.787 $[0.275]^{***}$	-0.449 $[0.251]^{*}$	-0.595 $[0.280]^{**}$
$\beta_2 : 1(Characteristic)$	1.723 $[0.383]^{***}$	-1.221 $[0.615]^{*}$	0.837 [0.522]	1.863 $[0.580]^{***}$
$\beta_3 : ProdIndex_{post}$ $\times 1(Characteristic)$	0.007 [0.270]	0.081 [0.210]	-0.638 $[0.292]^{**}$	-0.251 [0.334]
$\beta_1 + \beta_3$ p-value: $\beta_1 + \beta_3 = 0$	-0.746 0.010^{***}	-0.706 0.019^{**}	-1.087 0.003^{***}	-0.846 0.030^{**}
D.V. Mean	12.06	12.06	12.06	12.06
R^2	0.031	0.033	0.033	0.033
N	154,519	154,520	154,519	154,519
Fixed Effects				
Zip Code	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Quarter	Yes	Yes	Yes	Yes

Notes: The characteristic of interest is described in the header. In particular, the indicator in column (4) records whether the application is within 47 days to the next incentive step change. The production index is normalized. The last two rows in the top panel show the value of $\beta_1 + \beta_3$, which is the weather effect for customers with the characteristic of interest, and the corresponding p-value. All coefficients are multiplied by 100 for legibility. Other system characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

terms suggest that these characteristics do not systematically affect response to weather. Column (3) examines third-party ownership (TPO). TPO is a financing option that allows the household to acquire the system without paying a large up-front cost. The household signs a power-purchase agreement (PPA) or a lease with the solar company, who owns the system.[22] About half of all applications in the data are under TPO. It is unclear whether customers with TPO systems should be more or less sensitive to weather. On one hand, TPO partially insures the risk from weather variations by compensating the customer when production is particularly low. On the other hand,

[22]PPA allows the household to pay a low fixed rate for the electricity generated from the system, while the lease is similar to a loan which the household need to pay back.

selection into TPO is not random, as the option is created to appeal to customers who are more liquidity-constrained, risk-averse, or hesitant. The estimates suggest that they are much more sensitive to weather: the weather effect on them is more than twice of that on self-owned systems.

Column (4) examines heterogeneous effects based on whether the application is submitted near the next incentive step change. Step changes occur when the total solar capacity reserved in the program exceeds certain thresholds. Although the exact timing of the change is unknown, solar companies might still be able to anticipate an imminent step change. They might even take advantage of it and push marginal customers to sign up in order to secure a higher rebate.[23] Therefore, we might expect higher propensity to cancel and a stronger weather effect among these customers. I calculate the number of days to the next step change for each application based on the realized schedule.[24] The sample is split by whether this duration is less than 47 days, the 25th percentile of this variable (see Appendix Figure A.5 for the distribution of days to next step change).[25] The result suggests that these customers did have a substantially higher probability of cancellation. They are also slightly more sensitive to weather but this effect is insignificant.

Lastly, I examine whether there are heterogeneous responses across zip codes based on aggregate characteristics (as household-level characteristics are not available). These zip-code-level characteristics include median income, the fraction white population, the fraction with a bachelor degree, the fraction of urban households, average household size, average housing cost, whether the zip code is in Northern California, and whether it is in non-coastal counties. The results are reported in Appendix Table A.13. The only notable pattern is that the responses to weather are entirely driven by zip codes that have more urban households than rural ones. The rest of the zip code characteristics do not seem to split the sample in meaningful ways. This might be because these aggregate characteristics are not informative about the highly-selected group of

[23] According to a CSI staff, "... there was always a flurry of reservations and overall activity around step changes".

[24] Source: https://www.californiasolarstatistics.ca.gov/reports/budget_forecast/#Step.

[25] The results are robust to other thresholds in a similar range. There is no special reason to focus on the 25th percentile. The threshold should be far enough from the step change to have sufficient identifying variations while close enough to reflect imminent change.

solar customers (Harding and Rapson, 2013; Borenstein and Davis, 2015).

1.6 Mechanisms

In this section, I propose a number of psychological and rational mechanisms that might have driven the effects of weather in the solar market. Each mechanism is evaluated against the empirical evidence.

Projection bias. Projection bias is the tendency for people to exaggerate the degree to which their future preference will resemble the present one. This leads solar customers to overestimate the total utility of going solar when weather conditions bring higher marginal return and regret when conditions worsen. The empirical evidence on both cancellation and sign-up are consistent with the theoretical implications of projection bias in Section 1.2.

Misprediction of utility from a durable good comes in two types: mispredicting either the path of states or the utilities generated in those states. They generate the same predictions and have not been disentangled empirically in the literature (DellaVigna, 2009). Since solar panels are largely instrumental, I argue that mispredicting the objective states is more plausible.[26]

Salience effects. Bordalo, Gennaioli, and Shleifer (2013) propose a theory of context-dependent choice, where a consumer's attention is drawn to salient attributes of goods. I find empirical evidence that the solar customers respond most strongly to weather patterns associated with solar productivity. One explanation is that sunny weather makes the benefit from solar panels more salient relative to the cost, leading to more sign-ups. Likewise, gloomier weather reduces the benefit-cost ratio by reducing the salience of the benefit, thus leading to regrets. Therefore, this mechanism also generates implications that are consistent with the sign-up and cancellation results.

One way to distinguish projection bias from salience effects is to test if the consumers are

[26]This interpretation might not be valid for cases where there is a large subjective component in the utility of going solar, such as a customer with strong green preferences.

responding to the absolute "levels" of weather rather than deviations from the norm (Busse, Pope, Pope, and Silva-Risso, 2015). However, the test conflicts with our need to control for non-weather seasonal confounders. For this reason, the literature, including this study, has not been able to tell them apart.

Bayesian learning. Another possibility is that customers are in fact extracting useful information from weather fluctuations to guide their decision. Below, I discuss two main scenarios regarding how learning might occur.

First, the customers might need to learn about existing weather patterns because they were previously inattentive. However, the fixed effects in the regressions preclude the identification of this type of learning. Once the zip code fixed effects absorb most of the meaningful correlations between the short- and long-run weather, the former is no longer predictive of the latter (see Appendix Table A.14). It is also hard to reconcile with the cancellation results. By the time of the sign-up, the customers should have been informed by the installers on the estimated benefits. In particular, these estimates are based on simulations of local climate, thereby making it unnecessary for the customers to directly observe the weather.

Second, the customers might be learning about climate change, which might affect the desirability of solar power. Previous studies find that exposure to unusual heat increases people's belief in climate change (Li, Johnson, and Zaval, 2011; Deryugina, 2013). However, the temporal fixed effects should have absorbed the gradual movements driven by climate change, again precluding identification of meaningful learning. Furthermore, the effect of solar insolation remains highly significant after controlling for extreme heat or cold, while extreme heat has minimal effect on adoption (Table 1.2). Therefore, learning about climate change is also not likely to be the main mechanism.

Supply-side factors. Supply-side factors might also drive adoption. If solar companies ramp up their promotion efforts to take advantage of the good weather, it would directly increase sign-ups and might also help reduce cancellations if the latter is affected by visible marketing

efforts. The anecdotal evidence has been mixed. For instance, one salesman notes that it is easier to make a sale in the summer than the winter. On the other hand, my interview with a major company finds no high-level awareness or effort to incorporate current weather into the marketing plan. While the importance of this mechanism is difficult to gauge, it ultimately relies on people's psychological reactions to the weather. To the extent it is important the results should be interpreted as the overall effect, including the demand-side response together with any amplification effect from marketing practices.

Another possibility is that the solar workers are more productive in better weather. Solar salesmen and installers are outdoor workers whose labor supply and productivity are subject to weather fluctuations (Graff Zivin and Neidell, 2014). This could potentially change the entire interpretation because it is independent of any demand-side effect. In particular, one concern is that the installation might be repeatedly re-scheduled due to bad weather, eventually leading to cancellation. I partially test the veracity of this scenario by examining, among the completed projects, whether the completion duration is longer when the average weather (as measured by production index) is worse. There is no evidence of a significant relationship therein (see Appendix Table A.15). Moreover, I find little effect of extreme heat in Table 1.2. This is also inconsistent with the existing literature on heat and productivity where the biggest impact is often associated with extreme heat. Overall, there is little evidence in support of worker productivity being the main mechanism.

1.7 Conclusion

This paper studies the effect of short-run weather fluctuations on cancellations of signed solar contracts in California. This decision appears to respond most strongly to short-run weather patterns associated with solar panel productivity. The main result shows that customers whose sign-up for solar panels is followed by unfavorable weather are more likely to cancel their contract.

Furthermore, they are more likely to cancel when they experience worse weather after signing up relative to before, suggesting that the negative update changes their valuation of solar panels. I also find suggestive evidence that non-residential customers do not respond to weather, and the marginal ones respond more. The most plausible mechanism consistent with all these results is psychological, such as projection bias or a salience effect. Under such mechanisms, the customers rely too heavily on current weather to predict their future utility from solar panels.

These results suggest that projection bias or salience effects can substantially affect the demand for a big-ticket item such as a solar PV system. Compared to the other contexts studied in the literature, such biases in solar PV adoption create new welfare implications and policy challenges. Solar panels are under-adopted socially since electricity generated from fossil fuels is not correctly priced. They might also be under-adopted privately for reasons similar to the case of energy efficiency. Therefore, the undervaluation caused by these biases creates a much larger welfare loss than over-valuation. This highlights the importance of continual communications with the customers who are waiting for their systems to be installed, especially during periods of unfavorable weather. Given that cancellation reflects a wasteful cycle of over- and undervaluation created by weather fluctuations, it is also important to "de-bias" the customers and mitigate these cycles.

Finally, there are several limitations in this paper. For each I propose a direction for future work. The first limitation is not having the precise cancellation dates. While solar PV adoption is an interesting context where projections are continually updated within the relevant period, the lack of precise dates prevent me from exploring how the weather dynamics *within* this period are incorporated into decisions. Second, without demographic information of the customers, I am not able to fully examine whether the effects differ by education background, income, or household structure. Documenting these heterogeneities can help researchers deepen the current understanding of behavioral biases, and it can also help solar companies and behavioral programs to better target vulnerable populations. Third, this paper finds that information provided by the

solar companies or peer effects by neighbors do not eliminate such bias. Future studies can explore more effective strategies. Answering these questions are important for both economic modeling and policy design.

Acknowledgements

Chapter 1, in full, is currently being prepared for submission for publication of the material. Yanjun Liao. The dissertation author was the primary author of this chapter.

Chapter 2

How Hurricanes Sweep Up Housing Markets: Evidence from Florida

Abstract

This paper examines the impacts of hurricanes on the housing market and the associated implications for local population turnover. We directly characterize equilibrium dynamics in the housing market using micro-level estimates. For this purpose, we assemble a comprehensive dataset by combining housing transactions, parcel tax assessments, and hurricane history in Florida during 2000-2016. Our results show that hurricanes cause an increase in equilibrium prices and a concurrent decrease in the probability of transaction for homes in affected areas, both lasting up to three years. Together, these dynamics imply a negative transitory shock to the housing supply as a consequence of the hurricane. Furthermore, we match buyer characteristics from mortgage applications to provide the first buyer-level evidence on population turnover. We find that incoming homeowners in this period have higher incomes, leading to an overall shift in the local economic profile toward higher-income groups. Our findings suggest that market responses to destructive natural disasters can lead to uneven and lasting demographic changes in affected communities, even with a full recovery in physical capital.

2.1 Introduction

About three hurricanes make landfall every year along the Gulf and east coasts of the United States. In 2017 alone, these giant, spiraling tropical storms cost a total damage of $265 billion (USD).[1] While statistics of direct property losses and fatalities are readily available, many questions remain unanswered regarding the adjustments in the economy beyond immediate impacts. Little is known about how the distribution of physical capital responds to stochastic events like hurricanes. From a welfare perspective, the literature is only starting to investigate how individuals cope with natural disasters and their heterogeneous ability in doing so.[2] Answering these questions is important for assessing the long-run consequences of and adaptations to hurricanes.

Using transaction-level housing data from Florida, this paper studies the responses to hurricanes in housing markets for up to ten years after the hurricane strikes. The housing market offers a unique vantage point for understanding the interaction between natural forces and the local economy. Across localities, housing market conditions correspond closely to changes in population and long-run economic projections (Glaeser and Gyourko, 2005). They also determine who has access to a given location, which, in turn, affects access to economic opportunities and a range of important local amenities (Kling, Ludwig, and Katz, 2005; Kling, Liebman, and Katz, 2007; Chetty, Hendren, and Katz, 2016). Therefore, studying the housing market allows us to understand how hurricanes affect both the distribution of physical capital and the welfare of the affected population.

The contribution of this paper is two-fold. First, we estimate the causal impacts of

[1] Recent studies in climate science observe that hurricanes have been intensifying and traveling more slowly, possibly due to anthropogenic climate change (IPCC, 2013b; Kossin, 2018). Continuation of these trends would mean greater damages from hurricanes in the near term.

[2] This emerging literature is largely motivated by and focus on Hurricane Katrina. Earlier studies present descriptive evidence of migratory decisions made by Katrina evacuees (Groen and Polivka, 2008; McIntosh, 2008; Paxson and Rouse, 2008; Sastry and Gregory, 2014; Fussell, 2015). More recent studies identify causal impacts of Katrina on its victims' employment outcome and debt level (Gallagher and Hartley, 2017; Deryugina, Kawano, and Levitt, 2018).

hurricanes on the dynamics of housing prices and transaction probability in exposed areas. Together, they characterize the equilibrium adjustments in the housing market after the destructive impacts of hurricanes. While many factors in housing supply and demand can be affected by hurricanes, this approach identifies the dominant effect and provides evidence on whether the effect is temporary or more enduring. Second, we investigate the implications of these adjustments on affected neighborhoods, including potential changes in local demographics and housing characteristics. The direction of these changes is theoretically ambiguous: the large initial damage by hurricanes might induce further deterioration of affected communities, but it might also be neutralized in a full recovery, even lead to some type of creative destruction, as suggested by anecdotal evidence[3] and some past studies on aggregate growth effects (Skidmore and Toya, 2002; Crespo Cuaresma, Hlouskova, and Obersteiner, 2008; Pelli and Tschopp, 2017). This paper provides the first micro-level evidence on this question.

Our analysis focuses on Florida in 2000-2016. As the most hurricane-prone state in the U.S., Florida has experienced hurricanes with different strengths, ranging from severe outliers – such as Hurricane Andrew and Hurricane Katrina – to milder ones. Therefore, our results represent the impacts of an average hurricane. We assemble a detailed housing dataset by combining transaction records and county tax assessments. The transaction data cover 95% of the universe of housing transactions. The tax assessments provide rich information on hedonic characteristics of any given parcel in repeated observations, which we use to infer characteristics of the transacted home with accuracy. As such, this dataset allows us to identify and track individual parcels over time, observing most transactions and major renovations that took place between 2000 and 2016.

We use a staggered difference-in-differences framework that exploits the randomness in the paths and timing of several hurricanes. Treatment is defined, for individual parcels, as being exposed to hurricane-scale wind speed. We construct a measure of exposure based entirely on

[3] See, for example, https://www.newyorker.com/magazine/2011/03/28/creative-destruction.

physical and geographic attributes of the hurricanes. There is an important distinction between being exposed and being damaged. An exposed house is one that lies in the approximate area around the segment of the hurricane path with high wind speed, but *not* necessarily damaged. Essentially, our strategy compares areas exposed to a hurricane to unexposed ones before and after the hurricane.

We find that home prices *increase* in exposed areas in the three years following a hurricane. Compared to unexposed areas, home prices in exposed areas are 5% higher on average during this period, with a peak of 10% in the second year. This effect is identified in two models. The first uses variation within census tract while controlling for house characteristics, seasonality, and differential economic growth across counties. The second employs parcel fixed effects and restricts the identifying variation to be coming from repeated sales of the same structure. The direction and scale of the estimates are very similar across the two models, providing strong evidence that the effect is mainly driven by within-home appreciation rather than a shift in the composition of transacted homes.

We also find that the transaction probability of homes in exposed areas *falls* by 0.6 percentage points, or 6% of the baseline probability. The timing of this effect is the same as the price effect: both last around three years before returning to the baseline. Taken together, they suggest that the housing markets at exposed areas are dominated by a negative supply shock, but this effect is temporary. These results are consistent with the mechanism that hurricanes lead to significant physical damages to and destruction of part of the housing stock, thereby temporarily contracting available supply, and in turn, reducing sales and driving up prices. Housing shortage of this nature has been increasingly identified by media reporting on recent hurricane events.[4] Moreover, it takes time for hurricane victims to seek financial aid from insurance companies or federal agencies, and to eventually restore their damaged homes to habitable or salable conditions.

[4]See, for example, the Wall Street Journal's coverage of the impact of Hurricane Irma in Florida Keys (https://www.wsj.com/articles/hurricane-irma-destroyed-25-of-homes-in-florida-keys) or that of Hurricane Florence in North Carolina (https://www.wsj.com/articles/hurricane-florence-creating-housing-shortage-for-displaced-north-carolinians).

This explains the duration of the effect.

While the adjustment in the market equilibrium appears to be transitory, it generates lasting impacts on local demographics. We match a subset of our housing transactions to Home Mortgage Disclosure Act (HMDA) records to obtain income and demographic information of the home buyers.[5] Using this subsample, we show that the average income of new buyers increases nearly proportionally to the rise in home prices. This illustrates a greater ability of higher-income individuals to acquire homes even in disrupted market conditions. Moreover, both average transacted prices and buyer incomes return to the baseline level but not below in subsequent years. The three-year increase thus leads to a stock effect of around 28% of homes being occupied by households with higher income. In this light, hurricanes act as a place-based policy that leads to a substantial increase in neighbors with higher income, which could spur subsequent neighborhood improvements (Bayer, Ferreira, and McMillan, 2007; Card, Mas, and Rothstein, 2008; Guerrieri, Hartley, and Hurst, 2013; Diamond, 2016).

Our estimation strategy combines a staggered difference-in-differences framework and the repeat sales model. The former is common among studies that use variation from several hurricanes, but is often applied to aggregate data. On the other hand, the latter method is widely used in the real estate literature to estimate price effects that are free of confounding shifts in the composition of transacted homes. In particular, it is used in several papers to study the impact of a single hurricane on housing prices (Bin and Polasky, 2004; Hallstrom and Smith, 2005; Bin and Landry, 2013; Gibson, Mullins, and Hill, 2019). Our strategy extends this previous literature to a broader set of hurricanes while maintaining a credible identification of within-home price change. To our best knowledge, this is the first paper to employ this combination.

Our main point of departure from the literature on housing prices lies in the conceptual framework. These studies follow the tradition of using hedonic regressions to estimate the marginal willingness to pay for nonmarket goods, as first established in the seminal work by

[5]This subset of transactions are representative of the entire market in terms of home characteristics and price dynamics.

Rosen (1974).[6] Their analyses focus on the hurricanes' role in conveying risk information to households. As full capitalization requires a fixed housing stock, most of these studies deliberately avoid direct-hit locations and estimate the hurricane's effect on flood risk premium in near-miss areas. While perceived flood risk is a major factor in the adaptation to natural disasters, it is not the only way in which hurricanes affect the final allocation of housing through market interactions. For this reason, we compare exposed areas with unexposed ones, thereby obtaining estimates that reflect changes in the market equilibrium. This conceptual framework can be applied to other scenarios where the housing stock is exogenously altered by natural disasters, wars and social turmoil, private investments, or government policies.

This paper adds to a recent literature that estimates the impacts of hurricanes on a variety of economic outcomes. These include household responses (Gallagher, 2014; Gagnon and Lopez-Salido, 2015; Bleemer and Klaauw, 2017; Gallagher and Hartley, 2017; McCoy and Zhao, 2018), migration patterns (Paxson and Rouse, 2008; Boustan, Kahn, Rhode, and Yanguas, 2017), industry and labor market consequences (Groen and Polivka, 2008; McIntosh, 2008; Belasen and Polachek, 2009; Deryugina, Kawano, and Levitt, 2018; Seetharam, 2018), macroeconomic growth (Hsiang and Jina, 2014; Strobl, 2011), and government spending (Deryugina, 2017). Our analysis of the housing market not only adds an important outcome to this collection but also complements previous results on mortgage payback and migration to shed light on how market processes contribute to post-hurricane demographic adjustments. Our finding on higher housing prices is qualitatively similar to Murphy and Strobl (2010) and Aqeel (2011). Our transaction-level data, however, enables us to also observe transaction probability, which provides crucial evidence to clarify the nature of the equilibrium price shift.

This paper also contributes to a literature that examines the impacts of capital and infrastructure destruction on local economic activity. In different contexts of wartime destruction, past studies generally find a relatively rapid recovery, in line with our results on housing market

[6] See Kuminoff, Smith, and Timmins (2013) for a recent review on the broad literature of hedonic models.

adjustments (Ikle, 1951; Davis and Weinstein, 2002; Miguel and Roland, 2011; Feigenbaum, Lee, and Mezzanotti, 2017). In addition, we provide the first micro-level evidence on population turnover, which presents a more nuanced view of neighborhood changes.

Last but not least, our findings are related to several lines of research in urban economics, from taste-based sorting by individuals to neighborhood gentrification and urban transformation.[7] Our results suggest that destruction by exogenous natural events do not alter the fundamental forces underlying urban change. Moreover, our finding on buyer income in disrupted markets sheds light on potential distortions by credit constraints in taste-based sorting and raises further research and policy questions.

The remainder of the paper is organized as follows. Section 2.2 describes the data, Section 2.3 presents our estimation framework, Section 2.4 reports and interprets the results, and Section 2.5 concludes.

2.2 Data and Background

We have two objectives in this paper: first, to examine the equilibrium adjustments in the housing market following a hurricane; second, to understand the implications for population turnover. The first objective requires data that represent the universe of transactions, and the second requires demographic information on the home buyers. We build a comprehensive dataset that combines Florida housing transactions, tax assessments, demographics of mortgage holders, and hurricane exposure. This section gives an overview of the data and the background in this study.

[7] There is an extensive literature in each of these topic areas. Examples include Glaeser and Gyourko (2005), Card, Mas, and Rothstein (2008), Guerrieri, Hartley, and Hurst (2013), Albouy and Stuart (2014), Bayer, McMillan, Murphy, and Timmins (2016), Diamond (2016), and Diamond and McQuade (forthcoming)

2.2.1 Florida Housing Market

Florida is the third most populous state in the U.S., with varying demographic composition across locations (see Figure B.2 for the demographic composition in Florida counties and nearby states based on the 2010 Census). Most Florida counties have a population that is more than 70% white and less than 15% black. Several counties in Central and South Florida have a high median age at 45-55. Compared to nearby states, the population in Florida is older, significantly more white and less black, and with a larger percentage of Hispanics especially in South Florida.

Florida has a similar percentage of family households as surrounding states (see Figure B.4 for statistics on family and housing in Florida counties). On the other hand, both the homeowner vacancy rate (2-6%) and rental vacancy rate (10-20%) are significantly higher than nearby states. The percentage of occupied housing units is also lower in Central and South Florida. These patterns might be explained by a higher percentage of timeshares and vacation homes associated with its vibrant tourism industry.

The main source of our housing data is Zillow Transaction and Assessment Dataset (ZTRAX). We obtain data on Florida housing transactions between 2000 and 2016, which accounts for around 95% of all transactions over this time period. Each transaction record contains information on the timing, transaction price, mortgage profile (including loan amount and lender's name), location and condition of the property, along with buyers' and sellers' names. We exclude three types of transactions where the price might deviate from the market value: (1) non-arm's-length transactions[8]; (2) foreclosure sales[9]; (3) transactions that involve multiple homes on different parcels.

[8] An arm's length transaction is one in which the buyers and sellers act independently and do not have any relationship to each other. An example of a non-arm's-length transaction is one between family members, whose price often do not reflect market conditions. We rely on a combination of Zillow's internal code and the type of deed to determine the nature of the transaction.

[9] In theory, foreclosed properties might be auctioned off at their market values. In practice, however, the foreclosure process is subject to the negotiating process and idiosyncratic procedure of the bank. According to extensive analysis by the Zillow group, foreclosure sales are generally made at substantial discounts to comparable non-foreclosure sales. Moreover, foreclosed properties are often in more dilapidated conditions, which is unobservable to the econometrician.

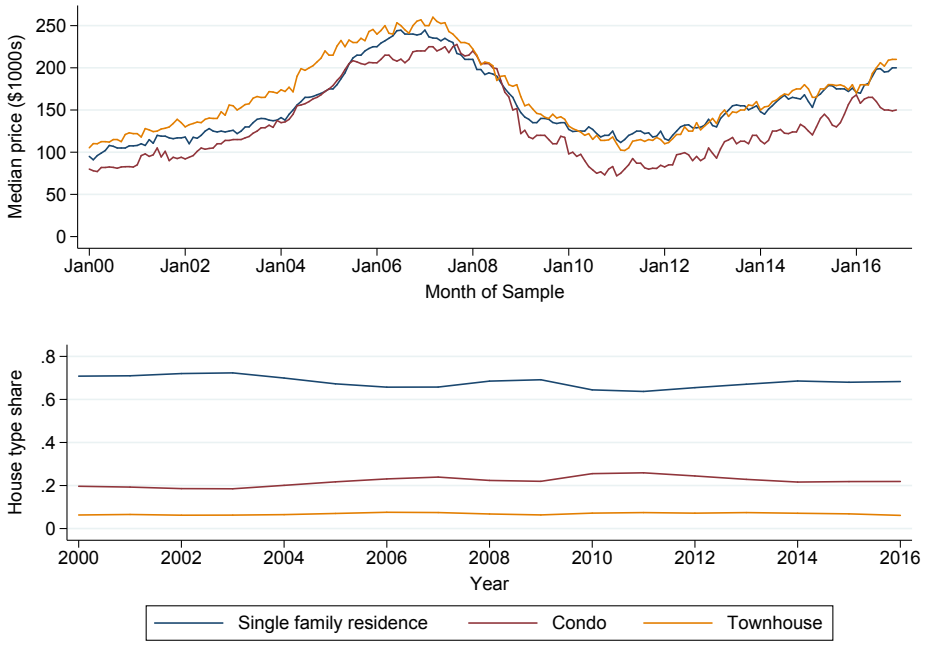


Figure 2.1: Florida Housing Market Sales and Composition

Source: Authors' calculation based on ZTRAX data.
Notes: The top panel plots the monthly median prices for the three major home types. The bottom panel plots the annual sales percentages of each home type. These time series are based on the full sample of Florida home buyers.

We also obtain parcel-level[10] assessment records over the same period from ZTRAX, which are originally generated by county assessor's offices. The assessment data contains an essential set of hedonic characteristics for each parcel, including square footage, year built and remodeled, lot size, number of rooms, number of bathrooms, number of units in building, and land use code, among others. Importantly, we observe multiple assessments for a single parcel and hence can track changes in these characteristics over time. We match the two datasets together by parcel and assessment year to ensure each house's condition at the time of the transaction is accurately reflected in the data.

[10] A parcel is also known as a lot or plot. It is a defined piece of real estate, usually resulting from the division of a large area of land.

Our data contains precise geographic coordinates for each parcel.[11] This has two major advantages. For one, we can determine the hurricane exposure of any home by directly calculating its distance from the hurricane tracks. We also use the coordinates in conjunction with detailed shapefiles[12] to accurately assign, for each parcel, the statistical geographic area in any given vintage. This allows us to match the housing data to mortgage records from HMDA with high accuracy, and to use fine geographic variation in our estimation.

Figure 2.1 plots the monthly median price by home type in the top panel. Median home prices experience large fluctuations. It started around $100,000 in 2000, rose to over $200,000 at the peak around 2007, declined back to $100,000 in 2010-2012, and gradually climbed back afterward as the economy recovered from the financial crisis. On the other hand, the lower panel shows that the composition of home types among sales remained stable over time. Single family residences account for around 70% of the sales, condominiums around 20%, and townhouses less than 10%. As we will match buyers with a mortgage (henceforth "borrowers") to HMDA records, we also report the time series of median price for the subset of all Zillow borrowers and the share of such sales (see Figure B.1 in the Appendix). The pattern of prices closely resembles those of all buyers although the share of sales with a mortgage decreases from 75% in 2007 to 50% in 2010 as the financial crisis took place.

2.2.2 Hurricane History and Exposure

Much of Florida is on a peninsula between the Gulf of Mexico and the North Atlantic. Its unique geography has exposed it to more hurricanes than any other U.S. state. The average Florida hurricane is much milder compared to the ones most studied in the past literature (*i.e.* Hurricane Katrina and Hurricane Andrew).

Between 1992 and 2017, a total of 15 hurricanes swept past parts of Florida. We obtain

[11] The longitude and latitude measures are down to five decimal places (1.11 meter at equator).

[12] Source: TIGER/Line Shapefiles, Census Bureau. Available at https://www.census.gov/geo/maps-data/data/tiger-line.html.

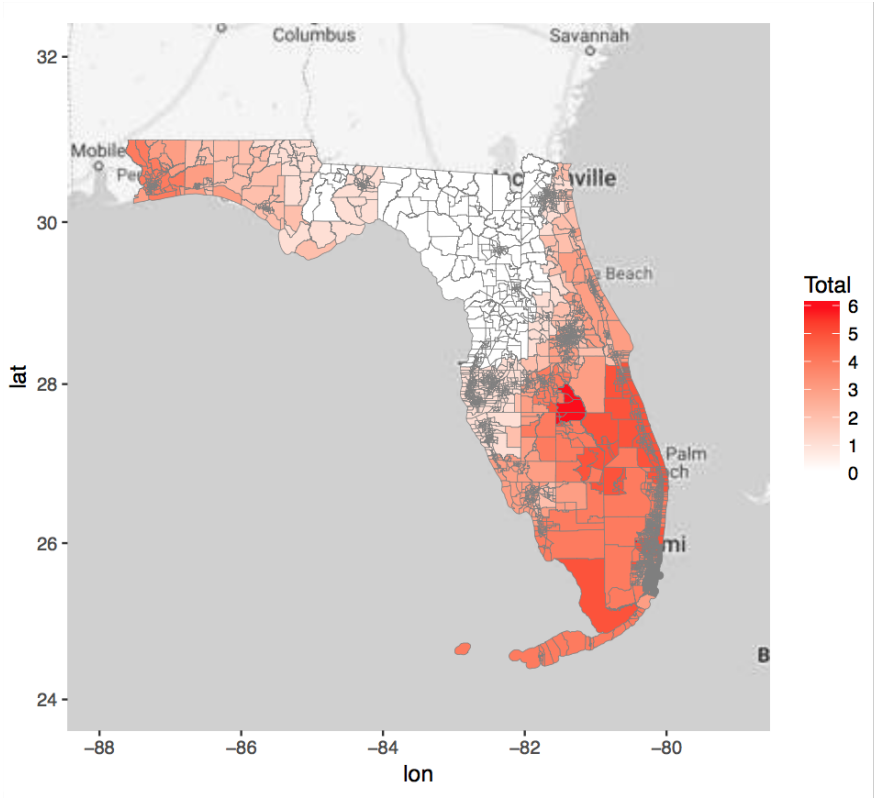


Figure 2.2: Exposure to Hurricanes by Census Tracts, 1992-2017

Notes: This graph shows the number of times a census tract is within the radius of hurricane-scale wind speed (64 nautical miles per hour) during 1992-2017. Calculations are based on detailed hurricane track points and census tract population centroid.

their physical attributes from the Tropical Cyclone Extended Best Track dataset.[13] Each tropical cyclone is recorded in six-hour intervals. Each observation reports the geographic coordinates of the center, maximum sustained wind speed, and maximum radial extent of 34, 50 and 64 nautical miles per hour (kn) wind.[14] We approximate the full hurricane path by linear interpolation between consecutive observations.[15]

Hurricane damages rise exponentially with wind speed. We are thus interested in determin-

[13]Data source: Demuth, DeMaria, and Knaff (2006). Available at http://rammb.cira.colostate.edu/research/tropical-cyclones/tc-extended-best-track-dataset/.

[14]A tropical cyclone is classified as a hurricane when the 1-minute sustained winds reach 64 kn (74 mph or 119 km/h). Thresholds for the SaffirSimpson hurricane wind scale are: Category 1, 74-95 mph; Category 2, 96110 mph; Category 3, 111129 mph; Category 4, 130156 mph; Category 5, 157+ mph.

[15]We also assume the hurricane center travels with constant speed between two consecutive observations, while wind speed and radii change linearly with time. See Appendix B.3.2 for more details on the interpolation procedure.

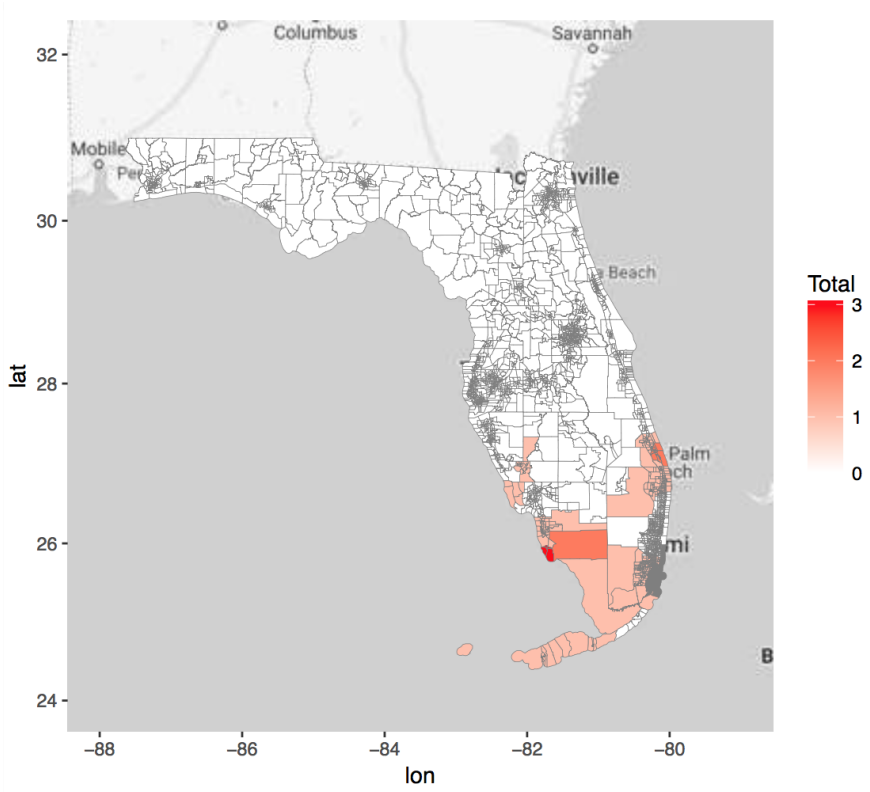


Figure 2.3: Severe Exposure by Census Tracts, 1992-2017

Notes: This graph shows the number of times a census tract is within the radius of category 3+ wind speed (96 nautical miles per hour) during 1992-2007. Calculations are based on detailed hurricane track points and census tract population centroid.

ing exposure to Category-3-and-beyond hurricane wind speed (96+ kn or 110+ mph) separately for heterogeneity analysis.[16] Since maximal reach radii of 96 kn is not provided in the dataset, we calculate our own measure by estimating a nonlinear relationship between wind speed and its reach radii (see Appendix B.3.1 for details on this procedure and the validity checks).

Throughout this paper, we define hurricane exposure by whether a location was ever within the reach of a 64 kn wind speed radius along a hurricane path. Severe exposure is similarly defined but with a 96 kn wind speed. This requires both proximity to the hurricane path and high sustained wind speed at that stretch of the path. The radius average 95 miles for observations

[16] We use the category 3 speed threshold both to be in line with previous literature, and because we do not believe the difference between the categories 1 and 2 thresholds is sufficient to produce measurable differences in outcomes. Category 4 wind speeds, on the other hand, almost never reach Florida shores over our hurricane time period.

with wind speed above 64 kn, and 45 miles for those above 96 kn. It is important to note that our exposure measure does *not* indicate actual damage. Houses lying outside of hurricane-scale wind speed radius, for example, might also be affected by heavy precipitation or flood, while those inside might sustain little damage due to strict building code or high elevation. Rather, we use this measure to identify homes located in areas that are likely to suffer from most direct impacts. Ultimately, we are interested in estimating equilibrium adjustments in these areas rather than the impact of being damaged.

Figure 2.2 shows the geographic distribution of hurricane events by census tract, defining exposure using their population-weighted centroids. Around 90% of tracts experience at least one hurricane event between 1992 and 2017, with high variation in frequency across locations. We plot the distribution of severe exposure events separately in Figure 2.3. This set of locations is much more limited (16%) and mostly concentrated in south Florida. The average census tract experiences 2.6 hurricane events and 0.16 severe exposure event between 1992 to 2017.

2.2.3 Home Mortgage Disclosure Act

The Home Mortgage Disclosure Act (HMDA), enacted by Congress in late 1975, requires all large financial institutions[17] to disclose all of their home lending activity every year. The loans reported are estimated to represent about 80 percent of all home lending nationwide. These records are made publicly available in an effort to improve home financing market transparency.[18] The HMDA data provide the date, property location (census tract), mortgage loan amount, purpose of application (purchase, improvement, or refinancing), mortgage lender's name and applicant demographics including annual income, gender, and race.

We merge the subset of successful loan applications for purchases from HMDA to our transaction data following the matching procedure described in Bayer, McMillan, Murphy, and

[17] By current definition, a large financial institution is one with more than $45 million in assets. This threshold is subject to yearly revision.

[18] See http://www.ffiec.gov/hmda for more details.

Timmins (2016). Matches are based on the year of each transaction, the census tract of the home, the loan amount (in 1000s), and the lender name.[19] The full procedure ultimately matches just over half of the original Zillow borrowers data, with no significant yearly variation in pairing success. This match rate is similar to that in Bayer, McMillan, Murphy, and Timmins (2016), which merges HMDA data to the universe of housing transactions in the Bay Area. The merged data enables us to look at the evolution of borrower demographic characteristics following hurricane events.

2.3 Econometric Framework

This section presents our research design. The first part of our analysis concerns post-hurricane adjustments in the housing market. We study two main outcomes that characterize the equilibrium – housing price and transaction probability – and use a separate regression model for each. The models differ in the dependent variable and the unit of analysis. Despite these outward differences, both essentially rely on the randomness of hurricane paths and timing as the identifying variation. In the second part of the analysis, we use a similar specification to the housing price model to examine the associated population turnover.

2.3.1 Housing Price Model

We model the transaction price of a home as follows:

$$log(Price_{ihmy}) = \sum_{\tau=-6}^{10} \beta_\tau Hurr_{imy}^\tau + \gamma' HouseChar_{iy} + \delta_{ht} + \delta_{hm} + \delta_{hcy} + \varepsilon_{ihmy} \qquad (2.1)$$

[19] See Appendix B.4 for more details on the matching procedure.

where i denotes an individual transaction, h the housing type,[20] m the month and y the year of the transaction, and t and c the census tract and county in which the transaction occurred, respectively.[21] The unit of analysis is an individual transaction. $log(Price_{ihmy})$ is the log of price in transaction i, which involves a house of type h and takes place in month m of year y. $Hurr^{\tau}_{imy}$ is a set of indicators specifying whether the transaction occurs τ years after the house was exposed to a hurricane ($\tau = 0$ refers to transactions in the first twelve months after a hurricane, $\tau = 1$ the next twelve months, and so on; a negative τ indicates the transaction happens before the hurricane). $HouseChar_{iy}$ is a set of house characteristics commonly used in hedonic models and include lot size, structural age, effective age,[22] number of stories and number of bathrooms. We control for the latter two characteristics flexibly using a set of value bins. These variables are taken from an assessment in the year of the transaction. For a given parcel, these variables might not remain the same over time and will be reflected in the data if the home is transacted multiple times.

We also account for both fixed and time-varying regional differences in housing attributes and local amenities by using a rich set of fixed effects. δ_{ht} is type-by-tract fixed effects, which absorb cross-sectional correlations in the likelihood of being hit by a hurricane and time-invariant local amenities, such as proximity to the coast. δ_{hm} is type-by-month fixed effects that control for the seasonality in both home prices and the timing of hurricanes. It is also important to account for the housing booms and busts during our sample period, especially given their uneven impacts across markets (Charles, Hurst, and Notowidigdo, 2018). We use type-by-county-by-year fixed effects (δ_{hcy}) to control for changing macroeconomic conditions at the county level.

The key variables of interest are the hurricane indicators. They are constructed based on

[20] We group homes into six main housing types based on land use classification in county tax assessments. They include single family residential (68.3%), condominium (21.7%), townhouse (6.83%), residential-multifamily (2.61%), vacation home (0.14%) and miscellaneous (0.4%).

[21] Census tracts and counties are defined according to the 2000 census definition throughout to maintain geographic consistency across time.

[22] The time (in years) since the house last had a major remodel.

the definition of hurricane exposure for individual parcels as described in Section 2.2.2.[23] Their identification requires that the idiosyncratic shocks to transaction prices are uncorrelated with them:

$$E[\varepsilon_{imy} \times Hurr_i^{\tau} | HouseChar_{iy}, \delta_{ht}, \delta_{hm}, \delta_{hcy}] = 0 \quad \forall \tau.$$

In essence, this is a staggered difference-in-difference (DID) framework with many hurricane treatments at different times.[24] Areas exposed to a hurricane are compared to unexposed ones before and after the hurricane in a repeated cross-section of housing transactions. Intuitively, because hurricane paths and timing are exogenous, exposed and unexposed areas are comparable absent the hurricane.

There are two main advantages of using this set of event time indicators rather than a single indicator as in a standard DID specification. First, the post-hurricane housing market might have transitional patterns that last an indeterminate duration. Analyzing these dynamics can provide key insights into the market adjustment process. Our set of post-hurricane indicators let them play out without imposing any restriction on the trend or duration. Second, the pre-hurricane indicators also allow for full flexibility in the pre-trends. This provides important evidence for us to assess the key identification assumption of DID, which requires the treatment and control group to have parallel trends absent the treatment. Our specification allows for powerful detection of differential pre-trends and provides a transparent partial test of the assumption.[25] To sum up, we adopt a data-driven approach to minimize misspecification and maximize the transparency of our research design.

Throughout this paper, we cluster standard errors at the county level to allow for correla-

[23] Specifically, we calculate exposure to each of the 15 hurricanes for each transacted home. We then construct the hurricane indicators based on the transaction timing relative to all hurricane events the house is exposed to.

[24] This design has been used by several studies of the economic impacts of natural disasters such as Belasen and Polachek (2009), Strobl (2011), Gallagher (2014), Hsiang and Jina (2014), Deryugina (2017), and Boustan, Kahn, Rhode, and Yanguas (2017)

[25] In fact, Deryugina (2017) points out that the parallel trend assumption can be relaxed given the exogeneity of the hurricane. See also Malani and Reif (2015). Our approach therefore provides extra assurance on the validity of our identifying variations.

tions in the idiosyncratic shocks to all transactions occurring in the same county over the entire sample period. We also estimate four variants of it, as described below.

Repeat sales. Hurricanes might change the composition of transacted homes. Equation (2.1) addresses this concern using a rich set of controls which greatly limits the extent of composition shifts in the estimates. To further eliminate within-tract selection of transacted attributes, we replace the census tract fixed effects with parcel-level ones:

$$1(Transacted)_{py} = \sum_{\tau=-6}^{10} \beta_\tau Hurr_{py}^\tau + \gamma' HouseChar_{py} + \delta_p + \delta_{hy} + \delta_{cy} + \varepsilon_{py}. \tag{2.2}$$

This approach restricts the identifying variation to price changes from repeated sales of the same home and holds all time-invariant characteristics fixed. This method is used to generate several major house price indexes[26] and also widely adopted in academic research (Cannaday, Munneke, and Yang, 2005; Hallstrom and Smith, 2005; Harding, Rosenthal, and Sirmans, 2007).

Wind intensity heterogeneities. We explore heterogeneous effects of different wind intensities, as stronger winds bring much more severe damages to the housing stock. Specifically, we split the exposed areas into those under category 1-2 wind speeds and those under stronger winds as discussed in Section 2.2.2. We estimate the dynamics of both groups simultaneously by interacting the indicator for each group with the event time indicators.

Other outcome variables. Hurricanes might change the set of homes transacted or the type of home buyers. Understanding any systematic shift is important for us to assess the nature of market responses to hurricanes. To do so, we replace the outcome variable in Equation (2.1) with a given characteristic of interest, such as lot size or log income of the buyer.

Standard DID. For ease of interpretation, we also estimate a standard DID model, defining treatment as being exposed to a hurricane in the last 36 months. The validity of this specification can be inferred from results on the event study, as the identification assumption is essentially the

[26]These indexes include the Case-Shiller Index, Federal Housing Finance Agencys (FHFA) monthly House Price Index, and CoreLogics LoanPerformance Home Price Index.

same for the two models.

2.3.2 Transaction Probability Model

A market equilibrium is characterized by both the price and the quantity. At the individual home level, transacted quantity is either zero or one. Formally, let $1(Transacted)_{py}$ be an indicator of whether a transaction record exists for parcel p in year y. We model its relationship to hurricane events as follows:

$$1(Transacted)_{py} = \sum_{\tau=-6}^{10} \beta_\tau Hurr_{py}^\tau + \gamma' HouseChar_{py} + \delta_p + \delta_{hy} + \delta_{cy} + \varepsilon_{py}. \tag{2.3}$$

where p, h, y, and c denote parcel, housing type, year, and county as before. The unit of analysis is now a parcel-year. We construct a balanced panel of most parcels that have been transacted during the sample period. We treat all parcel-year observations without a transaction record as having no transaction. This approach introduces measurement error if there are unreported transactions, and can even lead to biased estimates if the missing pattern is endogenous to or correlated with hurricane events. In our analysis, we take into account detectable patterns of missing records in a few county-years by omitting these entire counties from our sample. The combined sales in these counties account for less than 1% of total sales as they are all small counties. Therefore, our approach of omitting the entire sales histories does not yield any distinguishable differences in results from omitting only the missing years. We also include only parcels that have been built before 2000, so that the panel consistently tracks the housing stock that has been present at the beginning of the study period. Florida is subject to tight land supply due to geographic constraints (Saiz, 2010). This endogenously generates strict zoning regulation and hinders new developments, especially in dense urban areas.

We construct the panel by year rather than by month due to computational constraints. As most hurricanes take place in August, many transactions in the same calendar year might have

happened before the hurricane. This will severely attenuate the estimate of the year 0 hurricane indicator if we define the event time based on the calendar year. Instead, we redefine event years to begin in August and end in July. For example, a transaction in an area exposed to Katrina (August 2005) will have its year 0 indicator turned on *only if* it occurs between August 2005 and July 2006. This ensures a correct event year classification for most transactions, an uncontaminated event year -1 indicator, and a small and predictable downward bias in the estimate of the event year 0 indicator.[27]

Despite the differences in the data structure, this model shares many similarities to the previous one on housing price. We keep the same event-study specification. The model also controls for time-varying house characteristics observed from repeated assessments. Because the threat to identification is similar to that in the previous model, we also include parcel fixed effects to isolate within-home variation, county-by-year fixed effects to control for differential growth across counties, and type-by-year fixed effects for differential changes in home types. The identification assumption is also similar:

$$E[\varepsilon_{py} \times Hurr_p^{\tau} | HouseChar_{py}, \delta_p, \delta_{hy}, \delta_{cy}] = 0 \quad \forall \tau.$$

Our dataset includes three main parcel types: single family residence, condominium, and townhouse. Due to limited substitutability between these types, their dynamics might be different. Therefore, we also estimate this transaction probability model separately for each type. The type-by-year fixed effects are eliminated in these estimations. The standard errors are clustered at the county level throughout.

[27] Some hurricanes in our sample occur as late as October. The year 0 indicator might still get attenuated because it captures some transactions in areas affected by these hurricanes right before they actually took place. Given our approach to begin the year in August, this mechanical issue can only affect transactions occurring in the three months preceding hurricane Wilma (October 2005) or the two months preceding Frances and Jeanne (both dissipating in September 2004). The ensuing attenuation bias to event year 0 indicator is small and predictable. An alternative is to begin the year even later. However, this other alternative misclassifies some transactions of actually exposed homes to event year -1. We choose to remove a few observations from the event year -1 group but keep it uncontaminated: a future hurricane cannot have a causal effect on any of our outcomes before it hits because its path is unpredictable.

2.3.3 Variation in Hurricane Exposure

Hurricanes are rare events. Although their paths and timing are random, they might hit locations with distinctive characteristics and growth patterns by pure chance. This is an issue discussed and explicitly accounted for in the past literature (Deryugina, 2017). We alleviate this concern by using a longer hurricane history and hence a larger set of hurricanes. This approach reduces small sample bias[28] while preserving unbiasedness so long as hurricanes are exogenous. It also allows us to maintain the interpretation of our results as the effects of an average hurricane in Florida rather than a small, selected subset of hurricanes. From Figure 2.2, we can see that Florida hurricanes during 1992-2017 create a fairly disperse geographic distribution together. In this section, we take a closer look at variation over time as well as contribution by individual hurricanes.

For each separate hurricane, we first examine the percentage of transactions that are in its affected area, in each of the 10 years before and after its occurrence (see Table B.3 for a summary of these percentages). Looking across hurricanes, we obtain a snapshot of the composition of hurricane variations used in Equation (2.1), the housing price model.[29] These variation profiles illustrate that the seventeen-year width of our sample constrains the length of pre- and post-hurricane dynamics that we can estimate in our models. Event years -5 to 10 all contain variation from at least ten out of sixteen hurricanes and 8-12% of transactions in exposed areas. Event year -6 contain variation from six hurricanes and a total 5% impacted transactions. However, variation in event years -10 to -7 come from only two hurricanes, raising serious representativeness concerns.

We also generate similar variation profiles for Equation (2.3), the transaction probability model (see Table B.4). As this model is estimated using a balanced panel of parcels, the exposed

[28] "Small sample" refers to a limited set of hurricanes.

[29] The last row of each panel in Table B.3 reports cumulative exposure in each event time year. Note that this percentage is less than or equal to the sum of individual hurricane exposure because transacted homes may have been hit by multiple hurricanes in an event year.

percentage is constant for each hurricane across event years. Again, we see a similar pattern where earlier event years pick up variation from very few hurricanes, while event years starting from -5 capture exposure to more hurricanes. In these later years, exposed areas also account for around 10% of all parcels. Overall, these patterns attest to the similar nature of identifying variation used in the two models.

Based on these patterns, we exclude indicators for event years -10 to -7 from the estimation. This leaves us with six pre-periods and ten post-periods, comparable to the past literature that studies natural disasters using an event study design (Murphy and Strobl, 2010; Gallagher, 2014; Gallagher and Hartley, 2017; Deryugina, Kawano, and Levitt, 2018). We have also estimated models with shorter dynamics and found the results robust in both direction and scale.

Lastly, we examine variation in severe exposure (category 3 or greater wind speeds) in the same way (see Tables B.5). The fraction of transactions subject to high wind speeds is not only small (less than 1% in all years) but also limited to 5 hurricanes in total. This will greatly limit the power when we estimate the heterogeneous effects of higher wind speeds. In addition, the geographic concentration of severe exposure is high according to Figure 2.3. This limited set of locations are less likely to have parallel pre-trends with the rest of Florida.

2.4 Results

The results are presented in several subsections below. We begin by estimating the post-hurricane dynamics in housing prices for areas affected by hurricane while addressing potential composition effects in Section 2.4.1. Likewise, we estimate how transaction probabilities adjust for parcels in these areas in Section 2.4.2. Together, these two outcomes characterize equilibrium shifts in the housing market as it recovers from hurricanes, which we interpret further in Section 2.4.3. In Section 2.4.4, we examine the implications of these market adjustments on local population turnover. In Section 2.4.5, we explore whether the previous effects are heterogeneous

according to hurricane intensity.

2.4.1 Post-Hurricane Price Dynamics

We first estimate the post-hurricane dynamics of housing prices using the full sample of all transactions. The average price and home attributes in these transactions are reported in the first two columns in Table B.1. We use a flexible difference-in-differences specification as detailed in Equation (2.1). The standard errors were clustered by county over the entire sample length.

The estimated coefficients of event year indicators are plotted along with their 95% confidence intervals in Figure 2.4 (the full results are reported in the first column of Table B.6). None of the estimates for pre-hurricane indicators are statistically different from zero, supporting our choice of fixed effects that control for preexisting differences in average census tract prices, seasonality, and county-specific dynamics. Our event year 0 and 1 estimates suggest that hurricanes result in increases in home prices of 5% in the first and 10% in the second twelve-month periods after the strike. The surge in prices appears to end sometime in the next twelve-month period (event year 2) as the estimated increase relative to unaffected homes drops to 2% and is no longer statistically significant. All later event year estimates are small and not statistically distinguishable from zero. In short, the estimates suggest a temporary surge in prices in the three years immediately following a hurricane in exposed areas. In Table B.7, we provide additional results using a standard DID specification, defining treatment as being in event year 0-2, for the overall market as well as by home types. The main takeaway is, while there is an overall increase in home prices, the effect is particularly pronounced for condominia.

One possible interpretation is a shift in the distribution of transacted homes: in the aftermath of a hurricane, homes with greater hurricane resistance, quality or general desirability might be more likely to get transacted. In the extreme, the price surge could have been entirely driven by such composition shift without actually reflecting change in any individual home's

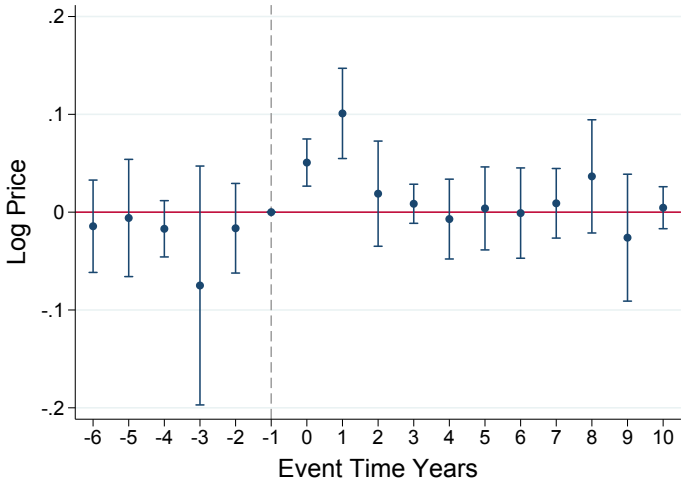


Figure 2.4: Hurricane Effects on House Prices – Full Sample

Notes: Estimates from Equation (2.1) are plotted with their 95% confidence intervals. The model controls for standard hedonic variables and census tract, month, and county-year fixed effects. Standard errors are clustered by county. The results are based on the full sample of Florida home sales (N = 7,408,259).

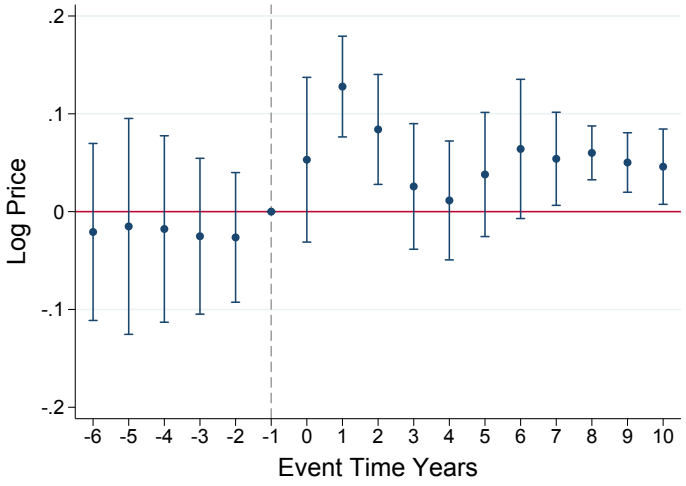


Figure 2.5: Hurricane Effects on House Prices – Repeat Sales

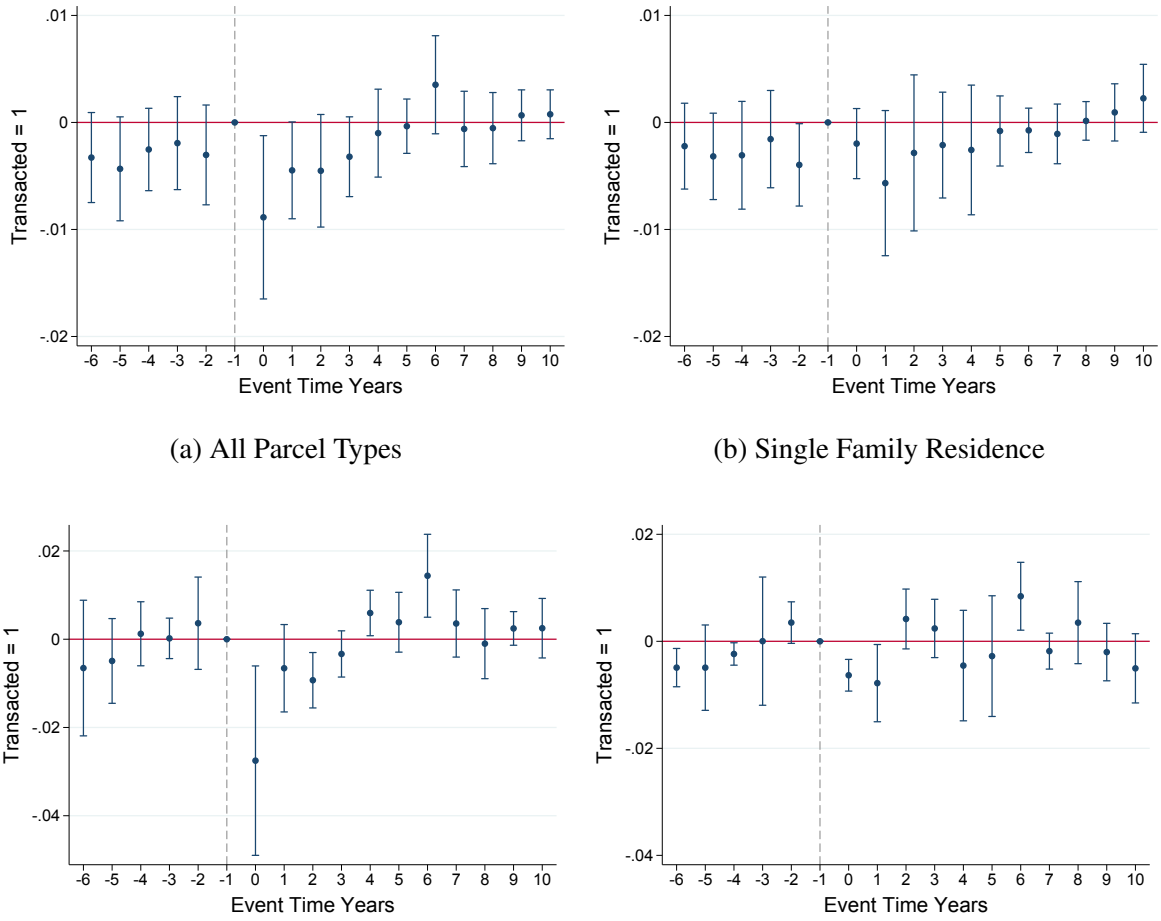
Notes: Estimates with parcel, month and county-year fixed effects are plotted with their 95% confidence intervals. The model also controls for age and effective age at time of sale. Standard errors are clustered by county. The results are based on only parcels with repeated transactions appearing in both pre- and post-hurricane periods (N = 1,338,384).

price. Our model partially addresses this possibility by controlling for a standard set of hedonic variables. The estimated effect is thus free of a shift in observed quality or any unobserved aspect that is strongly correlated with the observed characteristics.

Nevertheless, we cannot completely rule out this possibility due to limited observability of damage-related house characteristics such as base elevation or roof condition. To investigate this issue further, we re-estimate our price model with the inclusion of parcel fixed effects, and restrict the sample to only include homes which were transacted at least once both before and after a hurricane. The result is plotted in Figure 2.5 and reported in the second column of Table B.6. The standard errors are slightly larger due to the extensive fixed effects and a smaller sample size. On the other hand, this specification yields stronger evidence of parallel pre-trends which suggests greater credibility in identification. More importantly, the previous patterns are closely reproduced and if anything, larger in scale. Our point estimates imply that these homes on average sold at 5% higher prices in the event year immediately following the hurricane, as much as 14% in the next event year, and 8% in the third event year. These results suggest that the homes sold in the first two to three years in exposed areas appreciated relative to when they were sold outside of this post-hurricane window. Moreover, such within-home appreciation is likely to be the main driver of the price effect in the full sample, given its slightly larger magnitude. In order to better understand the mechanisms behind this finding, we now turn towards the potential hurricane-induced changes in the probability of home transaction.

2.4.2 Post-Hurricane Dynamics of Transaction Probability

We estimate the post-hurricane dynamics of transaction probability using Equation (2.3). As described in Section 2.3, we use a balanced panel of parcels where each observation is a parcel-year with an associated indicator for whether or not a transaction occurred in that year. Our analysis focus on only parcels with a structure built before 2000. The summary statistics are reported in Table B.2.



(a) All Parcel Types

(b) Single Family Residence

(c) Condominium

(d) Townhouse

Figure 2.6: Hurricane Effects on Transaction Probability by House Type

Notes: The estimates from Equation (2.3) are plotted with their 95% confidence intervals. Standard errors are clustered by county. Panel (a) plots the results based on all parcel types (N = 49,120,139), panels (b)-(d) are based on single family residence (N = 33,651,628), condominium (N = 12,836,476), and townhouse (N = 2,976,067), respectively. The model controls for effective age, parcel and county-year fixed effects. In panel (a), the model also includes type-year fixed effects.

The estimates from the full sample are plotted along with their 95% confidence intervals in Panel (a) of Figure 2.6 (as well as reported in the third column of Table B.6). The standard errors are again clustered by county over the entire sample length. Baseline probability of transaction is around 10%. Again, we see a stable pre-trend with all pre-hurricane indicators being statistically indistinguishable from zero. In event year 0, the estimate indicates a drop of 0.9 percentage points, or nine percent of the baseline probability, which is also statistically significant at the 5% level. The effect shrinks to 0.3-0.5 percentage points in the next three event years, and gradually returns to the baseline.

There are three main types of parcels in this dataset: single family residence, condominium, and townhouse. As shown in Table B.2, their characteristics are quite different and substitutability between them might be limited. Considering also the wide confidence intervals for the above estimates, it is likely that these market segments might have different equilibrium dynamics. Therefore, we also estimate the model for each type separately. These results are reported in panels (b)-(d) in Figure 2.6.[30]

Single family residences are the largest segment of the housing market and make up around 70% of all transactions. For them, we observe a negative but small effect that peaks at 5 percentage point in event year 1, and none of the event year indicators is statistically significant at the 5% level. In contrast, we observe larger changes in transaction probability for the latter two types of homes. Condominia, which usually represent 20% of transactions, are nearly 3 percentage points less likely to be transacted in event year 0, 1 percentage points in the following two event years, and around 0.5 percentage point in the next, before returning to the baseline. Townhouses make up the smallest share of the market at under 10%, and experience a drop of 1 percentage point in the first two event years.

Taken together, our results show a 6% decrease in transaction probability in the entire market, averaged over three years. We provide a more direct display of these scales by home type

[30] We use the same specification except for dropping house-type-by-year fixed effects.

in Table B.8, where we report estimates from a standard DID specification. The results again show substantial heterogeneity: the effects range from a 20% drop for condominia to only 3% for single family residences. We consider the potential sources of such heterogeneity in the next section as part of a discussion of mechanisms.

2.4.3 Mechanisms for Price and Transaction Probability Dynamics

In the previous two sections, we find that post-hurricane housing markets see an increase in housing prices and a concurrent fall in transactions. These dynamics together are consistent with a negative supply shock. We further find that this shock is temporary, as both price and probability appear to respond for two to three years and eventually return to the baseline.

To clarify, this does *not* imply that hurricanes only induce supply-side adjustments. In fact, localities exposed to hurricanes are likely to experience changes in housing demand through mechanisms such as people updating their beliefs in local disaster risk (Bin and Polasky, 2004; Hallstrom and Smith, 2005; Bin, Kruse, and Landry, 2008; Gibson, Mullins, and Hill, 2019), disruptions to local industries and labor market (Belasen and Polachek, 2009; Seetharam, 2018), among others. However, it is very hard to exhaust all possible factors in demand or supply that might shift after a hurricane and estimate the magnitude of each. Instead, we remain agnostic about them so far but choose to estimate directly the first-order effect driving the market equilibrium. Below, we interpret these findings further by discussing the nature of the negative supply shock and exploring a number of alternative explanations.

Mechanism 1: Physical damage and repair process.

Hurricanes generate extreme winds and severe flooding that can destroy or cause sufficient damages to buildings. As a result, part of the housing stock in exposed areas becomes uninhabitable in the short run, which explains the negative supply shock. It also creates a mechanical drop in the transaction probability because it is very hard to sell a house in a damaged state.

Affected homeowners may seek to temporarily relocate to habitable properties nearby through renting, which is covered by insurance companies and federal assistance in most cases. Demand increases in the rental market, which, in turn, put upward pressure on housing prices. This might partly account for the large price increase for condominia: a larger percentage of them are rental properties and might see more appreciation.

It takes time for homeowners to restore their homes to the original condition. For one, a large majority of affected homeowners seek post-hurricane financial relief in the form of insurance payment or FEMA assistance.[31] This requires homeowners to go through the filing process, wait for assessment of the damages to be performed, and settle potential disputes with the insurance company. This process can take months because federal institutions and private insurance companies are often overwhelmed by claims at times of large-scaled disasters. The findings in Gallagher and Hartley (2017) directly shed light on the time frame for households to obtain disaster aid: while victims of hurricane Katrina use disaster relief to pay down home loans, the effect did not stabilize until after a year.

The rebuilding or repairing process can also take a long time. Anecdotal evidence suggests that post-disaster recovery is often hindered by a shortage of construction labor. In the end, most people find it takes at least 18-24 months to repair, rebuild, or replace their home and possessions after a large loss. Notably, this duration is consistent with the observed effects on price and and probability. In particular, according to the Florida Condominium Act, the condominium association is responsible for repairing the exterior damages rather than the homeowners. This division of responsibility could further restrict condo supply in the market, which is also consistent with the particularly large effect in this market segment. Overall, physical damage and destruction provides a plausible explanation that matches the findings well.

[31] There are various types of financial relief homeowners can seek after a disaster. For example, wind damages are often covered by homeowner's insurance, while flood damages are separately covered by flood insurance. Homeowners might also seek relief from the federal government for uncovered damages through the Individual Assistance program (Kutz and Ryan, 2006; Michel-Kerjan, 2010; Hoople, 2013; Gallagher and Hartley, 2017; Deryugina, Kawano, and Levitt, 2018; Kousky, Michel-Kerjan, and Raschky, 2018).

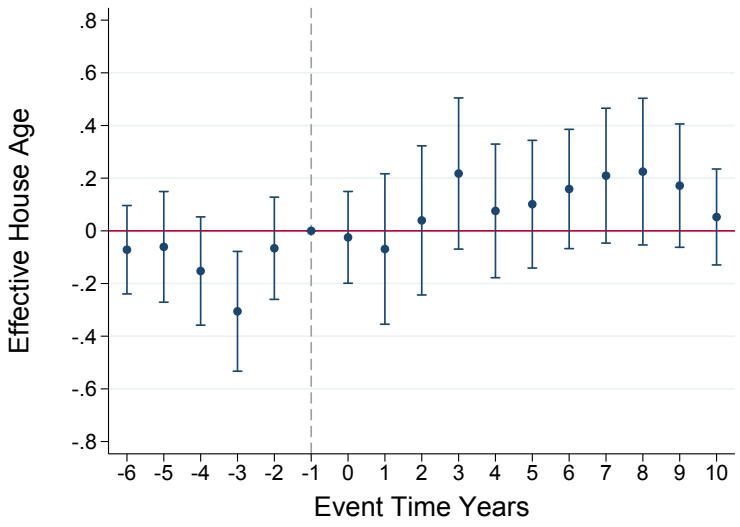


Figure 2.8: Changes in Characteristics of Transacted Homes – Effective Age

Notes: The estimates of difference-in-differences coefficients are plotted with their 95% confidence intervals. The dependent variable is effective house age, which is the number of years since last remodeled, if applicable. The specification is similar to Equation (2.1), which controls for census tract, month, and county-year fixed effects.

Mechanism 2: Quality improvement through rebuilding.

Another possibility is that the rebuilding process not only restores but also improves the quality of the home, thereby driving up its value. This could happen for a number of reasons. First, if the repair is funded by insurance payment or government assistance, the homeowner has an incentive to overuse the construction service (*e.g.* to better account for the possibility of future events through investment in more disaster-resistant structures and materials). Second, the building codes in Florida have been updated multiple times over the sample period largely for the purpose of enhancing resistance to natural disasters. The restoration project might thus be subject to stricter codes compared to its original construction.[32] Lastly, hurricanes might destroy some houses with particularly poor quality, sparing homeowners the fixed cost of tearing down a building and setting in motion a renovation process that would have taken place later. Any of

[32] A 2018 report on new residential building codes and enforcement systems by the Insurance Institute for Business & Home Safety gave Florida a rating of 95/100, the highest score among all the states it evaluated.

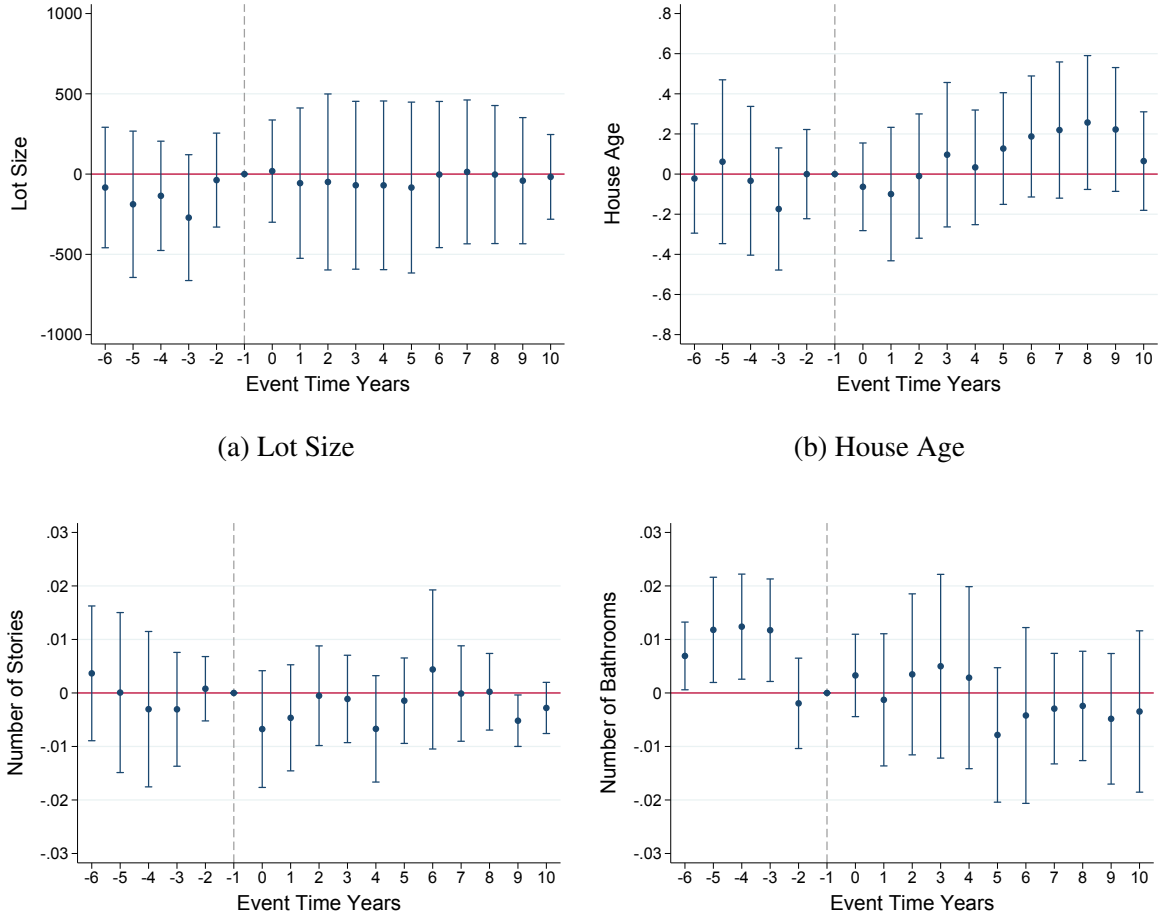
these scenarios is consistent with an increase in price. However, given what we know about repair time, it is questionable whether such rebuilding can be speedy enough to drive the peak increase in the first or second year after the hurricane.

We explore this possibility by estimating whether there is a systematic change in the effective age of houses transacted after a hurricane using a variant of Equation (2.1). Effective age records the years since the house was last remodeled. It is likely to capture (a) repairs funded by private insurance and (b) large renovations as both are likely to be permitted. On the other hand, it would fail to capture renovations that are not permitted, which is a caveat to this analysis. If this mechanism is important, we would expect the houses transacted during the adjustment period to have a smaller effective age because they have just been renovated. Instead, the estimates show no change at all (Figure 2.8), providing little empirical support for this mechanism.

Mechanism 3: Composition shift through information revelation.

As discussed in Section 2.4.1, we control for major hedonic characteristics and parcel fixed effects, which addresses most types of composition shift. Still, we cannot rule out a subtle composition effect: hurricanes provide an opportunity for the more disaster-resistant houses to overcome the information asymmetry problem on the market. These are houses that sustain minimal damage despite having been exposed to a hurricane, potentially due to better building materials or structural factors. In normal times, the home buyers might not be able to recognize their superior ability to withstand hurricanes. After a hurricane passes, however, these houses are valued at a premium because their quality has been revealed. This mechanism is consistent with the within-home appreciation and immune to our controls.

This mechanism generates two testable implications. First, hurricanes induce differential capitalization so that certain characteristics will be valued more. Second, hurricanes induce a composition shift toward these characteristics in transacted homes. Unfortunately, such characteristics – building materials, roof conditions, etc. – are not easily observable. Here, we provide



(a) Lot Size

(b) House Age

(c) Number of Stories

(d) Number of Bathrooms

Figure 2.9: Changes in Characteristics of Transacted Homes

Notes: The estimates of difference-in-differences coefficients are plotted with their 95% confidence intervals. Each panel has a different dependent variables: panels (a)-(d) correspond to lot size, house age, number of stories, and number of bathrooms, respectively. The specification is similar to Equation (2.1), controlling for census tract, month, and county-year fixed effects.

a partial test of these implications using observed characteristics that are predictors of disaster resistance. One such characteristic is house age: newer houses are more resistant to hurricanes because of improvements in building technology and more stringent building codes. Compared to effective age, the age of the house might better reflect the quality of the structure. Another characteristics is the number of stories, as taller buildings are more susceptible to wind damage.

To test for differential capitalization, we interact the treatment in the standard DID specification with each of the home characteristics. These interaction terms will capture any differential capitalization effect during the adjustment period. As reported in Table B.9, there is no differential capitalization effect for house age or being a one-story building. We also find a *smaller* effect of being recently renovated, which is contrary to the expectation that these homes might be more resistant to hurricanes. On the other hand, there is an *increase* in the value of lot size and a *decrease* in the value of bathrooms. These changes are consistent with the appreciation of habitable space due to the negative supply shock.

We also test for any systematic change in these variables among transacted homes, as predicted by the second testable implication. As shown in Figure 2.9, none of these characteristics see detectable change after the hurricane. If anything, there is a gradual rise in average house age, which is likely a mechanical effect due to aging of the housing stock.

2.4.4 Implications for Population Turnover

Equilibrium shifts in the housing market, as documented in the previous section, might change the match between housing and potential home buyers. In turn, these dynamics might have demographic implications for hurricane-battered communities. In this section, we turn our attention to how the average income level of home buyers evolves following hurricanes. The analysis is based on a subsample of home buyers for whom we find a valid match in the Zillow-HMDA matching procedure. Before we present the main results, we will first discuss the representativeness of this subsample.

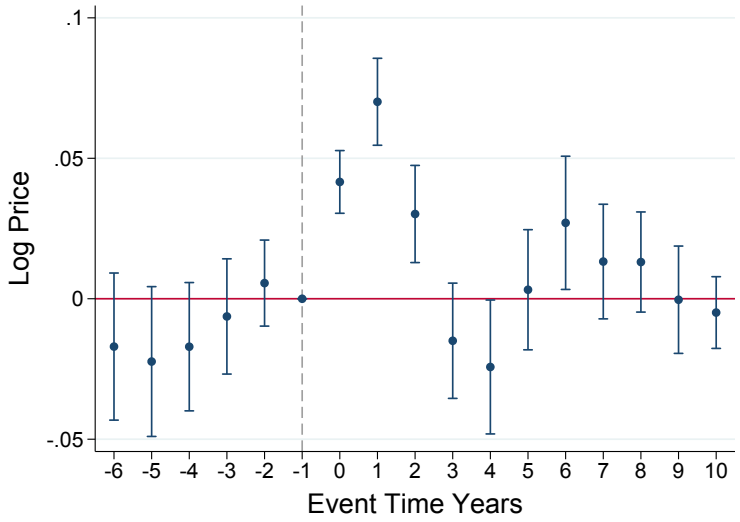


Figure 2.11: Hurricane Effects on House Prices – HMDA Sample

Notes: Estimates from Equation (2.1) are plotted with their 95% confidence intervals. The model includes standard hedonic variables and census tract, month, and county-year fixed effects. Standard errors are clustered by county. The results are based on the HMDA sub-sample (N = 1,967,334).

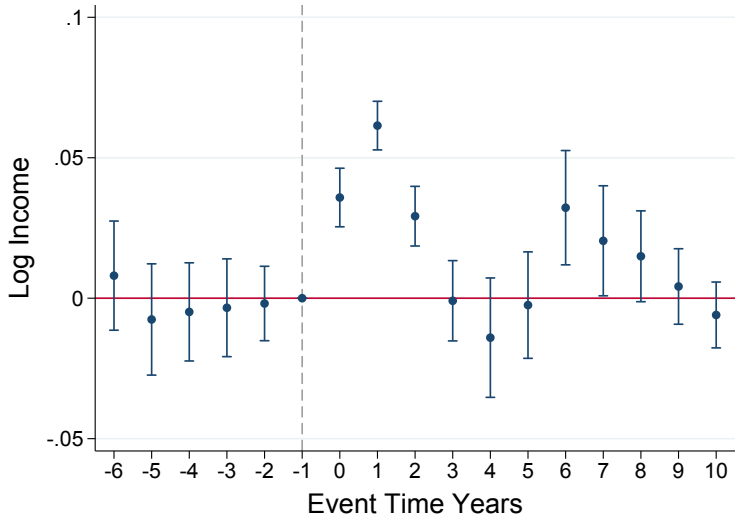


Figure 2.12: Hurricane Effects on Buyer Income – HMDA Sample

Notes: Estimates on buyer income are plotted with their 95% confidence intervals. The model includes standard hedonic variables and census tract, month, and county-year fixed effects. Standard errors are clustered by county. The results are based on the HMDA sub-sample (N = 1,881,208).

A hurricane might induce changes in the composition of mortgage borrowers that are different from the whole market. Lenders might change their lending criteria. If the would-be buyers are local, they might have lower liquidity level. To shed light on these potential mechanisms of selection, we estimate whether hurricanes affect the share of home sales involving a mortgage. There is no evidence of change (Figure B.8), although we cannot rule out the possibility of opposite selection forces canceling each other out. Next, we directly compare variables in the three samples: all transactions, transactions with a mortgage, and HMDA-matched transactions. Table B.1 reports the mean and median of various characteristics of transacted homes. The full sample appears to contain more outliers, or expensive homes with a large lot size. Since these are likely to be estates of the uber-wealthy, their transaction is not likely to be financed by a mortgage and thus would not show up in the borrower or HMDA subsample. Apart from that, most characteristics are comparable across samples in terms of the median. We also re-estimate Equation (2.1) on the subsamples to examine whether the price dynamics there are similar to that in the full sample. The price response in the HMDA subsample (Figure 2.11) indeed closely resembles those from both the full sample (Figure 2.4) and the borrower subsample (Figure B.6). Overall, the transactions in the HMDA subsample appear to be quite comparable to the full sample.

Using the HMDA subsample, we now estimate the effect of hurricanes on the average income of new buyers by replacing the outcome variable in Equation (2.1) with buyer income. We find the dynamics of post-hurricane income are strikingly similar to those of prices: the average income increases by around 4% in the first event year and nearly 7% in the second, then reverts to a 4% rise in the third before returning to their pre-hurricane baseline in later years. Since we control for home characteristics, this effect is not driven by better homes being transacted. Rather, homes of comparable quality as before are being transferred to buyers with higher income. Moreover, the subsequent dynamics exhibit some fluctuating but mostly positive effects. A back-of-envelope calculation suggests that a total of 28% of local homes are occupied by households

with higher income after the 3-year market adjustment period. Given that home ownership is a long-term decision, this pattern likely represents a lasting change in the economic profile of local communities toward higher-income groups. An influx of higher-income households might also trigger further gentrification of these neighborhoods (Card, Mas, and Rothstein, 2008; Guerrieri, Hartley, and Hurst, 2013). This raises concerns for future disaster costs, in terms of both private damages and the burden on federal budget.

The present finding, in conjunction with the market adjustment results, suggests that a negative shock to housing supply induces a selection of buyers with higher avarage income. We present a model in Appendix B.5 to explore underlying channels of such selection. The model features households with heterogeneous income and idiosyncratic taste for a given location. The first channel is preference, where willingness-to-pay for these locations are different across income groups. The second channel concerns credit constraints, which directly relate household income with the home price a household can afford (Greenwald, 2018). In the model, we derive two conditions that can explain the empirical relationship:

1. In an unconstrained market, valuation for hurricane-affected areas must be systematically higher for households with higher income.

2. The presence of a credit constraint relaxes the above requirement.

These conditions provide two plausible explanations for the positive income selection. First, higher-income households simply have higher willingness-to-pay for a given location. This is consistent with a mechanism of Tiebout sorting, as outlined in Banzhaf and Walsh (2008). In short, suppose communities can be ranked by amenity quality, which is valued more by higher-income households, households with similar income levels will sort into the same community. Within each community, the households' willingness-to-pay will line up with their income level, as lower-income households have a stronger tendency to substitute location benefit with consumption. However, this mechanism might not be able to account for cases where

household preferences for locations are not monotonically changing with income.[33] As suggested by the second condition, an alternative explanation is the presence of credit constraints. As credit constraints directly restrict the set of houses a household can afford, lower-income households face greater difficulty to adjust to the disrupted market conditions. By driving up housing prices, hurricanes exacerbate this pre-existing problem and further constrained their mobility. If an equitable recovery is the policy goal in post-disaster assistance, policymakers need to pay more attention to this interaction.

2.4.5 Heterogeneous Effects

In this section, we explore the heterogeneous effects of different wind speed exposure. For post-hurricane exposure, we separately define exposure to 64-96 kn wind speed (low-intensity), and 96 kn and above (high-intensity), as described in Section 2.2.2. We then estimate a variant of Equation (2.1) which allows for separate post-hurricane dynamics for each level of exposure.[34]

The results from this exercise are plotted in Figure 2.13. Most instances of exposure are of low intensity. Naturally, the price dynamics generated by low-intensity exposure are very similar to the overall pattern shown in Figure 2.4. The high-intensity exposure estimates, on the other hand, suffer from a lack of statistical power as a result of the significantly lower treatment saturation.[35] The general pattern is very similar: the average price jumps up in event year 0, peaks at event year 1, and tapers down in subsequent years. The point estimates are larger for high intensity, with a peak of 15% compared to 10% for low intensity. The effect also appears to be more persistent: the fourth year after high-intensity exposure still see a statistically significant effect of 5% higher price, whereas the corresponding estimate is small and insignificant for low intensity. Despite these patterns, the lack of statistical power constrains our ability to detect

[33] This is referred to as the "single crossing" property in Banzhaf and Walsh (2008).

[34] We pool the pre-hurricane indicators for the two types of exposure to enhance power.

[35] At most 1% of homes are ever sold from areas affected by category 3 and above wind speeds in any year before or after a hurricane, as such wind speeds rarely affected Florida over our sample period (recall Figure 2.3, with the caveat that it depicts tract-wide and not individual home category 3 hurricane exposure).

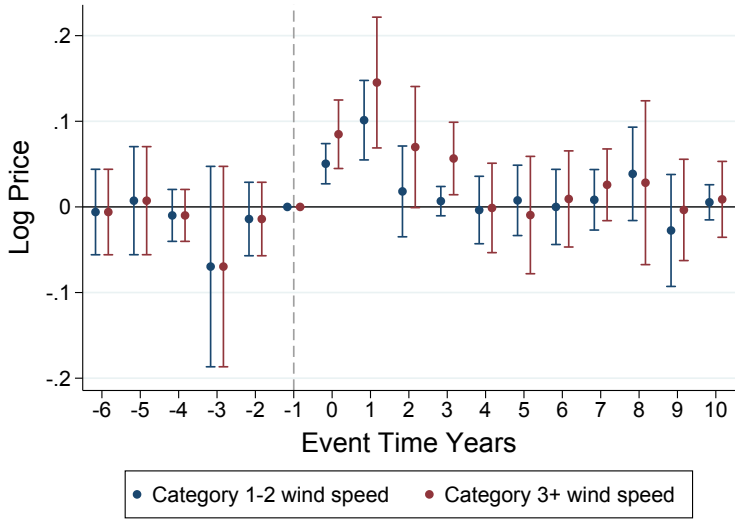


Figure 2.13: Heterogeneous Effects by Hurricane Intensity – Full Sample

Notes: Estimates for separate wind categories are plotted with their 95% confidence intervals. The model includes standard hedonic variables and census tract, month, and county-year fixed effects. Standard errors are clustered by county. The results are based on the full sample of Florida home buyers (N = 7,409,258).

differential effects. We find the same pattern for price and income in the HMDA subsample, where the high-intensity exposure is associated with a higher but imprecisely estimated peak (see Figures B.9 and B.10).

2.5 Discussions and Conclusion

This paper provides two sets of findings. First, we estimate the equilibrium dynamics in Florida housing markets in the wake of a hurricane. We find an increase in housing prices and a concurrent decrease in transaction probability, both lasting up to three years. The direction and duration of these effects are consistent with the mechanism in which part of the housing stock were first damaged by hurricanes and later restored. We find little empirical support for alternative explanations. Second, we examine demographic changes in local communities associated with these market adjustments, and find that incoming home buyers during the recovery period have

higher average income *conditional on* characteristics of homes transacted. This results in an enduring change in local economic profiles toward higher-income groups.

Several interesting implications emerge from these findings. For one, our results suggest there is a limit to the extent of any demand response. In the short run, any behavioral responses on the demand side of the market is dwarfed by the effect of the supply shock. It is likely that the potential home buyers are well informed about the general risk of hurricane exposure in Florida. For them, the realization of a hurricane conveys no more information beyond the very local effect pertaining to homes located in flood zones, as identified in previous studies. Moreover, we find remarkably little evidence of downward price adjustments beyond the response to the supply shock, which suggests that hurricanes do not fundamentally change the long-run demand for housing in affected areas. This is also illustrated in Figure B.7, where population trends in affected and unaffected counties are shown. There is no evidence of differential growth trends, whether or not they were recently impacted by a hurricane. This might be due to the growing importance of warm weather as an urban amenity in recent years. Locations with warm winter across the U.S., such as most of Florida, have seen steadily increasing demand (Glaeser and Gyourko, 2005).

In addition, we find that the transaction probability return to the baseline in three years, implying a close-to-full recovery in terms of the housing stock. This fast recuperation of physical capital is in line with past studies on capital destruction by war (Davis and Weinstein, 2002; Miguel and Roland, 2011; Feigenbaum, Lee, and Mezzanotti, 2017). In the case of hurricanes, however, this finding stands in contrast to Gallagher and Hartley (2017), who find a substantial portion of Katrina victims in New Orleans use insurance payments to pay down home mortgage and move away instead of repairing their homes. An important factor contributing to this difference is the incentive to rebuild, which is governed by the comparison between construction costs and housing prices (Gyourko and Saiz, 2004). Hurricane Katrina's impact in New Orleans is much more severe than most of the hurricanes studied in this paper, which might have driven up

reconstruction costs and affected local amenities and industries to the point of reducing long-run demand.

Finally, we return to the increase in average income of home buyers. As average buyer income increases after the strike without later dropping below their pre-hurricane baselines, these findings depict a richer demographic inhabiting hurricane-affected communities following a disaster. To the extent that such a demographic brings along more expensive assets, and spurs more economic development in these areas, this could result in more expensive future hurricane damage claims to publicly funded disaster relief organizations like FEMA. Moreover, the empirical relationship between housing supply and buyer income is interesting in and of itself. Our model highlights financial constraints as the potential drivers of such a relationship. This finding raises concerns on housing affordability issues in the recovery process, especially under hard credit constraints. We believe a more in-depth understanding of the roles played by hard and soft financial constraints is an important area for future work. It provides implications on promoting a more equitable recovery through housing policy from future (and likely more intense) natural disasters.

Acknowledgement

Chapter 2, in full, is currently being prepared for submission for publication of the material. Yanjun Liao; Yann Panassié. The dissertation author was a primary author of this chapter.

Data provided by Zillow through the Zillow Transaction and and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the authors and do not reflect the position of Zillow Group.

Chapter 3

Extreme Weather and the Politics of Climate Change: a Study of Campaign Contributions and Elections

Abstract

In this paper we study how extreme weather and natural disasters affect political outcomes such as campaign contributions and elections. We suggest that weather events associated with climate change may influence these outcomes by (1) leading individuals to update their beliefs about climate change and act on these beliefs politically; and (2) increasing the salience of climate change for individuals who are already politically active. In a short-run analysis, we find that the number of contributions to the Democratic Party increases in response to weekly average temperature, and that the effect is stronger among constituents in counties with more anti-environment incumbent politicians. In a medium-run analysis, we find suggestive evidence that after a natural disaster, total fundraising and the number of donors in an election cycle is higher if the incumbent has a more anti-environment stance, indicating a higher level of competition. Further, we find evidence that after a disaster, the more anti-environment the incumbent, the

higher the chance is of a challenger entering the race. Finally, we address alternative mechanisms and explanations for our results.

3.1 Introduction

Public opinion on key issues is thought to play a crucial role in shaping policies and elections in a democracy. Therefore, it is important to understand what factors contribute to the formation of these opinions as well as their political ramifications. One such key issue is climate change, which has received significant policy attention for decades.

In the United States, both the public and legislators remain divided when it comes to climate change, which has prompted several studies of Americans' attitudes towards it (Akerlof, Maibach, Fitzgerald, Cedeno, and Neuman, 2013; Howe, Mildenberger, Marlon, and Leiserowitz, 2015; Myers, Maibach, Roser-Renouf, Akerlof, and Leiserowitz, 2013; Spence, Poortinga, Butler, and Pidgeon, 2011; Zaval, Keenan, Johnson, and Weber, 2014). These studies document great heterogeneity in beliefs across the country and provide important insights into the factors that shape such beliefs. In particular, as extreme weather events and natural disasters have become widely associated with climate change (IPCC, 2013a), one recurring finding is that exposure to these events leads individuals to change their reported perceptions of the issue.

Even though weather anomalies have been shown to influence stated beliefs regarding climate change, less is known about how they may impact costly, real-world actions. Importantly, people may misreport their true preferences due to social or strategic considerations, so surveys of stated beliefs can be misleading. Furthermore, it is unclear whether these extreme weather events can, through changes in beliefs, have real-world political consequences. In particular, it remains an open question whether politicians will be held accountable for their positions on environmental issues as beliefs regarding climate change vary.[1]

[1] One reason changes in beliefs might not lead to political consequences is that climate change may not always be a top policy priority (Davis and Wurth, 2003; Guber, 2001).

In this paper, we present direct evidence of campaign finance and electoral responses to extreme weather events. We assemble a comprehensive dataset of extreme weather shocks, natural disasters, and U.S. House of Representative elections. These data allow us to examine various response margins, from campaign contributions to the competitiveness of elections and their outcomes. Moreover, we collect information on the environmental voting records of members of Congress to assess where they stand on the *anti-environment* to *pro-environment* spectrum.[2] Importantly, we test for differential effects of weather and disaster shocks depending on the environmental stance of incumbent politicians. This approach helps us shed light on whether environmental ideology is a driver of campaign contributions and political support for candidates in elections. Our results document a margin of political behavior in this context that is, to the best of our knowledge, novel in the literature. They also shed light on the mechanisms through which public opinion may shape climate change adaptation and mitigation policies.

Our study follows the literature closely in terms of choosing regression frameworks and constructing measures of weather shocks. In terms of their empirical strategies, previous studies can be classified into three categories. The first set of studies focuses on short-run weather shocks, over a period of a month or less (Egan and Mullin, 2012; Hamilton and Stampone, 2013; Joireman, Truelove, and Duell, 2010; Li, Johnson, and Zaval, 2011; Zaval, Keenan, Johnson, and Weber, 2014). The second set of studies uses medium-run temperature shocks, over a period of a month to a year (Deryugina, 2013). The third set of studies focuses on medium-run natural disaster shocks, also over a period of a month to a year (Lang and Ryder, 2016; Sisco, Bosetti, and Weber, 2017; Spence, Poortinga, Butler, and Pidgeon, 2011). In order to relate to the existing literature, in our study we focus on both short-run and medium-run temperature variations as well as medium-run natural disaster shocks.

In the short-run analysis, we examine how weekly temperature shocks affect contributions to Democratic candidates through ActBlue, an online fundraising platform, over the 2006-2012

[2]These terms are used for concise communication with the reader and do not necessarily represent the views of the authors on these issues or the politicians involved.

period. The identification relies on two features. First, temperature shocks are measured by deviations of weekly mean temperature from the historical average in the same month and location, which eliminates most cross-sectional variation and seasonality that may be correlated with unobserved confounding factors. Second, we control for a rich set of fixed effects including county, week-in-sample, and state-by-election cycle. The results reveal an extensive-margin response: a 1 ^{o}F increase in weekly average temperature has a contemporaneous effect of a 1.2% increase in the contribution rate, and a cumulative monthly effect of 2.7%. In our main analysis, we do not detect any intensive-margin effect. When looking across incumbent characteristics such as party membership and environmental attitudes, we find stronger responses to temperature shocks among constituents with more anti-environment incumbents. Together, these results suggest that following a temperature shock, Democratic candidates are rewarded for their pro-environment stance when the incumbents in the contributor's district are more anti-environment. This effect is mainly driven by an increase in the number of contributions.

In the medium-run analysis, we start by exploring how natural disasters in an election cycle interact with an incumbent's stance on environmental issues to influence both campaign finance and electoral outcomes. We examine the universe of individual and political action committee (PAC) contributions to candidates in U.S. House of Representatives races during election cycles 1990-2012. Importantly, this analysis includes both Republican and Democratic candidates. We use two distinct regression specifications. The first is our main specification and exploits variation in the incumbent's stance on environmental issues regardless of party affiliation, whereas the second utilizes within-party variation. Using the main specification, we find that after a natural disaster, total fundraising and the number of donors in an election cycle is higher if the incumbent has a more anti-environment stance, and there is some suggestive evidence that the effect is stronger for donations to challengers than for incumbents. Further, we find that after a disaster, the more anti-environment the incumbent is the higher the chance of a challenger entering the race, leading to a slightly lower re-election probability for the incumbent. When we

use within-party variation in incumbents' environmental stances, we find that estimates generally point in the same direction as before, but they are now imprecisely estimated and statistically indistinguishable from zero. The one exception is the probability of an incumbent entering the race as a response to a natural disaster, which is still higher the more anti-environment the incumbent is and remains statistically significant. In addition to natural disasters, we also focus on medium-run temperature shocks by using a similar methodology but treating hot and cold cycles as separate events. While we do not have enough power to rule out null results, the magnitude and direction of the effects of hot weather events are similar to that of natural disasters. In the case of cold weather events, we find that the effects are typically opposite in sign, which points to the possibility that in this context people react differently to hot and cold weather abnormalities. These results complement the short-run analysis since we are able to utilize the universe of contributions to House candidates, both Republican and Democrat, online and offline.

We explore a number of possible explanations for our results and argue that the most plausible explanation is the *environmental preference mechanism*. We suggest that there are two margins at play within this mechanism. First, extreme weather events prompt people to update their beliefs about climate change and become more politically active as a result. Second, extreme weather events make environmental issues more salient, serving as a call to action for people with static beliefs who are politically active. Furthermore, we note that in a competitive environment such as the electoral system, people on both sides of the climate change debate may be galvanized by extreme weather events, either independently or as a response to the other side's actions.

This paper contributes to several research areas. Firstly, it is one of the few existing studies that adopt a revealed preference approach to studying the effects of weather shocks on people's beliefs about climate change. Previous studies have focused on outcomes such as Google searches (Herrnstadt and Muehlegger, 2014; Lang and Ryder, 2016) and Twitter posts (Sisco, Bosetti, and Weber, 2017; Moore, Obradovich, Lehner, and Baylis, 2019). Li, Johnson, and Zaval (2011) show that respondents in a survey donated more money to an environmental charity if they

thought that day was warmer than usual, although this donation came from the fee they were awarded for completing the study. Importantly, the outcomes in this study are more costly and spontaneous, thereby providing more meaningful reflections of people's beliefs. Moreover, we also directly examine the political processes through which public opinions on climate change might shape policies, which has been a goal of this literature.

Secondly, our results contribute to the current understanding of the motivations for political giving. Our results show that the responses to short-run temperature shocks in terms of ActBlue contributions occurr mostly on the extensive margin. This is consistent with the mainstream view that individuals make campaign contributions for ideological reasons (Barber, 2016; Bonica, 2014; Ensley, 2009; Francia, Green, Herrnson, Powell, and Wilcox, 2003) and that they derive direct utility from contributing to their candidate of choice, as if they were consuming an ideologically-motivated consumption good (Ansolabehere, Figueiredo, and Snyder, 2003). Our findings are also consistent with the idea that online fundraising platforms like ActBlue have enabled such "political consumption" by significantly lowering transaction costs (Karpf, 2013). Our medium-run results provide insights into PAC contributions, whose motivation has not been unanimously agreed upon in the literature. While a prevalent theory is that PAC contributions have a *quid pro quo* nature, recent studies reveal that ideological considerations are also at play (Barber, 2016; Bonica, 2013; Bonica, 2014; Bonica, 2013; Snyder, 1990). Our evidence that PAC contributions also respond to natural disaster shocks lends further support to the ideological mechanism.

Thirdly, our results shed light on whether politicians are held accountable by constituents for their policy positions. On one hand, there is evidence to suggest that voters make seemingly irrational decisions, since same-day weather conditions or financial windfalls from lotteries have been shown to affect voting outcomes (Gomez, Hansford, and Krause, 2007; Bagues and Esteve-Volart, 2016; Meier, Schmid, and Stutzer, 2016; Jachimowicz, Menges, and Galinsky, 2017). On the other hand, there is also evidence that incumbents are held partially accountable

for their roles in disaster preparedness and post-disaster relief. For example, Arceneaux and Stein (2006) show that voters punish the incumbent mayor following a flood if they believed the city was responsible for flood preparation. Gasper and Reeves (2011) show that electorates may punish presidents and governors for severe weather damage, but that only the president is punished if a request by the governor for federal assistance is rejected. Healy and Malhotra (2009) show that voters reward the incumbent presidential party for delivering disaster relief spending. Our analysis complements these previous studies by exploring legislative elections to the U.S. House of Representatives. We also go beyond the direct impact of natural disasters in our context, by focusing on differential political consequences for incumbents with varying environmental stances. Further, our results shed light on those in Herrnstadt and Muehlegger (2014), who show that congresspersons are more likely to vote in favor of environmental legislation following natural disasters in their state. While there are multiple possible channels for their results, a higher probability of being challenged could certainly put pressure on incumbents to change voting behavior.

Finally, our findings also have important policy implications. Even though environmental issues are usually not front-and-center in U.S. elections, we demonstrate that the electorate is responsive to the salience of these issues. However, we caution that these responses may not be rational, since people may process shocks with psychological biases.[3] Further, these responses may also reflect a suboptimal allocation of attention. Nevertheless, our findings suggest that scientific outreach or other approaches that raise issue salience have the potential to induce substantial changes in political behavior.

The remainder of this article is organized as follows. Section 3.2 describes the data sources we draw upon for our analysis, while Section 3.3 describes our empirical strategy. In Section 3.4 we report and discuss the results. We conclude in Section 3.5.

[3]For example, Gallagher (2014) examines flood insurance take-up following flood events and finds a pattern indicative of availability bias or other forms of Bayesian learning with incomplete information.

3.2 Data

3.2.1 Database on Ideology, Money in Politics, and Elections

The political data we use come from the Database on Ideology, Money in Politics, and Elections (DIME) (Bonica, 2016). This database includes over 100 million campaign contributions made by individuals and organizations to candidates in local, state, and federal elections from 1979 to 2016. The main source of the data are administrative records from the Federal Election Commission (FEC). In addition to campaign finance data, the database includes characteristics of the candidates receiving contributions, as well as information on election outcomes.[4]

For our study of the impact of short-run weather shocks on political campaign contributions, we use a subsample of the individual contributions data from DIME. The reason is that while individual contributions have dates assigned to them, these dates do not always match the contribution date. Instead, they may indicate the date the campaign or candidate filed these contributions. Since we are interested in people's response to short-run, time-varying weather shocks, it is essential for us to make use of accurate date information. In order to circumvent this problem, we focus on contributions made through the online fundraising platform ActBlue, since the reported date matches the date of the contribution in this sample. We assess the implications of using ActBlue data in the following section.

For our study of the political consequences of natural disasters and medium-run weather shocks, we use the recipients file of the DIME database. This file contains information at the election cycle-by-candidate level and includes the total amount of funds raised by candidates from different sources, the seat sought, and the result of the election.

[4]For a detailed description of the database and data sources, please visit https://data.stanford.edu/dime.

3.2.2 ActBlue

In our short-term study, we focus on campaign contributions made through ActBlue, which is an online fundraising platform for Democratic candidates. The site was founded in 2004 and its popularity rose quickly thereafter. In our sample, ActBlue accounts for 4.3% of contributions and 0.8% of the total amount contributed to Democratic candidates.[5]

The main advantage of using ActBlue data is that the dates on ActBlue records are accurate, as they are electronically recorded at the time the contribution is made. Offline contributions, on the other hand, are at risk of being inaccurate in a non-random fashion, with the associated date often corresponding to the campaign's filing date. Naturally, relying on accurate date information is crucial for estimating responses to short-run weather variations. Another advantage of using ActBlue is that the contributions made on the site are typically small in quantity (see Figure C.2). For our purposes, these donations are very relevant, since they correspond to more spontaneous, lower stakes contribution decisions that may be affected by short-run weather variations.[6]

However, there are two concerns with using only ActBlue data for our short term study. The first concern is that the lack of an established Republican equivalent of ActBlue leaves us with only donations to Democrats.[7] This feature of our data does not allow us to see how donations to Republicans would respond, which is a shortcoming of our strategy. However, in the following sections we propose alternative methodologies to address this concern. The second concern is that it is unclear whether using ActBlue data will yield results that are representative of how contributions to Democrats as a whole respond to weather shocks. For example, it is possible that Internet contributors may be different in fundamental ways from the rest of contributors who still make up the majority. This issue has been investigated by Karpf (2013) and Wilcox (2008). Karpf (2013) suggests that the Internet brings about an increase in small donors by lowering

[5]Conversely, the total amount of contributions to Democrats is about 24 times the number of ActBlue contributions, and the total amount contributed to Democrats is 122 times the amount contributed through ActBlue. We keep these numbers in mind when assessing the magnitude of our coefficients later on.

[6]For an example of how contributions are made to a Democratic candidate through ActBlue, see Figure C.1.

[7]Rightroots, Big Red Tent, and Slatecard are examples, but their popularity has been far lower than ActBlue's.

transaction costs. They also suggest that this change has facilitated the flow of campaign funds towards more polarizing candidates. Meanwhile, Wilcox (2008) finds that Internet donors are much younger than other donors, but that those giving small amounts to Democrats online are actually similarly likely to consider themselves "ideologically extreme" than as larger donors are. However, these findings are taken from surveys conducted in the year 2000, while our main time period is 2006-2012, during which Internet use was more prevalent among the general population. For our purposes, even though Internet contributors may not be a mirror image of the general contributing population, focusing on these contributions allows us to hone in on lower cost, spontaneous decisions that may be affected by weather variations.

In order to further alleviate these representativeness concerns, we would ideally correlate changes over time in ActBlue donations to changes in non-ActBlue donations, given that we exploit time-varying weather shocks in our analysis. However, there are two difficulties associated with doing this. First, as stated above, the date information for the non-ActBlue data is unreliable. Second, ActBlue was founded in 2004 and has become more popular since then, meaning that the trend of donations made through ActBlue will likely differ from the trend of overall Democratic donations. However, even though exploiting the time dimension may be difficult, we can explore whether ActBlue data do a good job at explaining the cross-section of total donations to Democrats. In order to do this, we regress total donation amounts and counts at the state-by-election cycle level on ActBlue donations and counts. If the cross-section of ActBlue donations is representative of the total Democratic cross-section, it should have high explanatory power. Additionally, to account for the fact that ActBlue becomes more popular over time and may represent a larger portion of total donations, we let our coefficients vary by election cycle in alternative regressions.

The results of these regressions are in Table C.2. The first two columns refer to the total amount contributed and the next two refer to the number of of contributions. As can be seen in column (1), simply including the amount donated through ActBlue is a strong predictor of total donations, leading to an R^2 of 0.74. When we allow the effect to vary by election cycle, as

in column (2), the explanatory is even higher, with an R^2 of 0.86. When we consider counts of donations instead of amounts donated, the fit is slightly better, with an R^2 of 0.83 and 0.88 in columns (3) and (4), respectively. Finally, an interesting feature of Table C.2 is the time-varying estimates in columns (2) and (4). The estimates for earlier years tend to be larger than in later years, revealing that over time the portion of ActBlue donations in total Democratic donations is rising.[8]

3.2.3 League of Conservation Voters Scorecard

In order to capture the stance of incumbent politicians on environmental issues, we use the League of Conservation Voters (LCV) scorecard (also known as the *National Environmental Scorecard*). The LCV scorecard assigns percentage scores to U.S. congresspersons based on their voting records regarding environmental legislation introduced during a particular year.[9] According to the terminology used by the LCV, if a politician aligns with the LCV opinion on a vote, it is marked as a *pro-environment* action; conversely, if the politician does not align with the LCV on a vote, it is marked as an *anti-environment* action (League of Conservation Voters, 2007). For conciseness, in this paper we will follow this terminology and refer to politicians who frequently align with the LCV as pro-environment and to those who don't as anti-environment.[10]

More specifically, LCV scores range from zero to one with pro- and anti-environment voting records on either side of the spectrum. In this paper, we subtract the original scores from one so that a score of zero indicates that the politician has disagreed with the LCV on 0% of the votes selected (pro-environment); conversely, a score of one indicates that the politician has disagreed with the LCV on 100% of the votes selected (anti-environment).[11]

[8] It is worth pointing out that this trend stabilizes during the 2012 election cycle.

[9] The legislation included in the scorecard arises from a consensus among leading environmental and conservation organizations in the U.S.

[10] Disclaimer: these terms are used to facilitate communication with the reader and do not necessarily represent the views of the authors on these issues or the politicians involved.

[11] For more information about the LCV scorecard, please visit the LCV website at http://scorecard.lcv.org/.

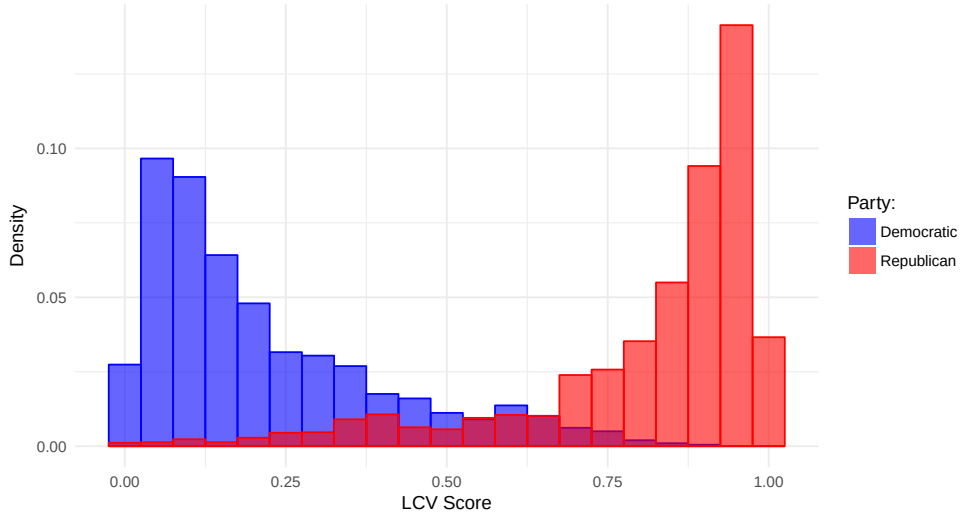


Figure 3.1: LCV Score Distribution by Party Affiliation

If politicians tend to vote along party lines when it comes to environmental issues, then we would expect to see a divide in the LCV scores of Democrats versus Republicans. This is certainly the case, as can be seen in Figure 3.1, where we plot the LCV score of U.S. congresspersons. As shown in the figure, most Democrats fall into the 0-0.25 range, meaning that they disagree with the LCV on less than 25% of the relevant votes. On the other hand. most Republicans fall in the 0.75-1 range, meaning that they disagree with the LCV more than 75% of the time. However, judging from the remaining mass in the 0.25-0.75 region, there is still substantial within-party variation in environmental voting records.

Additionally, the LCV score is an important indicator of whether the politician is a climate change denier. To show this, we obtain information on which congresspersons in the 112th caucus are climate change deniers from the site ThinkProgress.org.[12] Linking this information with LCV score data, we show that the probability of being a climate change denier is 51% for politicians with LCV scores above 0.5. Conversely, the probability of being a climate change denier for politicians with LCV scores below 0.5 is zero.

[12]See the article "The Climate Zombie Caucus Of The 112th Congress" at https://thinkprogress.org/the-climate-zombie-caucus-of-the-112th-congress-2ee9c4f9e46/.

3.2.4 Weather Shocks

We obtain historical weather data from the Global Historical Climatology Network Daily (GHCN-D) database. This database contains daily observations of maximum temperature and precipitation from more than 8,000 weather stations throughout the United States during 1960-2014. Using this information, we construct measures of county-level weather.[13]

We construct two measures of daily temperature shocks, which we later aggregate to the appropriate time frames for our analyses. The first measure is the daily deviation in maximum temperature from the historical climate normal in each county and month:

$$TmaxDev_{cmd} = Tmax_{cmd} - \overline{Tmax}_{cm}$$

where c is county, m is month, and d is day-in-sample. $Tmax_{cmd}$ is the contemporaneous daily maximum temperature in county c. $\overline{Tmax}_{cm}$ is the long-run average of maximum temperature for this county in the same month, calculated over the 30 preceding years. The second measure is a pair of indicators for whether the maximum daily temperature is abnormally high or low, compared to historical temperature distributions:

$$TmaxLow_{5,cmd} = 1(Tmax_{cmd} \leq Tmax_{5,cm})$$

$$TmaxHigh_{5,cmd} = 1(Tmax_{cmd} \geq Tmax_{95,cm})$$

where $Tmax_{5,cm}$ is the 5th percentile of the distribution of maximum temperatures in the same county and month over the 30 preceding years, and $Tmax_{95,cm}$ is the corresponding 95th percentile. As a result, $TmaxLow_{5,cmd}$ is an indicator for whether the contemporaneous temperature is lower than the 5th percentile of the historical distribution, whereas $TmaxHigh_{95,cmd}$ indicates whether it is higher than the 95th percentile of that distribution.

[13] If there is more than one weather station present in a given county, we take the average over all weather stations.

For our short-run analysis at the county-week level, we aggregate these daily measures by week, in order to obtain our independent variables of interest. We construct $TmaxDev_{cw}$, which is the average of $TmaxDev_{cmd}$ over the week, and is our primary temperature shock measure. We also construct $TmaxHigh_{cw}$ and $TmaxLow_{cw}$, which are the sums over the week of $TmaxHigh_{5,cmd}$ and $TmaxLow_{5,cmd}$, respectively. Importantly, these measures capture different aspects of weather shocks. Further, we also construct similar measures of precipitation deviations which we use as controls in our regressions.

For our medium-run analysis, we first calculate the number of hot days, defined as those above the 90th percentile of the historical distribution, experienced by the average person in each congressional district and election cycle.[14] We then rank district-cycle observations by this variable and assign *hot* status to those cycles in the top quartile. Similarly, we assign *cold* status to a district-cycle if it is in the top quartile ranked by number of cold days, defined as those below the 10th percentile of the historical distribution.

3.2.5 Natural Disasters

We obtain official disaster declaration data from the Federal Emergency Management Agency (FEMA) between 1990 and 2012. There are a total of 2,206 climate-related disasters, with the largest categories being severe storms, hurricanes, floods, fires and snow events (see Table C.1 for a detailed breakdown of disaster types). Importantly, these official records contain the period of the incident and the specific counties affected. Most declarations are not statewide.

Because we analyze the impact of natural disasters at the congressional district level, we need to assign disaster status to these for each election cycle. In order to do this, we first calculate the fraction of the population in a district living in counties hit by disasters,[15] and then assign disaster status to a district if this fraction exceeds 50%. It should be noted that this might not

[14] Our procedure makes use of the MABLE/Geocorr crosswalks developed by Missouri Census Data Center (2017), which partitions the population in a congressional district into its overlapping counties using Census data.

[15] This procedure also uses MABLE/Geocorr crosswalks (Missouri Census Data Center, 2017).

be the exact threshold at which natural disasters become salient politically, potentially leading to measurement error. However, in our data the majority of district-cycle observations have a fraction of the population affected of either zero or one, so adjustments to the threshold would not have substantial impact on our results.

3.3 Empirical Framework

Existing literature suggests that people might update their beliefs about climate change as a result of contemporaneous weather shocks, or even after a prolonged period of unusual temperature. Similarly, natural disasters may also lead to belief updating, given the salience of these events. Given that there are multiple possible relevant time frames and types of weather event, we examine the impacts of weather shocks both in the short-run and in the medium-run, as well as the medium-run impacts of natural disasters. In this section, we outline our empirical strategy for doing so.

3.3.1 Short-Run Weather Impacts

We first analyze the impact of weekly weather shocks on contributions to Democrats through ActBlue. Since Democratic candidates tend to be more pro-environment than non-Democratic candidates, we expect these donations to be responsive to weather shocks through the environmental preference mechanism we have proposed.

The estimating equations we use to study the impact of weather shocks on campaign contributions take the following general form:

$$Y_{cw} = \gamma' Weather_{cw} + \delta_w + \delta_c + \delta_{se} + \varepsilon_{cw} \tag{3.1}$$

where c is county, w is week-in-sample, s is state, and e is election cycle. Y_{cw} is the outcome of

interest, which can be either (1) the contribution rate (the number of contributions per million people), or (2) the average amount per contribution. These outcomes are calculated based on contributions made through the ActBlue platform. Furthermore, $Weather_{cw}$ is a vector of weather variables, which includes measures of temperature conditions as the key regressors and those of precipitation conditions as controls. Our coefficients of interest are the ones in γ' corresponding to the temperature variables.

We use three different specifications of $Weather_{cw}$ in our main analysis. The first specification includes only weekly contemporaneous deviations of temperature and precipitation from long-run climate normals, as they are expected to be the most relevant. It takes the following form:

$$Weather_{cw} = [TmaxDev_{cw}, PrcpDev_{cw}]^T.$$

One thing to note about this specification is that it does not allow for weather from past weeks to influence contribution decisions. However, it may be possible that the decision to contribute is delayed with respect to exposure to weather shocks. To allow for this, our second specification is:

$$Weather_{cw} = [TmaxDev_{cw}, ..., TmaxDev_{c,w-4}, PrcpDev_{cw}, ..., PrcpDev_{c,w-4}]^T$$

where we have added four lags of both temperature and precipitation deviations. This way, we capture the influence of weather shocks occurring during the past month. Furthermore, another weakness of our first specification is that it may not adequately capture current weather for contributions occurring earlier on in the week, since our measures are weekly averages. In order to address this, we propose the following third specification:

$$Weather_{cw} = [\overline{TmaxDev}^{2W}_{c,w}, \overline{PrcpDev}^{2W}_{c,w}]^T$$

where $\overline{TmaxDev}^{2W}_{c,w} = \frac{1}{2}(TmaxDev_{cw} + TmaxDev_{c,w-1})$ and similarly $\overline{PrcpDev}^{2W}_{c,w} = \frac{1}{2}(PrcpDev_{cw} +$

$PrcpDev_{c,w-1}$). Therefore, we average deviations in temperature and precipitation over the current and previous weeks.

In all of our specifications we also control for week-in-sample (δ_w), county (δ_c), and state-by-election cycle (δ_{se}) fixed effects. The county fixed effects absorb time-invariant factors in each county such as general political preferences and contribution behavior. The week-in-sample fixed effects control for confounding national events and the exponential growth of the platform itself. Finally, the state-by-cycle fixed effects account for slower-moving changes across states, such as whether the current president is politically aligned with the state, or new policies adopted by the State. We cluster standard errors at the county level.

Aside from the main specifications outlined above, we also consider four extensions. First, we estimate the effects of positive and negative temperature deviations separately, to shed light on whether people respond to these differently. Second, we estimate effects separately for each quarter of the election cycle, to study how these vary with the progression of campaigns. Third, we consider alternative measures of our independent variables by focusing on counts of extreme temperature events instead of average temperature. Fourth, we extend our main regressions to allow for heterogeneous effects depending on the environmental stance of the incumbents in the contributor's place of residence, to rule out unobservable confounding factors that may drive all contributions across time and location, and not only those that are environmentally motivated. More specifically, we examine whether counties where the majority of the population lives in districts represented by anti-environment incumbents exhibit stronger responses to weather shocks.[16] For more details on these additional specifications, see Appendix C.3.

[16]This is what we would expect as long as the Democratic candidates receiving contributions on ActBlue are more pro-environment on average.

3.3.2 Medium-Run Natural Disaster Impacts

Aside from studying weather shocks, we are interested in how fundraising and elections are affected by natural disasters in the medium run. Specifically, we study how this relationship varies depending on the environmental stance of the incumbent politician. In order to explore this issue, we focus on races for the U.S. House of Representatives during election cycles 1990-2012. We study campaign finance outcomes, such as the total funds raised and the funds raised by the challenger or incumbent separately. Further, we consider electoral outcomes such as the probability of the race being competitive and the probability of the incumbent being re-elected.

One concern that our approach will have to to address is that natural disasters may have significant effects on campaign contributions and other political outcomes through channels unrelated to environmental preferences and beliefs. For example, Stevens (2001) documents that following the September 11 terror attacks, individuals substituted away from campaign contributions and towards charitable giving. We expect this to be relevant in the case of natural disasters, since they often entail tragic consequences and loss of property.

In order to address the above concern, our research design consists in comparing the outcomes of congressional districts experiencing natural disasters whose incumbent politicians have an anti-environment voting record to the outcomes of other districts experiencing natural disasters but whose incumbents exhibit pro-environment voting records. By studying differential impacts by the environmental stance of incumbents, we hope to isolate the environmental preference mechanism. From an econometric standpoint, we run regressions of the following form:

$$Y_{de} = \beta_1 Disaster_{de} + \beta_2 LCV_{de} + \beta_3 Disaster_{de} \times LCV_{de} + \delta_d + \delta_e + \varepsilon_{de} \tag{3.2}$$

where Y_{de} is an outcome for a race in congressional district d during election cycle e; $Disaster_{de}$ is an indicator variable for whether over half of the population in a congressional district lives in counties affected by a natural disaster during that cycle; LCV_{de} is the LCV score of the incum-

bent;[17] and δ_d and δ_e are fixed effects for congressional district and election cycle, respectively. We cluster standard errors at the state level.

Our coefficient of interest is β_3. We interpret this coefficient as the difference in the outcome of a congressional district affected by a natural disaster whose incumbent congressperson has the most anti-environment voting record ($LCV = 1$), and the outcome of a similar, disaster-struck congressional district whose incumbent congressperson has the most pro-environment voting record possible ($LCV = 0$). Given that a one unit difference in the LCV score is a very large difference, we suggest scaling our estimates by the standard deviation of the LCV score in order to interpret them properly. Since the standard deviation of the LCV score is 0.2, we interpret our coefficients by dividing them by five.[18]

While the LCV score captures precisely the political dimension we care about, it is important to consider how it relates to party affiliation. As we can see in Figure 3.1, even though there is certainly within-party variation in the LCV score, the two are closely related. This raises a limitation of model (3.2): following a natural disaster, if people react differently to incumbents from different parties for non-environmental reasons, then the above coefficient of interest would be picking up on these factors as well. To address this issue, we propose an extension of the above model. The proposed model is:

$$\begin{aligned} Y_{de} =& \beta_1 Disaster_{de} + \beta_2 LCV_{de} + \beta_3 Disaster_{de} \times LCV_{de} \\ &+ \beta_4 R_{de} + \beta_5 Disaster_{de} \times R_{de} + \delta_d + \delta_e + \varepsilon_{de} \end{aligned} \tag{3.3}$$

where all variables are defined as in model (3.2) and we have now added R_{de}, an indicator variable for whether the incumbent is a Republican, as well as an interaction of this variable with $Disaster_{de}$. Our coefficient of interest is still β_3, which is now identified using variation in the

[17] In order to incorporate all available information at the time of the race, we average the LCV score of politicians for that election cycle and all past election cycles, using this measure throughout in our regressions.

[18] The standard deviation we use is that of the LCV score after controlling for the politician's party, which is 0.2. Without controlling for the politician's party the standard deviation is 0.32, which is slightly larger.

LCV score within the incumbent's political party.

It is important to note that model (3.3) comes with both advantages and disadvantages with respect to model (3.2). The main advantage, as previously discussed, is that it addresses the concern that people may respond to disasters differently depending on the incumbent party for reasons unrelated to the environment. The disadvantage is that the model does not make use of of valuable cross-party variation in environmental stances, which is perhaps the most visible and available to people when making decisions. Therefore, we consider these models to be complementary and keep these features in mind when interpreting results.

3.3.3 Medium-Run Weather Impacts

In the same way that campaign contributions and elections may be affected by natural disasters, they may also respond to shocks to medium-run temperature. The main difference with respect to natural disasters is that in this case there can be two distinct shocks: a hot weather shock and a cold weather shock. Since people may respond to these shocks differently, they enter our regressions separately. Using the same notation as in the natural disasters section, we study the impact of medium-run weather as follows:

$$Y_{de} = \beta_1 Hot_{de} + \beta_2 Cold_{de} + \beta_3 LCV_{de} + \beta_4 Hot_{de} \times LCV_{de} + \beta_5 Cold_{de} \times LCV_{de} + \delta_d + \delta_e + \varepsilon_{de} \tag{3.4}$$

where all variables are as previously defined, and Hot_{de} and $Cold_{de}$ are indicators for whether the election cycle was particularly hot or cold for a given district in an election cycle, constructed as described in Section 3.2. Aside from including these variables, we also add their interaction with the LCV score, following our earlier methodology.[19]

Our coefficients of interest are β_4 and β_5. As above, we interpret these coefficients as the difference in the outcome of a congressional district undergoing an unusually hot (cold) cycle,

[19] Standard errors are clustered at the state level.

whose incumbent congressperson has the most anti-environment voting record ($LCV = 1$), and the outcome of a similar district whose incumbent congressperson has the most pro-environment voting record possible ($LCV = 0$). Again, to ensure a reasonable interpretation of estimates we scale them by dividing by five. Finally, it is straightforward to extend this methodology to make use of within-party variation in the LCV score following a disaster, by adding R_{de} and its interaction with LCV as in model (3.3) above.

3.4 Results

In this section, we present our results in three parts: (1) short-run weather impacts on ActBlue contributions, (2) medium-run natural disaster impacts on campaign contributions and election outcomes, and (3) medium-run weather impacts on the same set of outcomes as in (2).

3.4.1 Short-Run Weather Impacts

In the short-run analysis, we investigate how ActBlue contributions are affected by temperature shocks in current and previous weeks. We examine two main outcomes. The first outcome is the contribution rate, defined as the number of contributions per million people in a county. This variable captures extensive-margin responses, i.e. whether temperature shocks motivate more or fewer contributions. The second outcome is the average amount per contribution, calculated as the total amount contributed divided by the number of contributions for each county-week. Absent any extensive-margin responses, this outcome measures intensive-margin responses, i.e. whether temperature shocks motivate larger or smaller donations from regular contributors. However, if extensive-margin responses are present, this outcome captures both the intensive-margin responses and potential changes in the composition of contributors.

In our sample period of 2006-2012, each county receives on average 2.4 donations per week or around 15.4 per million people, as can be seen in Table 3.1. The average amount of these

Table 3.1: Summary Statistics

Variable	N	Mean	Std. Dev.	Min.	Max.
ActBlue, 2006-2012 (county-week)					
Amount ($)	935,201	151.72	2061.05	0	583663.8
Count	935,201	2.42	23.10	0	5315
Average amount	935,201	13.19	125.67	0	32500
Count (per 1M pop)	935,201	15.40	135.84	0	38848.92
Population	935,201	110665.3	337187.7	403	9974868
Mean LCV	830,316	0.6720	0.3219	0	0.9800
Republican incumbent	830,316	0.6220	0.4849	0	1
Anti-env incumbent	830,316	0.6945	0.4606	0	1
Short-run Weather, 2006-2012 (county-week)					
Tmax dev. (F)	935,201	0.4453	6.621	-37.60	37.51
Tmax positive dev. (F)	935,201	2.789	4.036	0	37.51
Tmax negative dev. (F)	935,201	2.344	3.805	0	37.60
Tmax low ($<$ 5th pct)	935,201	0.3185	0.7855	0	7
Tmax high ($>$ 95th pct)	935,201	0.4719	1.0652	0	7
Prcp dev. (1/10mm)	935,201	0.0609	13.5375	-49.91	468.54
Natural Disasters, 1990-2012 (congressional district-cycle)					
Election cycle	4,478	2001.03	7.01	1990	2012
Disaster indicator	4,478	0.58	0.49	0	1
Num. donors (C)	4,095	232.18	1074.94	0	27122
Receipts ($1,000) (C)	4,095	326.72	703.71	0	9825.57
Num. donors (I)	4,095	536.83	1151.29	0	43718
Receipts ($1,000) (I)	4,095	1076.15	1108.93	6.77	25894.72
Receipts PACs ($1,000) (C)	4,095	42.57	101.39	0	1122.45
Receipts PACs ($1,000) (I)	4,095	466.07	387.56	0	4082.6
Receipts Ind. ($1,000) (C)	4,095	194.61	436.17	0	6248.91
Receipts Ind. ($1,000) (I)	4,095	545.11	745.35	0	21836.54
Competitive election	4,478	0.74	0.44	0	1
Unopposed election	4,478	0.18	0.38	0	1
Open race election	4,478	0.09	0.28	0	1
LCV score	4,478	0.5	0.34	0	1
Republican incumbent	4,478	0.47	0.5	0	1
Incumbent wins	4,095	0.95	0.22	0	1
Hot indicator	4,478	0.25	0.43	0	1
Cold indicator	4,478	0.25	0.43	0	1

donations is $13.2, demonstrating that ActBlue contributions are usually small. Table 3.1 also provides summary statistics on weather variables. The mean temperature deviation is 0.45 °F, since positive deviations are larger than negative ones on average. This pattern is also captured by the extreme temperature bins, as the number of extremely hot days exceeds the number of extremely cold ones.[20] This general warming trend is a common observation in the literature.

The estimates from equation (3.1) are reported in Table 3.2. In columns (1)-(3), we examine responses in the contribution rate. In column (1), the main weather variable is temperature in the current week. As detailed in Section 3.2.4, this variable is constructed as the deviation from the long-run temperature normal. Precipitation deviation is also included in the model as a control. The estimate on temperature is positive and significant at the 1% level. This suggests that more positive or less negative deviations are associated with a higher number of contributions. In particular, a 1 °F increase in weekly temperature is associated with about 0.19 additional contributions per million people, or an approximately 1.2% increase in the contribution rate relative to the mean.[21]

Column (2) adds four lags of temperature deviations. This specification allows us to examine whether temperature deviations from previous weeks might affect contemporaneous contributions. The estimated dynamics have two remarkable features. First, the estimates are across-the-board positive and significant. This pattern is inconsistent with a harvesting mechanism, where a temperature shock simply shifts the timing of contributions but not the overall amount. Instead, these effects could represent a net increase in contributions. Second, the impact of a temperature shock appears to decay over time, as the estimates are lower for temperature deviations that took place longer ago. The contemporaneous effect is smaller than in column (1), but the cumulative effect is larger: a 1 °F increase in weekly temperature is associated with a cumulative effect of 0.42 additional contributions per million population, or a 2.7% increase in

[20] Extremely hot days are those above the 95th percentile of the historical distribution, while extremely cold days are those below the 5th percentile of that distribution.

[21] γ_0 is the coefficient on $TmaxDev_{c,w}$. The contemporary effect as a percentage of the mean of the dependent variable is $\widehat{\gamma}_0/MeanD.V. = 0.186/15.40 \approx 1.2\%$.

Table 3.2: Actblue Donation Responses to Short-run Temperature Shocks

Dep. Var.	(1) Count /1M pop	(2) Count /1M pop	(3) Count /1M pop	(4) Avg. amount	(5) Avg. amount	(6) Avg. amount
$TmaxDev_{c,w}$	0.186***	0.134***		0.016	0.013	
	(0.057)	(0.045)		(0.047)	(0.046)	
$TmaxDev_{c,w-1}$		0.103***			–0.043	
		(0.035)			(0.034)	
$TmaxDev_{c,w-2}$		0.054***			0.055	
		(0.017)			(0.041)	
$TmaxDev_{c,w-3}$		0.072***			0.035	
		(0.025)			(0.035)	
$TmaxDev_{c,w-4}$		0.055**			–0.035	
		(0.023)			(0.032)	
$\overline{TmaxDev}^{2W}_{c,w}$			0.287***			–0.007
			(0.083)			(0.054)
R^2	0.209	0.204	0.209	0.054	0.054	0.054
N	944,172	935,201	941,672	944,172	935201	941,672
Mean D.V.	15.45	15.40	15.42	13.13	13.19	13.15
County F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Week F.E.	Yes	Yes	Yes	Yes	Yes	Yes
State-cycle F.E.	Yes	Yes	Yes	Yes	Yes	Yes

Notes: point estimates from equation (3.1) are shown. The dependent variable in columns (1)-(3) is the number of ActBlue contributions per 1 million population. The dependent variable in columns (4)-(6) is the average amount per ActBlue contribution. The sample consists of ActBlue contributions by week and county. Standard errors (in parenthesis) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

the contribution rate relative to the mean.[22]

The comparison between columns (1) and (2) suggests that omitting temperature lags might have led us to overestimate the contemporaneous effect and underestimate the overall effects of a temperature shock. Furthermore, column (2) shows that the current and previous week temperature deviations have the largest impact. In column (3), we take the average of deviations in these two weeks and use it as our main independent variable. As expected, the estimate is again positive and significant at the 1% level. Below, we will use this as our specification of reference and estimate variants of it in later extensions of our analysis.

Next, we re-estimate our models using the average contribution amount as the outcome variable. The results are in columns (4)-(6) of Table 3.2. In this case, all estimates are small and statistically insignificant, with no recognizable pattern. This could mean that temperature shocks do not induce intensive-margin responses or changes in contributor composition that are strong enough to be statistically detectable. It is also possible that these changes go in opposite directions and cancel each other out.

A feature of the specification used above is the implicit assumption that reducing a negative shock by one degree and increasing a positive one by the same amount has the same quantitative effect. However, positive and negative shocks might not be viewed symmetrically by individuals when it comes to belief formation. In Table C.3, we report separate estimates of the effects of positive and negative shocks. This is implemented by replacing each deviation regressor above with a pair of variables that separately capture the absolute values of its positive and negative components.[23] The results suggest that the observed effects are mainly driven by variation in positive deviations.

So far, our temperature measures are all in the form of deviations from the long-run norm. An alternative is to use measures of extreme temperature events because they might be more salient. As detailed in Section 3.2.4, we also construct variables that count the number

[22] $(\widehat{\gamma_0}+...+\widehat{\gamma_4})/MeanD.V. = (0.134+0.103+0.054+0.072+0.055)/15.40 \approx 2.7\%$.

[23] For more details on the specification, see Appendix C.3.

of extremely hot or cold days in a week. Table C.4 reports estimates based on these measures. Again, we find effects on the contribution rate but not the average amount. Specifically, one more extreme-hot day in a week is associated with a contemporaneous increase of 0.35 contributions per million people, or 2.3% of the mean contribution rate. Further, the cumulative monthly increase in this case is 7%.[24] For one more extreme-cold day, the contemporaneous and cumulative effects are decreases of 6.6% and 15.8%, respectively. These results share a number of qualitative similarities with the ones in Table C.3. They both show that heat shocks have extensive-margin impacts, but not intensive-margin ones. They both find these effects lasting up to a month and decreasing over time. On the other hand, the two sets of results differ on the effects of cold shocks. The average deviation specification suggests no effect, while the cumulative extreme weather approach suggests significant and negative effects. This suggests that a cold spell might not be as salient as one additional extremely cold day. It might also be related to the political discourse surrounding extremely cold weather events. In the past, some politicians have used extremely cold weather to argue against climate change, which could plausibly change public opinion in the opposite direction (Pierre-Louis, 2017).

Overall, our results so far are consistent with two potential mechanisms. First, temperature shocks prompt people to update their beliefs about climate change and become more politically aligned with Democratic candidates, who are typically pro-environment. Second, temperature shocks make environmental issues more salient in the election, reminding pro-environment donors to express their preferences through contributions. As these two mechanisms are similar, we do not attempt to disentangle them and refer to them collectively as the environmental preference mechanism. However, there are still a few alternative explanations for our results that are unrelated to environmental reasons. For example, weather is known to change voting behavior through psychological channels (Meier, Schmid, and Stutzer, 2016; Jachimowicz, Menges, and Galinsky, 2017). It might also affect time use or the expediency of online versus other contribution

[24]The contemporaneous effect of an additional hot day is $\widehat{\gamma}_0/MeanD.V. = 0.353/15.40 \approx 2.3\%$ of the mean of the dependent variable.

channels.[25]

To address these concerns and shed light on the mechanisms behind our results, we examine heterogeneous effects based on incumbent characteristics.[26] To enhance statistical power, we estimate effects by using a single temperature variable, the mean deviation in the current and past week.[27] Furthermore, we also include a measure that characterizes the population-weighted environmental stance of the incumbent congresspersons in a county, and its interaction with the temperature variable. We are interested in the coefficient associated with the interaction term, which shows how the effects of weather shocks vary according to incumbent characteristics.[28] Three characteristics are examined in separate regressions: (1) population-weighted mean LCV score (mean = 0.672); (2) whether over half of the population has a Republican incumbent (mean = 62.2%); and (3) whether over half of the population has an incumbent with an LCV score above 0.5 (mean = 69.5%). In this analysis, we restrict our sample to competitive races. The summary statistics suggest that the counties in our sample tend to have incumbents who are unfavorable to environmental protection and more likely to be Republicans. This could be due to people supporting a Democratic challenger being more active in online contributions than those supporting a Democratic incumbent.

The results are reported in table 3.3. In terms of the baseline effects, as before, we find a positive and significant effect of temperature deviations on the contribution rate, but not on the average contribution amount. Importantly, by examining the interaction term we can now see that the effect on contributions is larger when the incumbents have a more unfavorable view of environmental protection. This is true for all three measures of incumbents' environmental positions. The scale is also economically important. According to column (1), when the mean

[25] Section 1.6 presents a detailed discussion of the alternative mechanisms.

[26] We do not observe which candidates receive the contribution in the ActBlue data, only the place of residence of the donor. This limits our investigation to incumbent characteristics. While many contributions are directed to candidates outside of the district of residence of the contributor, we think this is a meaningful margin of giving behavior to study. For example, environmentally motivated donors may look to other congressional district races if the district they reside in is a very safe seat held by an anti-environment politician.

[27] Baseline estimates are reported in Table 3.2, columns (3) and (6).

[28] For more details on the specification, see Appendix C.3.

Table 3.3: Heterogeneous Effects by Incumbent Characteristics

Incumbent Char.	LCV		Republican		1(LCV>0.5)	
Dep. Var.	(1) Count /1M pop	(2) Avg. amount	(3) Count /1M pop	(4) Avg. amount	(5) Count /1M pop	(6) Avg. amount
$\overline{TmaxDev}_{c,w}^{2W}$	0.193**	−0.115	0.228**	−0.102	0.241**	−0.110
	(0.098)	(0.118)	(0.103)	(0.075)	(0.110)	(0.088)
IncChar	10.299**	2.016	5.336**	0.764	7.043**	1.538
	(4.658)	(1.318)	(2.674)	(0.854)	(3.129)	(0.954)
$\overline{TmaxDev}_{c,w}^{2W} \times IncChar$	0.152**	0.128	0.107***	0.117**	0.079**	0.118*
	(0.068)	(0.110)	(0.037)	(0.056)	(0.037)	(0.118)
N	830,316	830,316	830,316	830,316	830,316	830,316
R^2	0.2070	0.0550	0.2070	0.0551	0.2070	0.0551
Mean D.V. (All/R/Anti)	12.29	11.09	11.12	8.16	10.33	7.48
Mean D.V. (All/D/Pro)	12.29	11.09	14.22	15.93	16.75	19.30
County F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Week F.E.	Yes	Yes	Yes	Yes	Yes	Yes
State-cycle F.E.	Yes	Yes	Yes	Yes	Yes	Yes

Notes: point estimates from equation (C.3) are shown. The dependent variables in columns (1), (3), and (5) is the number of ActBlue contributions per 1 million population. The dependent variable in columns (2), (4), and (6) is the average amount per ActBlue contribution. The temperature shock measure is the average temperature deviation in the current and past week. The IncumbChar variable is the mean LCV in columns (1)-(2), whether the county is dominated by Republican incumbents in columns (3)-(4), and whether the county is dominated by incumbents with anti-environment voting records in columns (5)-(6). Standard errors (in parenthesis) are clustered by county. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

LCV score increases by one standard deviation, the scale of the positive effect goes up by 15.8% of the baseline effect, corresponding to that of a county represented fully by pro-environment incumbents.[29] Further, the effect of temperature shocks in Republican-dominated counties is 47% larger than in Democrat-dominated ones, according to column (2). The corresponding difference in column (3) is 32.7%.[30] These results suggest that following a temperature shock, Democratic candidates are rewarded for their pro-environment stance when the incumbents in the contributors

[29]LCV incremental effect: $SD(\overline{LCV}) \times \widehat{\beta}_3/\widehat{\beta}_1 = 0.2 \times 0.152/0.193 \approx 15.8\%$.

[30]Republican incremental effect: $\widehat{\beta}_3/\widehat{\beta}_1 = 0.107/0.228 \approx 47\%$. Anti-environment incremental effect: $\widehat{\beta}_3/\widehat{\beta}_1 = 0.079/0.241 \approx 32.7\%$.

district are more anti-environment.[31]

Next, we explore how effects vary depending on the progression of campaigns. We use a similar specification, interacting the two-week average deviation measure with a set of eight indicators for quarters in the election cycle. This allows us to obtain a separate estimate for each quarter. The results are plotted in Figure 3.2 and reported in Table C.5.

For the contribution rate, the estimates are positive and significant for quarters 2-5 and 8. As expected, the effect is especially pronounced in the last quarter, as the election date draws near and campaigning efforts ramp up. The second largest effect is in quarter 3, or June to August in the year prior to the election. This might be driven by presidential candidates declaring their candidacy around that time. When we examine average amount as an outcome, the pattern is flipped. The effect is large and negative in quarter 8, followed by quarter 3, and fluctuates around 0 for the rest. This pattern is most likely driven by a change in the composition of the contributors – heat shocks draw in more small-amount contributions in this case.[32] Given this negative impact on the average contribution amount, we also examine the contribution amount per million people, which is a combined outcome of the two margins. As the extensive and intensive margins often change in opposite directions, the overall effect is mostly small except for quarters 5 and 8, where the increase in the extensive margin appears to dominate, leading to increased contribution amounts in those quarters.

Lastly, we perform a back-of-the-envelope calculation to infer the effects of temperature shocks on total Democratic contributions, instead of only ActBlue contributions. This calculation

[31] It is also possible to examine the coefficients on the incumbent characteristics, which are all positive and significant when looking at contribution rates. Since we have controlled for county fixed effects, this parameter identifies the increase in the contribution rate corresponding to an increase in anti-environment incumbents within a county. Together, these results suggest that people make compensatory contributions when politicians ideologically different from them are elected in their district.

[32] When interpreting the pattern, one caveat is that the average-amount variable is constructed by dividing total contribution amount by the number of contributions. This procedure might transform any measurement error in the denominator into non-classical ones in the outcome. The estimates might thus be biased, the confidence interval invalid, especially when the mean of the denominator is close to zero. A remedy of this issue is beyond the scope of this paper. In quarters 3 and 8, however, the number of contributions is sufficiently large that we are more confident in the estimates.

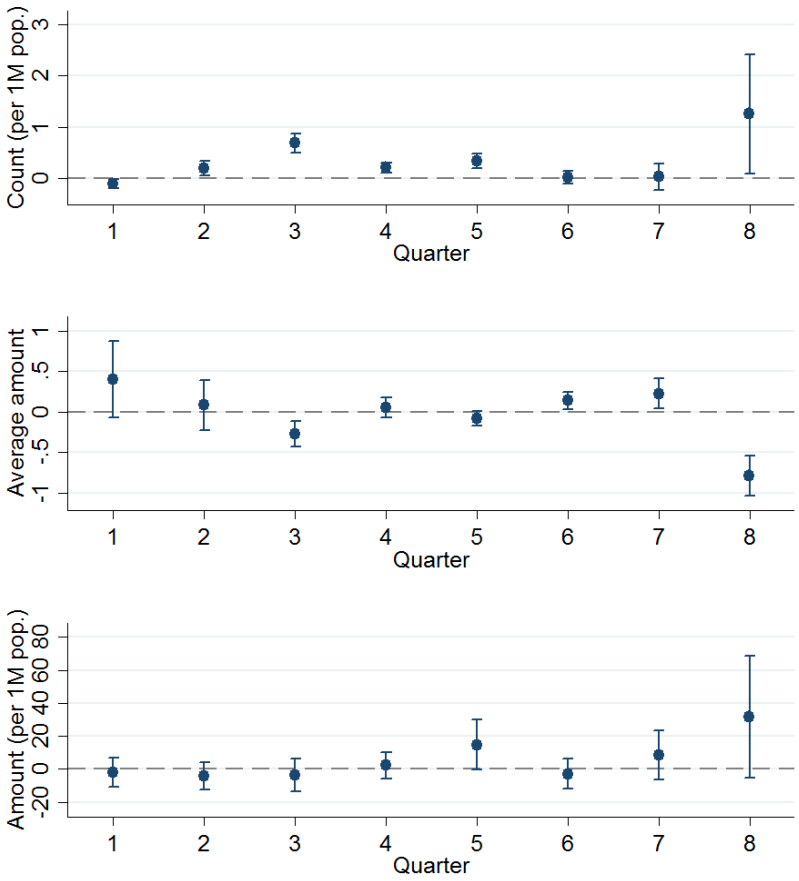


Figure 3.2: Variations of Estimates Across Quarters in Election Cycle

Notes: Point estimates from equation (C.4) and 95 percent confidence intervals are shown. The outcome variables, as displayed next to the y-axis, are based on ActBlue records. Standard errors are clustered by county. All regressions control for county, week-in-sample, and state-by-cycle fixed effects.

allows us to gauge the actual magnitudes of our previous estimates. In our sample, the total number of Democratic contributions is 24 times that of ActBlue contributions, and the total amount is 122 times larger. Using these numbers and our estimates in Table 3.2, we find that the contemporaneous effect of a $1\,^{o}$F increase in weekly mean temperature corresponds to a total increase of 3.2 contributions or $215.6 per million people per week.[33] The corresponding cumulative effects are 10 contributions and $672.6. It should be noted that this calculation relies on the assumption that total Democratic donations and ActBlue donations react similarly to weather shocks. In reality, we may expect ActBlue donations to react more strongly given the small and spontaneous nature of these contributions, meaning that these calculations are likely to represent upper bounds on the actual effects.

3.4.2 Medium-Run Natural Disaster Impacts

In the previous section, we show that heat shocks are associated with more contributions to Democratic candidates through ActBlue. As they are often the more pro-environment candidates, it is consistent with the mechanism that people update their beliefs about climate change and express their preference for pro-environmental policies through contributions. However, with only contribution records for the Democrats, it is difficult to fully disentangle this environmental preference channel from other factors that drive contributions in general.

In this section, we study the impact of natural disasters on campaign finance and elections. This analysis complements the previous results, as we explicitly account for politicians' environmental attitudes and include contributions to both Democratic and Republican candidates. As suggested by previous studies, natural disasters can draw public attention to climate change (Lang and Ryder, 2016; Sisco, Bosetti, and Weber, 2017) and they also bring about political ramifications for the incumbents (Arceneaux and Stein, 2006; Gasper and Reeves, 2011; Healy

[33] Δ number of contribution = $\widehat{\gamma_0} \times ratio(Dem/ActB) = 0.134 \times 24 = 3.22$. Δ total amount = Δ number of contributions $\times$ average amount $\times ratio(Dem/ActB) = 0.134 \times 13.19 \times 122 = 215.63$.

and Malhotra, 2009). Building on this literature, we hypothesize that anti-environment leaning incumbents will be held accountable for their environmental stance when a natural disaster strikes, leading to increased support for challengers. However, we do not necessarily expect support for these incumbents to remain unchanged, since they may intensify their fundraising efforts as a response to the increased support for challengers. In other words, we are agnostic about what happens to incumbent support.

In order to study the impact of natural disasters on campaign finance and elections, we examine how the impact of a natural disaster varies depending on the environmental voting record of the incumbent politician (LCV score), while focusing on competitive races. We run regressions that follow the specifications in Equations (3.2) and (3.3). When interpreting the magnitude of our coefficients, we divide them by five so that they correspond to the effect of a one-standard-deviation difference in the LCV score of a candidate.[34]

We start by studying the effects of natural disasters on the amount of funds raised during the election in Table 3.4. Panel A contains the results using our main specification, described in Equation (3.2). In column (1) we can see that total fundraising following a natural disaster is higher in those districts with more anti-environment incumbents. More precisely, a one-standard-deviation increase in the LCV score of the incumbent translates to a $78,000 increase in total fundraising during a cycle (5.1% D.V. mean), when a natural disaster strikes. We can further break down the fundraising effect by distinguishing between the types of donor. In columns (2) and (3) we can see that funds from PACs go up by about $20,000 (3.8% D.V. mean) and funds from individuals go up by about $38,000 (4.6% D.V. mean).

In order to assess whether there is increased support for challengers, incumbents, or both, we turn to columns (4)-(6) in Table 3.4. In column (4), we find that a one-standard-deviation

[34]By doing this, we make sure we are taking into account the variation in the LCV variable. Since the standard deviation of the LCV score (after controlling for the politician's party) is 0.2, we interpret our coefficients by dividing them by five. This allows us to interpret results as those corresponding to a one-standard-deviation difference in the LCV score of a candidate, instead of a one-unit difference, which would correspond to the entire range of the variable.

increase in the LCV score of the incumbent translates to a $39,000 increase in fundraising by challengers during a cycle, when a natural disaster strikes. This result substantiates the hypothesis of increased support for challengers in races with anti-environment incumbents. In column (5), we find that a one-standard-deviation increase in the LCV score of the incumbent also translates to a $39,000 increase in fundraising by incumbents following a natural disaster. These results are consistent with the hypothesis that incumbents may perceive the strengthened support for the challengers and react by increasing fundraising efforts themselves. Furthermore, even though fundraising increases for both the incumbent and the challengers, we can assess the relative strength of these effects by studying the share of funds raised by the challenger. The results of this exercise are in column (6). According to these estimates, the share of funds going to challengers is higher when the incumbent is anti-environment, although the estimate is not statistically significant. However, it is worth pointing out that the magnitude of the effect is economically non-negligible: a one-standard-deviation increase in LCV is associated with a 0.6 p.p. increase over a baseline of 19.6 p.p. Further, another way to assess whether the increase in support is greater for challengers or incumbents is by comparing the effects in columns (4) and (5) to their respective dependent variable means. In this case, the effect for challengers appears to be stronger, as the baseline level of donations to incumbents is higher and the estimated effects are similar. While the evidence is weak, it is consistent with the hypothesis that the increase in support may be stronger for challenges than for incumbents.

Panel B of Table 3.4 includes the results of our within-party specification described in Equation (3.3). As previously mentioned, this specification is not necessarily superior to our main specification, since within-party differences in environmental stances may be harder for people to perceive. With the exception of the share of funds raised by challengers, all the estimates on the interaction term are of the same sign as those in Panel A, albeit with varying relative magnitudes and levels of statistical significance. For example, the effects on giving from PACs and giving to the incumbent remain significant and of similar size as those in Panel A, whereas the effects

Table 3.4: The Effects of Natural Disasters on Amount Raised ($1,000)

Dep. Var.	(1) Total	(2) Total (PAC)	(3) Total (Ind.)	(4) Challenger	(5) Incumbent	(6) Share (C)
			Panel A: Main Specification			
Disaster	-316.4***	-78.15***	-164.8***	-184.7***	-131.7**	-0.0328***
	(70.39)	(24.16)	(55.17)	(46.02)	(54.61)	(0.0122)
LCV	-253.2	-87.08*	-39.90	-226.7**	-26.53	-0.0453*
	(305.8)	(48.97)	(259.3)	(107.3)	(221.0)	(0.0249)
Disaster × LCV	389.7***	101.8***	191.8**	194.1**	195.5**	0.0299
	(125.6)	(35.78)	(88.96)	(93.66)	(73.00)	(0.0245)
R^2	0.413	0.533	0.388	0.262	0.457	0.287
			Panel B: Within-Party Specification			
Disaster	-300.2***	-79.88***	-152.3***	-168.3***	-131.8**	-0.0247**
	(72.39)	(24.44)	(55.54)	(48.70)	(56.14)	(0.0112)
LCV	228.1	33.41	200.2	-11.27	239.4	0.0371
	(406.4)	(86.51)	(332.0)	(176.0)	(295.7)	(0.0385)
Disaster × LCV	255.7	122.2*	84.72	49.67	206.1*	-0.0422
	(227.9)	(61.79)	(154.2)	(156.7)	(115.5)	(0.0272)
R^2	0.414	0.534	0.388	0.263	0.458	0.290
Observations	3299	3299	3299	3299	3299	3299
Mean D.V.	1535.8	535.7	823.8	405.6	1130.2	0.196
Cycle F.E.	Yes	Yes	Yes	Yes	Yes	Yes
C.D. F.E.	Yes	Yes	Yes	Yes	Yes	Yes

Notes: point estimates from equations (3.2) and (3.3) are shown. The dependent variable in columns (1)-(5) is the amount of money raised in an election cycle from different sources in a given district, expressed in thousands of dollars. The dependent variable in column (6) is the share of total funds raised by the challengers. The sample includes competitive elections. Standard errors are clustered at the State level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

on giving to challengers and by individuals are reduced considerably and are not significant. Therefore, under this specification, we do not find evidence that giving to challengers is higher following a disaster if the incumbent has a more anti-environment voting record.

Having examined natural disaster impacts on fundraising, it is natural to also look at the number of donors to challengers and incumbents. We do so in Table 3.5, and start by examining the results for our main specification in Panel A. The coefficients in columns (1)-(3) are all positive and significant, indicating that, following a natural disaster, there is a higher number of donors from all sources if the incumbent has a more anti-environment stance. For example, in column (1) we can see that a one-standard-deviation increase in the LCV score of the incumbent translates to 97 additional donors when there is a natural disaster. When we break it down by giving to incumbent versus challengers, the effect sizes are about 42 and 55 additional donors, respectively. As in the case of fundraising, we have seen that the number of donors increases both in the case of challengers and incumbents, and we would again like to explore the relative strength of these effects by looking at the share of donors contributing to challengers. In column (4), we find a similar effect, indicating that the share of donors to challengers is slightly higher, although it is not significant. However, we again note that the effect for challengers is larger when comparing it to its average, suggesting that the effect on giving to challengers might be relatively stronger.

In panel B, we report estimates from our within-party specification. In this case, none of the estimates are statistically significant. The estimated effect on giving to challengers is now negative, although the confidence interval is very wide, including economically significant effects in either direction. Therefore, the results in this panel are uninformative of whether within-party variation in the LCV score influences the number of donors following a natural disaster.

Next, we take the analysis a step further by exploring how natural disasters affect elections. There are several reasons for this. First, the campaign finance consequences of natural disasters, as shown above, may in turn affect electoral outcomes. Second, natural disasters may affect

Table 3.5: The Effects of Natural Disasters on the Number of Donors

Dep. Var.	(1) Total	(2) Challenger	(3) Incumbent	(4) Share (C)
	Panel A: Main Specification			
Disaster	-335.1***	-118.2**	-216.8**	-0.0301**
	(118.5)	(50.78)	(105.0)	(0.0139)
LCV	-515.5	47.11	-562.6	0.00104
	(492.9)	(153.4)	(392.1)	(0.0279)
Disaster × LCV	484.9**	211.4*	273.5*	0.0334
	(201.1)	(111.1)	(155.5)	(0.0264)
R^2	0.297	0.216	0.320	0.283
	Panel B: Within-Party Specification			
Disaster	-305.3**	-88.64	-216.7**	-0.0207
	(122.0)	(63.64)	(106.2)	(0.0130)
LCV	-167.3	184.4	-351.7	0.0978**
	(611.3)	(236.5)	(507.9)	(0.0407)
Disaster × LCV	220.9	-58.37	279.2	-0.0507
	(306.4)	(252.9)	(172.6)	(0.0333)
R^2	0.297	0.217	0.321	0.286
Observations	3299	3299	3299	3299
Mean D.V.	862.6	288.2	574.4	0.208
Cycle F.E.	Yes	Yes	Yes	Yes
C.D. F.E.	Yes	Yes	Yes	Yes

Notes: point estimates from equations (3.2) and (3.3) are shown. The dependent variable in columns (1)-(3) is the number of donors in an election cycle from different sources in a given district. The dependent variable in column (4) is the share of total donors corresponding to the challengers. The sample includes competitive elections. Standard errors are clustered at the State level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the composition of elections by influencing the choice of prospective challengers to enter the race. Third, natural disasters may prompt issue voting by environmentally conscious constituents and directly influence the results of the election. Therefore, we focus on the following four outcomes: whether the election is competitive (i.e. there is a challenger), whether the incumbent runs unopposed, whether there is an open seat election (i.e. the incumbent does not run for re-election), and whether the incumbent is re-elected.

The results of this analysis are in Table 3.6. In panel A we present results from our main specification, with the first three columns focusing on the election type and the fourth column focusing on the election outcomes. When it comes to the election type, we can see that if the incumbent in a disaster-struck district has a higher LCV score, they are more likely to face a challenger in the race. For a one-standard-deviation increase in the LCV score, the probability of the race being competitive following a disaster is almost 3% higher. As a result, the effect on the incumbent running unopposed is negative and similar in magnitude. The probability of an open seat election taking place is unaffected. Importantly, these results are in line with our hypothesis of increased support for challengers, although they relate to the presence of challengers (extensive margin) and not the strength of the support they receive (intensive margin). Potential challengers may join the race simply because of the increased funds they are able to raise, or because they recognize an opportunity to run on a pro-environment platform given the incumbent's record. As for the election outcome, we examine the impacts on the incumbent's re-election probability in column (4). The estimates show that the probability of the incumbent winning the election is lower following natural disasters if the incumbent has an anti-environment voting record. Specifically, for a one-standard-deviation difference in the LCV score this effect is about 1.3%. Note that there are at least two possible mechanisms behind this effect: first, given that we observe an increased probability of the election being competitive, this can mechanically lead to a reduced probability of an incumbent victory, as the incumbent is now more likely to face a challenger. Second, the effect may be driven by stronger opposition to the incumbent as a result of their voting record,

independently of the presence of a challenger.

In panel B we examine the effects of natural disasters on elections using our within-party specification. The results look similar to those in panel A. Importantly, we still find evidence that if the incumbent in a disaster-struck district has a higher LCV score, they are more likely to face a challenger in the race. This result, which is robust and significant across specifications, is in line with our hypothesis of increased support for challengers. However, under this specification we no longer find evidence that the probability of the incumbent being re-elected is reduced, although the estimate is still negative.[35]

3.4.3 Medium-Run Weather Impacts

In the previous section we studied how natural disasters affected campaign finance and elections in the medium run. However, it is possible for other weather events to have a similar effect in the medium run, such as temperature shocks. In this section, we focus on how these shocks influence our political outcomes. Importantly, we treat temperature shocks as separate from natural disasters for two reasons: first, temperature shocks can be either hot weather shocks or cold weather shocks, each with their own possible ramifications. While people may interpret extremely hot weather as evidence of climate change, this may not be the case for extremely cold weather, which some may interpret as evidence against climate change, as we argued in the section on short-run impacts. Second, temperature shocks may be less salient than natural disasters, in part because the latter often results in property damage.

We start by looking at the effect of extreme temperature on the amount raised in an election in Table C.6. As explained in previous sections, we now study the effect of two events at

[35]Note how these findings related to election composition may affect the interpretation of our earlier results on amount raised and the number of donors. Since we focus on competitive elections, and the interaction of *LCV* and *Disaster* influences whether an election is competitive in the first place, our earlier results may capture a selection effect as well as a direct effect. The selection effect is likely to induce weaker challengers – those who would have not entered the race in absence of a natural disaster – to enter the race. This would likely bias our results towards lower fundraising, especially in the case of challengers, which could explain the absence of an effect on the share of funds going to challengers in Table 3.4.

Table 3.6: The Effects of Natural Disasters on Elections

Dep. Var.	(1) Competitive	(2) Unopposed	(3) Open Seat	(4) Incumbent Re-Election
		Panel A: Main Specification		
Disaster	-0.0352	0.0391	-0.00390	0.0470**
	(0.0246)	(0.0264)	(0.0158)	(0.0189)
LCV	-0.103**	0.0484	0.0547	0.104**
	(0.0493)	(0.0512)	(0.0410)	(0.0436)
Disaster × LCV	0.141***	-0.117**	-0.0239	-0.0629**
	(0.0411)	(0.0444)	(0.0232)	(0.0254)
R^2	0.169	0.232	0.103	0.171
		Panel B: Within-Party Specification		
Disaster	-0.0358	0.0353	0.000543	0.0420**
	(0.0241)	(0.0265)	(0.0176)	(0.0181)
LCV	-0.0853	-0.0402	0.125**	0.0243
	(0.0786)	(0.0602)	(0.0614)	(0.0477)
Disaster × LCV	0.147**	-0.0842	-0.0632	-0.0166
	(0.0640)	(0.0686)	(0.0531)	(0.0370)
R^2	0.169	0.233	0.104	0.173
Observations	4478	4478	4478	4095
Mean D.V.	0.737	0.178	0.0855	0.948
Cycle F.E.	Yes	Yes	Yes	Yes
C.D. F.E.	Yes	Yes	Yes	Yes

Notes: point estimates from equations (3.2) and (3.3) are shown. The dependent variable in columns (1)-(3) is the probability of a congressional race being of a certain type (competitive, unopposed, open seat). The dependent variable in column (4) is the probability that the incumbent is re-elected. Columns (1)-(3) include all elections and column (4) excludes open seat elections. Standard errors are clustered at the State level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

once: whether the election cycle was abnormally hot or whether it was abnormally cold. In line with our previous analysis, we interpret the interaction terms to see how the effects of weather events vary with the LCV score. We focus on specifications of the type in Equation (3.4). We can see that for abnormally hot cycles, the estimated coefficients on the interaction term are positive, as in the case of natural disasters, but they are insignificant and imprecisely estimated. The effects for abnormally cold cycles are also typically positive (with the exception of donations to PACs) and imprecisely estimated. In general, these results are not informative of the fundraising consequences of abnormal weather in the medium run. Next, we look at effects on the number of donors in Table C.7. The estimated effects are typically positive for abnormally hot election cycles and typically negative (with the exception of donors to challengers) for abnormally cold cycles. However, as before, these are imprecisely estimated. As in the natural disasters section, we also explore the effects of extreme temperatures on elections in Table C.8. For hot cycles, the effects point in the same direction as natural disasters, with races more likely to be competitive and a lower probability of incumbent re-election. However, these effects are smaller and not statistically significant. In the case of cold cycles, these relationships are largely reversed, although the estimates are still not statistically significant.

Given the level of statistical significance of our estimates, we cannot rule out that these are null results. A possible explanation for this is the salience of temperature events compared to natural disasters, as mentioned above. However, it is worth noting that in some of these results suggest the effects of hot and cold cycles are opposite in sign, with the effects of hot cycles looking more like those of natural disasters. While suggestive at best, these results point to the possibility that in this context people react differently to hot and cold weather abnormalities.

3.4.4 Mechanisms and Limitations

Throughout this paper, we have proposed the environmental preference mechanism – which contains two separate margins – as the driver of our results. First, we suggest that natural

disasters and weather shocks lead people to update their beliefs about climate change, which in turn enables them to act on these beliefs politically. Second, we suggest that extreme weather events simply increase the salience of climate change and other environmental issues, which in turn incentivizes those who have already made up their mind on these issues to take action. We do not attempt to distinguish between these two mechanisms, and acknowledge that our observations may stem from either or both of these. However, there may still be other possible explanations for these results. In this section, we address limitations and alternative mechanisms, discussing them in the context of our findings.

One mechanism that is especially relevant to our results regarding weather shocks is time use. If short- or medium-run weather shocks affect time use, which in turn may affect giving behavior, then this could be a driving force behind our results. This is especially important in the case of online giving, since if weather leads people to spend more time indoors then this could expose them to more opportunities for online giving. Importantly, if time spent indoors is driving our results, then results should be similar for hot weather shocks and cold weather shocks since it has been shown that both types of shock can lead to more time spent indoors (Graff Zivin and Neidell, 2014). However, in both short-run and medium-run analyses, we find that hot and cold events generally have different, sometimes opposite effects on campaign contributions. In addition, in the medium-run analysis we include contributions from all sources, and the patterns are generally similar even though the effects are often not statistically significant. Therefore, time spent indoors is not likely to account for our results.

In the case of natural disasters, an important alternative mechanism is that of factors that are correlated with the LCV score but unrelated to incumbents' stances on environmental issues. For example, if pro-environment candidates are also more willing to pass disaster relief packages for those affected, this may explain the increase in funds and support for these candidates following disasters. While this is a possibility, we argue that these factors are not likely to be driving our results for two reasons. First, the observed effects of unusually hot weather in the

short and medium-run are in the same direction as those of natural disasters, and there is no obvious policy position regarding hot weather other than a politician's stance on environmental issues. This is especially true in the case of short-run weather variations. Second, to alleviate this concern we study the effects of natural disasters using a within-party specification, since policy positions on several issues are determined along party lines. We find that estimates generally point in the same direction as in our main specification, despite party affiliation being perhaps the most visible source of information on politicians' environmental positions.

Finally, it is possible that there are other psychological explanations for our results. For example, Meier, Schmid, and Stutzer (2016) explore the link between rainy weather, risk aversion, and voting for status quo candidates. This link between short term weather and emotions could be a confounder to the extent that emotions affect individuals' incentives to make political campaign contributions. However, we observe similar patterns in the medium term, as well as stronger short-term effects in counties with more anti-environment incumbents, so our results cannot be entirely driven by the emotional consequences of short-run weather.

3.5 Conclusion

In this paper, we study the impacts of extreme weather events on campaign contributions and electoral outcomes in the United States. As these events are often considered signs of climate change, our analyses place particular emphasis on testing for differential constituent responses based on the incumbent politician's views on environmental issues. In a short-run analysis, we find that weekly temperature shocks lead to a higher number of online donations to Democratic candidates, especially in counties with a greater share of anti-environment incumbents. In a medium-run analysis, we find that when a natural disaster strikes, challengers are more likely to enter congressional races the more anti-environment the incumbent is. We also find suggestive evidence that natural disasters lead to increased overall funding in congressional races where the

incumbent is more anti-environment.

Throughout the paper, we have argued that our results are driven by the environmental preference mechanism. We have suggested that this mechanism is a combination of two margins. First, extreme weather may lead people to update their beliefs in climate change, which in turn enables them to act on these beliefs politically. Second, extreme weather events may simply increase the salience of climate change, which in turn incentivizes those who have already made up their mind on these issues to take action.

Irrespective of the margin at play, the results in this paper suggest that politicians' policy positions on environmental issues are taken into account by people when responding to extreme weather events, and that these responses may have political consequences. Further, these findings suggest additional mechanisms for results in previous studies showing that congresspersons are more likely to vote in favor of environmental legislation following natural disasters in their state (Herrnstadt and Muehlegger, 2014). Put together, these behaviors from constituents, candidates, and legislators are consistent with a representative democracy at work.

Even though our results have important implications as they stand, both margins of the environmental preference mechanism have their own ramifications. If the belief-updating margin is at play, one concern is the extent to which such updating is rational; as is often pointed out in the literature, people tend to process shocks with psychological and ideological biases (Tversky and Kahneman, 1974). If this is the case, our findings caution that such biases might be expressed through the political system with unclear welfare consequences. On the other hand, if the salience margin is at play, one might instead be concerned about the electorate having limited attention. If this is the case, election outcomes might not represent well-defined collective preferences, since the salience of different policy issues may vary due to stochastic events. Given these ramifications, it is important to distinguish which margin of the environmental preference mechanism is predominantly at play. Therefore, we believe this is an important question for future work.

In addition to distinguishing between the two margins of the environmental preference mechanism, the results presented in this paper pose a series of additional questions and possible extensions. Firstly, a question raised by this work is whether the politicians themselves react to the salience of climate change by adjusting their narratives when it comes to speeches and soliciting contributions. Secondly, an important player which is missing from our analysis are environmental advocacy groups. Future research should focus on the role these groups play in disseminating information and forming opinions following extreme weather events. Finally, it is an open question whether the behavior observed here generalizes to other policy areas in which the event of interest has a stochastic component, like terrorism or gun violence.

Acknowledgement

Chapter 3, in full, is currently being prepared for submission for publication of the material. Yanjun Liao; Pablo Ruiz Junco. The dissertation author was a primary author of this chapter.

Bibliography

[1] Karen Akerlof, Edward W Maibach, Dennis Fitzgerald, Andrew Y Cedeno, and Amanda Neuman. "Do people personally experience global warming, and if so how, and does it matter?" In: *Global Environmental Change* 23.1 (2013), pp. 81–91.

[2] David Albouy and Bryan Stuart. *Urban population and amenities: The neoclassical model of location*. Tech. rep. National Bureau of Economic Research, 2014.

[3] Hunt Allcott. "Consumers' perceptions and misperceptions of energy costs". In: *American Economic Review* 101.3 (2011), pp. 98–104.

[4] Hunt Allcott. "Paternalism and energy efficiency: an overview". In: *Annual Review of Economics* 8 (2016), pp. 145–176.

[5] Hunt Allcott. "The welfare effects of misperceived product costs: Data and calibrations from the automobile market". In: *American Economic Journal: Economic Policy* 5.3 (2013), pp. 30–66.

[6] Hunt Allcott and Sendhil Mullainathan. "Behavior and energy policy". In: *Science* 327.5970 (2010), pp. 1204–1205.

[7] Hunt Allcott, Sendhil Mullainathan, and Dmitry Taubinsky. "Energy policy with externalities and internalities". In: *Journal of Public Economics* 112 (2014), pp. 72–88.

[8] Stephen Ansolabehere, John M. de Figueiredo, and Jr Snyder James M. "Why is There so Little Money in U.S. Politics?" In: *Journal of Economic Perspectives* 17.1 (2003), pp. 105–130.

[9] S Aneeqa Aqeel. "House Prices and Vacancies after Hurricane Katrina: Empirical Analysis of the Search & Matching Model". 2011.

[10] Kevin Arceneaux and Robert M Stein. "Who is held responsible when disaster strikes? The attribution of responsibility for a natural disaster in an urban election". In: *Journal of Urban Affairs* 28.1 (2006), pp. 43–53.

[11] Shahzeen Z Attari, Michael L DeKay, Cliff I Davidson, and Wändi Bruine De Bruin. "Public perceptions of energy consumption and savings". In: *Proceedings of the National Academy of Sciences* 107.37 (2010), pp. 16054–16059.

[12] Maximilian Auffhammer, Solomon M Hsiang, Wolfram Schlenker, and Adam Sobel. "Using weather data and climate model output in economic analyses of climate change". In: *Review of Environmental Economics and Policy* 7.2 (2013), pp. 181–198.

[13] Manuel Bagues and Berta Esteve-Volart. "Politicians Luck of the Draw: Evidence from the Spanish Christmas Lottery". In: *Journal of Political Economy* 124.5 (2016), pp. 1269–1294.

[14] Erin Baker, Meredith Fowlie, Derek Lemoine, and Stanley S Reynolds. "The Economics of Solar Electricity". In: *Annual Review of Resource Economics* 5.1 (2013), pp. 387–426.

[15] H Spencer Banzhaf and Randall P Walsh. "Do people vote with their feet? An empirical test of Tiebout". In: *American Economic Review* 98.3 (2008), pp. 843–63.

[16] Michael Barber. "Donation Motivations: Testing Theories of Access and Ideology". In: *Political Research Quarterly* 69.1 (2016), pp. 148–159.

[17] Patrick Bayer, Fernando Ferreira, and Robert McMillan. "A unified framework for measuring preferences for schools and neighborhoods". In: *Journal of Political Economy* 115.4 (2007), pp. 588–638.

[18] Patrick Bayer, Robert McMillan, Alvin Murphy, and Christopher Timmins. "A dynamic model of demand for houses and neighborhoods". In: *Econometrica* 84.3 (2016), pp. 893–942.

[19] Ariel R Belasen and Solomon W Polachek. "How disasters affect local labor markets the effects of hurricanes in Florida". In: *Journal of Human Resources* 44.1 (2009), pp. 251–276.

[20] Okmyung Bin, Jamie Brown Kruse, and Craig E Landry. "Flood hazards, insurance rates, and amenities: Evidence from the coastal housing market". In: *Journal of Risk and Insurance* 75.1 (2008), pp. 63–82.

[21] Okmyung Bin and Craig E Landry. "Changes in implicit flood risk premiums: Empirical evidence from the housing market". In: *Journal of Environmental Economics and Management* 65.3 (2013), pp. 361–376.

[22] Okmyung Bin and Stephen Polasky. "Effects of flood hazards on property values: evidence before and after Hurricane Floyd". In: *Land Economics* 80.4 (2004), pp. 490–500.

[23] Zachary Bleemer and Wilbert Van der Klaauw. “Disaster (over-) insurance: the long-term financial and socioeconomic consequences of Hurricane Katrina”. 2017.

[24] Bryan Bollinger and Kenneth Gillingham. “Peer effects in the diffusion of solar photovoltaic panels”. In: *Marketing Science* 31.6 (2012), pp. 900–912.

[25] Bryan K Bollinger and Kenneth Gillingham. “Learning-by-doing in solar photovoltaic installations”. 2014.

[26] Adam Bonica. “Database on Ideology, Money in Politics, and Elections: Public version 2.0 [Computer file].” In: *Stanford, CA: Stanford University Libraries.* (2016). <https://data.stanford.edu/dime>.

[27] Adam Bonica. “Ideology and Interests in the Political Marketplace”. In: *American Journal of Political Science* 57.2 (2013), pp. 294–311.

[28] Adam Bonica. “Mapping the Ideological Marketplace”. In: *American Journal of Political Science* 58.2 (2014), pp. 367–386.

[29] Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. “Salience and consumer choice”. In: *Journal of Political Economy* 121.5 (2013), pp. 803–843.

[30] Severin Borenstein. “The redistributional impact of nonlinear electricity pricing”. In: *American Economic Journal: Economic Policy* 4.3 (2012), pp. 56–90.

[31] Severin Borenstein and Lucas W Davis. “The distributional effects of US clean energy tax credits”. In: *Tax Policy and the Economy, Volume 30*. University of Chicago Press, 2015.

[32] Leah Platt Boustan, Matthew E Kahn, Paul W Rhode, and Maria Lucia Yanguas. *The effect of natural disasters on economic activity in us counties: A century of data*. Tech. rep. National Bureau of Economic Research, 2017.

[33] Meghan R Busse, Christopher R Knittel, and Florian Zettelmeyer. “Are consumers myopic? Evidence from new and used car purchases”. In: *American Economic Review* 103.1 (2013), pp. 220–56.

[34] Meghan R Busse, Devin G Pope, Jaren C Pope, and Jorge Silva-Risso. “The psychological effect of weather on car purchases”. In: *The Quarterly Journal of Economics* 130.1 (2015), pp. 371–414.

[35] California ISO. *2012 Summer Loads and Resource Assessment*. Tech. rep. 2012.

[36] Roger E Cannaday, Henry J Munneke, and Tyler T Yang. “A multivariate repeat-sales model for estimating house price indices”. In: *Journal of Urban Economics* 57.2 (2005), pp. 320–342.

[37] David Card, Alexandre Mas, and Jesse Rothstein. “Tipping and the Dynamics of Segregation”. In: *The Quarterly Journal of Economics* 123.1 (2008), pp. 177–218.

[38] Tom Y Chang, Wei Huang, and Yongxiang Wang. “Something in the air: Pollution and the demand for health insurance”. In: *The Review of Economic Studies* 85.3 (2018), pp. 1609–1634.

[39] Kerwin Kofi Charles, Erik Hurst, and Matthew J Notowidigdo. “Housing booms and busts, labor market opportunities, and college attendance”. In: *American Economic Review* 108.10 (2018), pp. 2947–94.

[40] Raj Chetty, Nathaniel Hendren, and Lawrence F Katz. “The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment”. In: *American Economic Review* 106.4 (2016), pp. 855–902.

[41] Michael Conlin, Ted O’Donoghue, and Timothy J Vogelsang. “Projection bias in catalog orders”. In: *The American Economic Review* (2007), pp. 1217–1249.

[42] Christine Lasco Crago and Ilya Chernyakhovskiy. “Are policy incentives for solar power effective? Evidence from residential installations in the Northeast”. In: *Journal of Environmental Economics and Management* 81 (2017), pp. 132–151.

[43] Jesús Crespo Cuaresma, Jaroslava Hlouskova, and Michael Obersteiner. “Natural disasters as creative destruction? Evidence from developing countries”. In: *Economic Inquiry* 46.2 (2008), pp. 214–226.

[44] Russell Davidson and James G MacKinnon. *Estimation and inference in econometrics*. New York: Oxford University Press, 1993.

[45] Donald R Davis and David E Weinstein. “Bones, bombs, and break points: the geography of economic activity”. In: *American Economic Review* 92.5 (2002), pp. 1269–1289.

[46] Frank L Davis and Albert H Wurth. “Voting preferences and the environment in the American electorate: The discussion extended”. In: *Society & Natural Resources* 16.8 (2003), pp. 729–740.

[47] Melissa Dell, Benjamin F Jones, and Benjamin A Olken. “What do we learn from the weather? The new climate-economy literature”. In: *Journal of Economic Literature* 52.3 (2014), pp. 740–98.

[48] Stefano DellaVigna. “Psychology and economics: Evidence from the field”. In: *Journal of Economic Literature* 47.2 (2009), pp. 315–372.

[49] Julie L Demuth, Mark DeMaria, and John A Knaff. "Improvement of Advanced Microwave Sounding Unit tropical cyclone intensity and size estimation algorithms". In: *Journal of Applied Meteorology and Climatology* 45.11 (2006), pp. 1573–1581.

[50] Tatyana Deryugina. "How do people update? The effects of local weather fluctuations on beliefs about global warming". In: *Climatic Change* 118.2 (2013), pp. 397–416.

[51] Tatyana Deryugina. "The fiscal cost of hurricanes: disaster aid versus social insurance". In: *American Economic Journal: Economic Policy* 9.3 (2017), pp. 168–98.

[52] Tatyana Deryugina, Laura Kawano, and Steven Levitt. "The economic impact of Hurricane Katrina on its victims: evidence from individual tax returns". In: *American Economic Journal: Applied Economics* 10.2 (2018), pp. 202–33.

[53] Olivier Deschenes and Michael Greenstone. "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather". In: *American Economic Review* 97.1 (2007), pp. 354–385.

[54] Rebecca Diamond. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000". In: *American Economic Review* 106.3 (2016), pp. 479–524.

[55] Rebecca Diamond and Timothy McQuade. "Who wants affordable housing in their backyard? An equilibrium analysis of low income property development". In: *Journal of Political Economy* (forthcoming).

[56] Patrick J Egan and Megan Mullin. "Turning personal experience into political attitudes: The effect of local weather on Americans perceptions about global warming". In: *The Journal of Politics* 74.3 (2012), pp. 796–809.

[57] EIA. *Sources of Heating for California Homes*. 2016.

[58] Michael J. Ensley. "Individual Campaign Contributions and Candidate Ideology". In: *Public Choice* 138.1/2 (2009), pp. 221–238. ISSN: 00485829, 15737101.

[59] James J Feigenbaum, James Lee, and Filippo Mezzanotti. "Capital Destruction and Economic Growth: The Effects of Shermans March, 1850-1920". 2017.

[60] Peter L. Francia, John C. Green, Paul S. Herrnson, Lynda W. Powell, and Clyde Wilcox. *The Financiers of Congressional Elections: Investors, Ideologues, and Intimates*. Columbia University Press, 2003, pp. 42–68.

[61] Elizabeth Fussell. "The long-term recovery of New Orleans population after Hurricane Katrina". In: *American Behavioral Scientist* 59.10 (2015), pp. 1231–1245.

[62] Etienne Gagnon and David Lopez-Salido. “Small price responses to large demand shocks”. In: *Journal of the European Economic Association* (2015).

[63] Justin Gallagher. “Learning about an infrequent event: evidence from flood insurance take-up in the United States”. In: *American Economic Journal: Applied Economics* 6.3 (2014), pp. 206–33.

[64] Justin Gallagher and Daniel Hartley. “Household finance after a natural disaster: The case of Hurricane Katrina”. In: *American Economic Journal: Economic Policy* 9.3 (2017), pp. 199–228.

[65] John T Gasper and Andrew Reeves. “Make it rain? Retrospection and the attentive electorate in the context of natural disasters”. In: *American Journal of Political Science* 55.2 (2011), pp. 340–355.

[66] Matthew Gibson, Jamie T Mullins, and Alison Hill. “Climate risk and beliefs: Evidence from New York floodplains”. 2019.

[67] Kenneth Gillingham, Amelia Keyes, and Karen Palmer. “Advances in Evaluating Energy Efficiency Policies and Programs”. In: *Annual Review of Resource Economics* 10 (2018), pp. 511–532.

[68] Edward L Glaeser and Joseph Gyourko. “Urban decline and durable housing”. In: *Journal of Political Economy* 113.2 (2005), pp. 345–375.

[69] Brad T Gomez, Thomas G Hansford, and George A Krause. “The Republicans should pray for rain: Weather, turnout, and voting in US presidential elections”. In: *Journal of Politics* 69.3 (2007), pp. 649–663.

[70] Joshua Graff Zivin and Matthew Neidell. “Temperature and the allocation of time: Implications for climate change”. In: *Journal of Labor Economics* 32.1 (2014), pp. 1–26.

[71] William Greene. “The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects”. In: *The Econometrics Journal* 7.1 (2004), pp. 98–119.

[72] Daniel Greenwald. “The mortgage credit channel of macroeconomic transmission”. 2018.

[73] Jeffrey A Groen and Anne E Polivka. “The effect of Hurricane Katrina on the labor market outcomes of evacuees”. In: *American Economic Review* 98.2 (2008), pp. 43–48.

[74] Deborah Lynn Guber. “Voting preferences and the environment in the American electorate”. In: *Society & Natural Resources* 14.6 (2001), pp. 455–469.

[75] Veronica Guerrieri, Daniel Hartley, and Erik Hurst. "Endogenous gentrification and housing price dynamics". In: *Journal of Public Economics* 100 (2013), pp. 45–60.

[76] Joseph Gyourko and Albert Saiz. "Reinvestment in the housing stock: the role of construction costs and the supply side". In: *Journal of Urban Economics* 55.2 (2004), pp. 238–256.

[77] Daniel G Hallstrom and V Kerry Smith. "Market responses to hurricanes". In: *Journal of Environmental Economics and Management* 50.3 (2005), pp. 541–561.

[78] Lawrence C Hamilton and Mary D Stampone. "Blowin in the wind: Short-term weather and belief in anthropogenic climate change". In: *Weather, Climate, and Society* 5.2 (2013), pp. 112–119.

[79] John P Harding, Stuart S Rosenthal, and CF Sirmans. "Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model". In: *Journal of Urban Economics* 61.2 (2007), pp. 193–217.

[80] Matthew Harding and David Rapson. "Do voluntary carbon offsets induce energy rebound? A conservationists dilemma". 2013.

[81] Justine S Hastings and Jesse M Shapiro. "Fungibility and consumer choice: Evidence from commodity price shocks". In: *The Quarterly Journal of Economics* 128.1 (2013), pp. 1449–1498.

[82] Jerry A Hausman. "Individual discount rates and the purchase and utilization of energy-using durables". In: *The Bell Journal of Economics* (1979), pp. 33–54.

[83] Andrew Healy and Neil Malhotra. "Myopic voters and natural disaster policy". In: *American Political Science Review* 103.3 (2009), pp. 387–406.

[84] Evan Herrnstadt and Erich Muehlegger. "Weather, salience of climate change and congressional voting". In: *Journal of Environmental Economics and Management* 68.3 (2014), pp. 435–448.

[85] Daniel Hoople. "The Budgetary Impact of the Federal Governments Response to Disasters". In: *Congressional Budget Office* (2013).

[86] Peter D Howe, Matto Mildenberger, Jennifer R Marlon, and Anthony Leiserowitz. "Geographic variation in opinions on climate change at state and local scales in the USA". In: *Nature Climate Change* 5.6 (2015), p. 596.

[87] Solomon M Hsiang and Amir S Jina. *The causal effect of environmental catastrophe on long-run economic growth: Evidence from 6,700 cyclones*. Tech. rep. National Bureau of Economic Research, 2014.

[88] Jonathan E Hughes and Molly Podolefsky. “Getting Green with Solar Subsidies: Evidence from the California Solar Initiative”. In: *Journal of the Association of Environmental and Resource Economists* 2.2 (2015), pp. 235–275.

[89] Fred Charles Ikle. “The effect of war destruction upon the ecology of cities”. In: *Social Forces* 29.4 (1951), pp. 383–391.

[90] IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2013, p. 1535. ISBN: ISBN 978-1-107-66182-0.

[91] IPCC. “Summary for Policymakers”. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2013. Chap. SPM, 130. ISBN: ISBN 978-1-107-66182-0.

[92] Koichiro Ito. “Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing”. In: *The American Economic Review* 104.2 (2014), pp. 537–563.

[93] Jon Jachimowicz, Jochen Menges, and Adam D. Galinsky. “How Wind Speed Affects Voting Decisions (May 31, 2017)”. Columbia Business School Research Paper No. 16-82. Available at SSRN: https://ssrn.com/abstract=2868054. 2017.

[94] Jeff Joireman, Heather Barnes Truelove, and Blythe Duell. “Effect of outdoor temperature, heat primes and anchoring on belief in global warming”. In: *Journal of Environmental Psychology* 30.4 (2010), pp. 358–367.

[95] David Karpf. *The Internet and American political campaigns*. In *The Forum* (Vol. 11, No. 3, pp. 413–428). De Gryuter. 2013.

[96] Jan Kleissl. *Improving Economics of Solar Power through Resource Analysis, Forecasting, and Dynamic System Modeling*. Tech. rep. 2013.

[97] Jeffrey R Kling, Jeffrey B Liebman, and Lawrence F Katz. “Experimental analysis of neighborhood effects”. In: *Econometrica* 75.1 (2007), pp. 83–119.

[98] Jeffrey R Kling, Jens Ludwig, and Lawrence F Katz. “Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment”. In: *The Quarterly Journal of Economics* 120.1 (2005), pp. 87–130.

[99] James P Kossin. “A global slowdown of tropical-cyclone translation speed”. In: *Nature* 558.7708 (2018), p. 104.

[100] Botond Kőszegi and Adam Szeidl. “A model of focusing in economic choice”. In: *The Quarterly Journal of Economics* 128.1 (2012), pp. 53–104.

[101] Carolyn Kousky, Erwann O Michel-Kerjan, and Paul A Raschky. “Does federal disaster assistance crowd out flood insurance?” In: *Journal of Environmental Economics and Management* 87 (2018), pp. 150–164.

[102] Nicolai V Kuminoff, V Kerry Smith, and Christopher Timmins. “The new economics of equilibrium sorting and policy evaluation using housing markets”. In: *Journal of Economic Literature* 51.4 (2013), pp. 1007–62.

[103] Gregory D Kutz and John J Ryan. *Hurricanes Katrina and Rita Disaster Relief: Improper and Potentially Fraudulent Individual Assistance Payments Estimated to be Between 600 Million and 1.4 Billion: Testimony Before the Subcommittee on Investigations, Committee on Homeland Security, House of Representatives*. US Government Accountability Office, 2006.

[104] Stefan Lamp. *Sunspots that matter: the effect of weather on solar technology adoption*. Tech. rep. TSE Working Paper, 2018.

[105] Corey Lang and John David Ryder. “The effect of tropical cyclones on climate change engagement”. In: *Climatic change* 135.3-4 (2016), pp. 625–638.

[106] League of Conservation Voters. *National Environmental Scorecard, First Session of the 115th Congress*. <http://http://scorecard.lcv.org/sites/scorecard.lcv.org/files/LCV_Scorecard-2017-Full.pdf>. 2007.

[107] Ye Li, Eric J Johnson, and Lisa Zaval. “Local warming: daily variation in temperature affects beliefs and concern about climate change”. In: *Psychol Sci* 22 (2011), pp. 454–459.

[108] George Loewenstein, Ted O’Donoghue, and Matthew Rabin. “Projection Bias in Predicting Future Utility”. In: *The Quarterly Journal of Economics* (2003), pp. 1209–1248.

[109] Tiziana de Magistris and Azucena Gracia. “Assessing Projection Bias in Consumers Food Preferences”. In: *PloS one* 11.2 (2016), e0146308.

[110] Anup Malani and Julian Reif. “Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform”. In: *Journal of Public Economics* 124 (2015), pp. 1–17.

[111] Shawn J McCoy and Xiaoxi Zhao. “A City under Water: A Geospatial Analysis of Storm Damage, Changing Risk Perceptions, and Investment in Residential Housing”. In: *Journal of the Association of Environmental and Resource Economists* 5.2 (2018), pp. 301–330.

[112] Molly Fifer McIntosh. “Measuring the labor market impacts of Hurricane Katrina migration: evidence from Houston, Texas”. In: *American Economic Review* 98.2 (2008), pp. 54–57.

[113] Armando Meier, Lukas Schmid, and Alois Stutzer. “Rain, Emotions and Voting for the Status Quo”. Available at SSRN: https://ssrn.com/abstract=2868316. 2016.

[114] Erwann O Michel-Kerjan. “Catastrophe economics: the national flood insurance program”. In: *Journal of Economic Perspectives* 24.4 (2010), pp. 165–86.

[115] Edward Miguel and Gerard Roland. “The long-run impact of bombing Vietnam”. In: *Journal of Development Economics* 96.1 (2011), pp. 1–15.

[116] Missouri Census Data Center. *MABLE/Geocorr 90/2k/14: Geographic Correspondence Engine.* <http://mcdc.missouri.edu/websas/geocorr_index.shtml>. 2017.

[117] Frances C Moore, Nick Obradovich, Flavio Lehner, and Patrick Baylis. “Rapidly declining remarkability of temperature anomalies may obscure public perception of climate change”. In: *Proceedings of the National Academy of Sciences* 116.11 (2019), pp. 4905–4910.

[118] Anthony Murphy and Eric Strobl. “The impact of hurricanes on housing prices: evidence from US coastal cities”. 2010.

[119] Teresa A Myers, Edward W Maibach, Connie Roser-Renouf, Karen Akerlof, and Anthony A Leiserowitz. “The relationship between personal experience and belief in the reality of global warming”. In: *Nature Climate Change* 3.4 (2013), p. 343.

[120] Christina Paxson and Cecilia Elena Rouse. “Returning to new orleans after hurricane katrina”. In: *American Economic Review* 98.2 (2008), pp. 38–42.

[121] Martino Pelli and Jeanne Tschopp. “Comparative advantage, capital destruction, and hurricanes”. In: *Journal of International Economics* 108 (2017), pp. 315–337.

[122] Kendra Pierre-Louis. *Its Cold Outside. Cue the Trump Global Warming Tweet.* The New York Times. Available at <www.nytimes.com/2017/12/28/climate/trump-tweet-global-warming.html>. 2017.

[123] Jacquelyn Pless and Arthur A van Benthem. “Pass-Through as a Test for Market Power: An Application to Solar Subsidies”. In: *American Economic Journal: Applied Economics* (forthcoming).

[124] Sherwin Rosen. “Hedonic prices and implicit markets: product differentiation in pure competition”. In: *Journal of Political Economy* 82.1 (1974), pp. 34–55.

[125] Albert Saiz. “The geographic determinants of housing supply”. In: *The Quarterly Journal of Economics* 125.3 (2010), pp. 1253–1296.

[126] James M Sallee. “Rational inattention and energy efficiency”. In: *The Journal of Law and Economics* 57.3 (2014), pp. 781–820.

[127] Narayan Sastry and Jesse Gregory. “The location of displaced New Orleans residents in the year after Hurricane Katrina”. In: *Demography* 51.3 (2014), pp. 753–775.

[128] Ishuwar Seetharam. “The Indirect Effects of Hurricanes: Evidence from Firm Internal Networks”. 2018.

[129] Uri Simonsohn. “Weather to go to college”. In: *The Economic Journal* 120.543 (2010), pp. 270–280.

[130] Matthew R Sisco, Valentina Bosetti, and Elke U Weber. “When do extreme weather events generate attention to climate change?” In: *Climatic Change* 143.1-2 (2017), pp. 227–241.

[131] Mark Skidmore and Hideki Toya. “Do natural disasters promote long-run growth?” In: *Economic Inquiry* 40.4 (2002), pp. 664–687.

[132] James M. Snyder. “Campaign Contributions as Investments: The U.S. House of Representatives, 1980-1986”. In: *Journal of Political Economy* 98.6 (1990), pp. 1195–1227. ISSN: 00223808, 1537534X.

[133] Alexa Spence, Wouter Poortinga, Catherine Butler, and Nicholas Frank Pidgeon. “Perceptions of climate change and willingness to save energy related to flood experience”. In: *Nature Climate Change* 1.1 (2011), p. 46.

[134] Paul W Stackhouse Jr, David Westberg, James M Hoell, William S Chandler, and Taiping Zhang. *Prediction Of Worldwide Energy Resource (POWER) Agroclimatology Methodology*. Tech. rep. NASA Langley Research Center and SSAI/NASA Langley Research Center, 2015.

[135] Allison Stevens. *Despite Terrorism, Candidates Make Slow Return to Fundraising*. The Hill. Available at <www.hillnews.com/102401>. 2001.

[136] Eric Strobl. “The economic growth impact of hurricanes: evidence from US coastal counties”. In: *Review of Economics and Statistics* 93.2 (2011), pp. 575–589.

[137] Amos Tversky and Daniel Kahneman. “Judgment under uncertainty: Heuristics and biases”. In: *science* 185.4157 (1974), pp. 1124–1131.

[138] Clyde Wilcox. *Internet fundraising in 2008: A new model?* In *The Forum* (Vol. 6, No. 1). De Gryuter. 2008.

[139] Lisa Zaval, Elizabeth A Keenan, Eric J Johnson, and Elke U Weber. “How warm days increase belief in global warming”. In: *Nature Climate Change* 4.2 (2014), p. 143.

Appendix A

Chapter 1 Appendix

A.1 Additional Figures

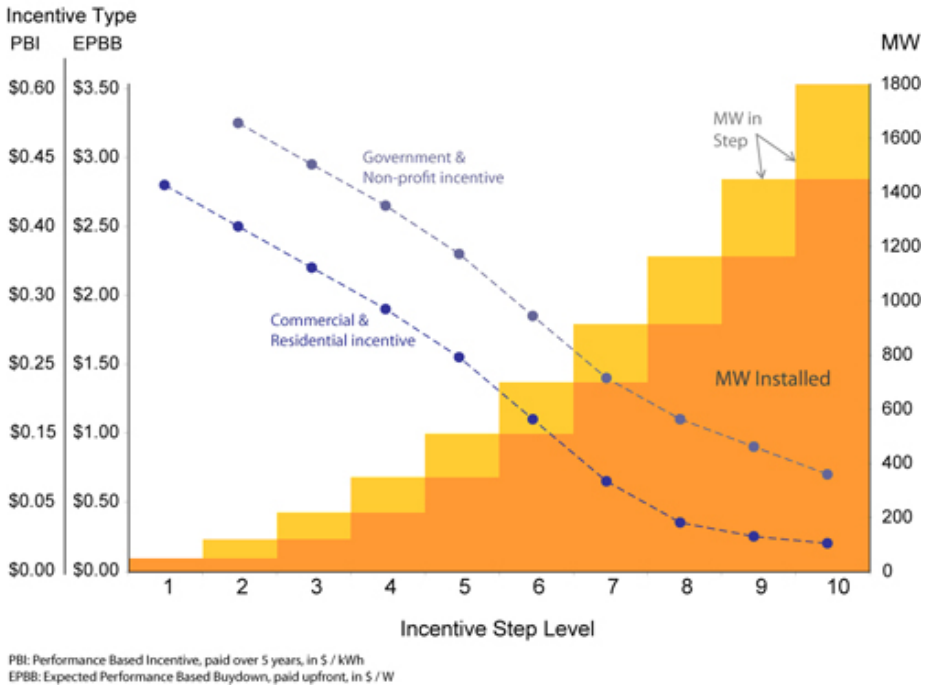


Figure A.1: CSI Incentive Step Design

Source: http://www.gosolarcalifornia.ca.gov/csi/rebates.php

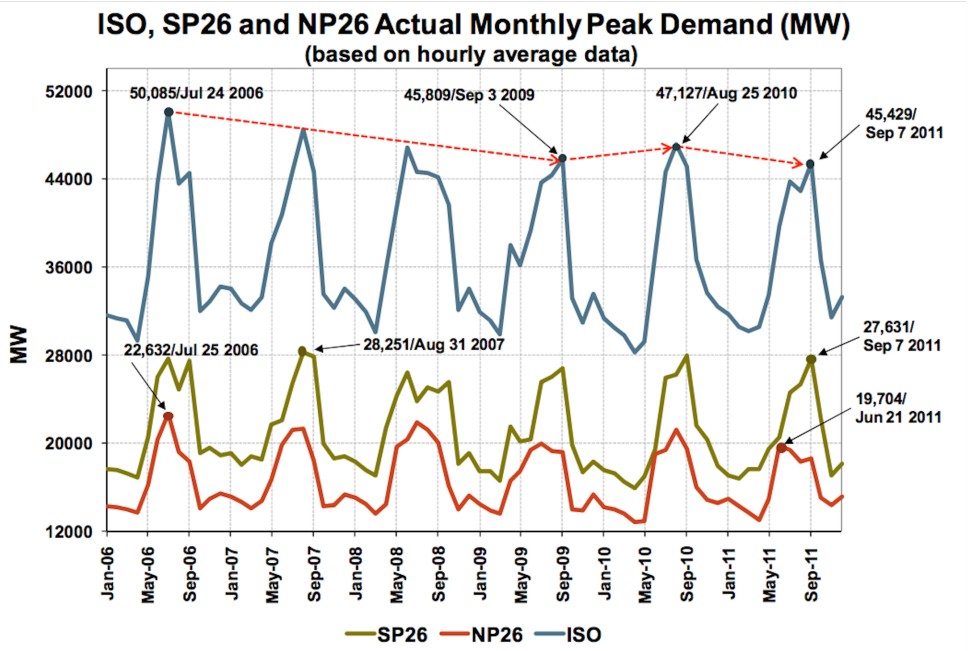


Figure A.2: Monthly Peak Demand in California, 2006-2011

Source: California ISO (2012)
Notes: This graph shows the actual monthly peak demand for southern California (SP26), northern California (NP26), and the entire California ISO system over the years 2006-2011.

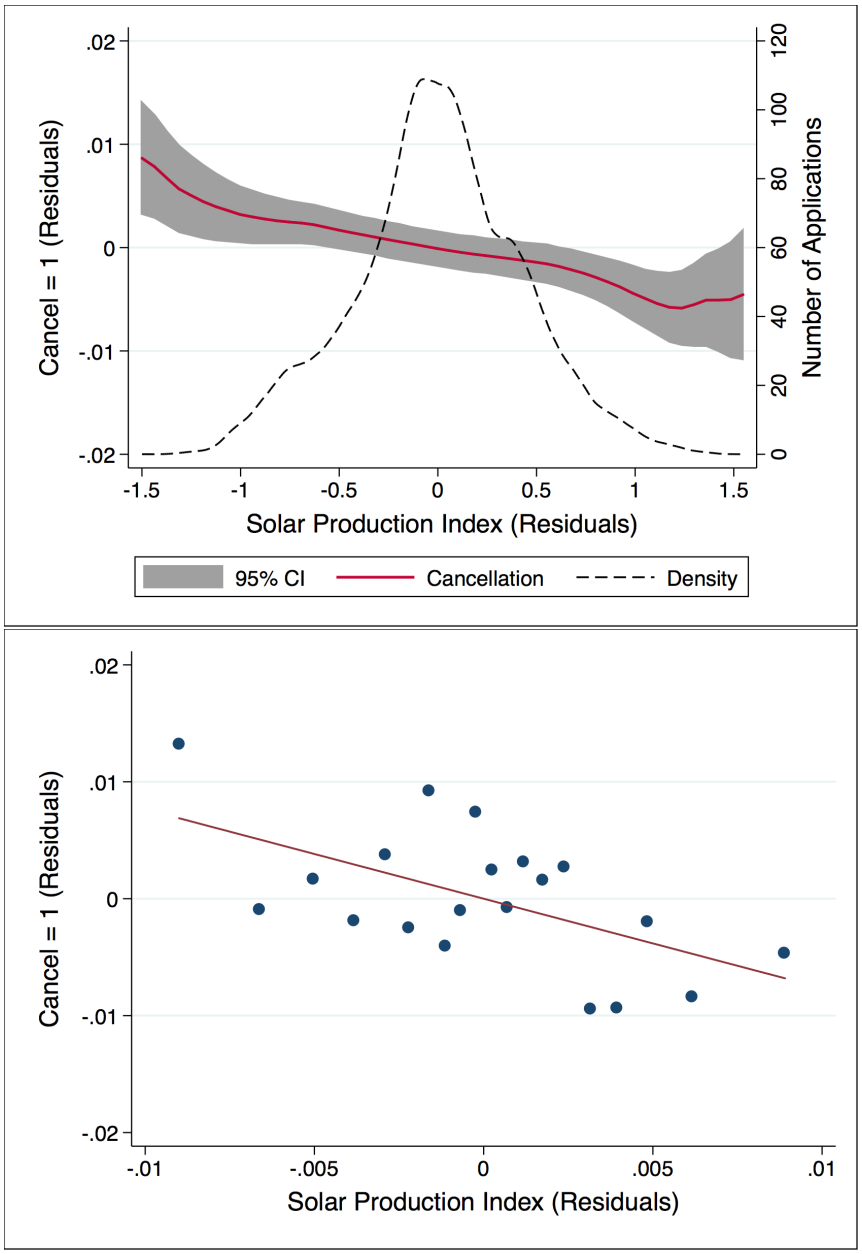


Figure A.3: Support for Linear Specification

Notes: These graphs verify the linear relationship between the probability of cancellation and solar production index using two approaches. The upper panel fits the relationship with a flexible local polynomial, and the lower one uses a binned scatter plot by dividing the observations into twenty bins.

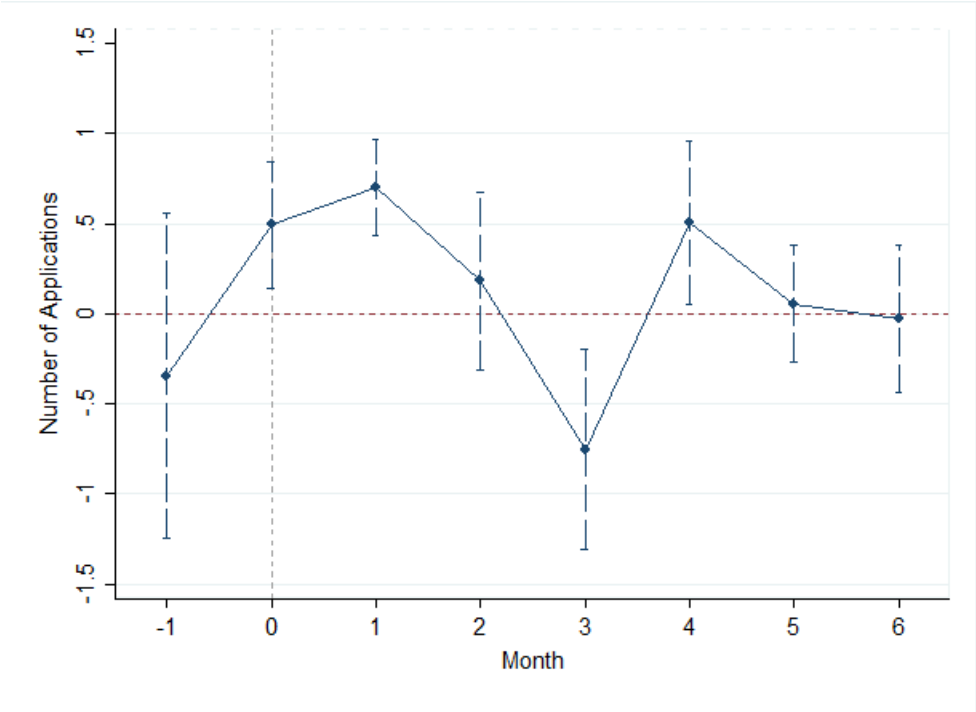


Figure A.4: Harvesting Effects in Sign-ups

Notes: This graph plots the contemporaneous and lagged effects of the production index on sign-ups. The dependent variable is the number of applications by zip code by month, and the main regressors are the average production index in the same month, its lead and six lags. The controls include monthly economic conditions as well as zip code, month-in-year, and year fixed effects. Point estimates and 95 percent confidence intervals are shown. Standard errors are clustered by county.

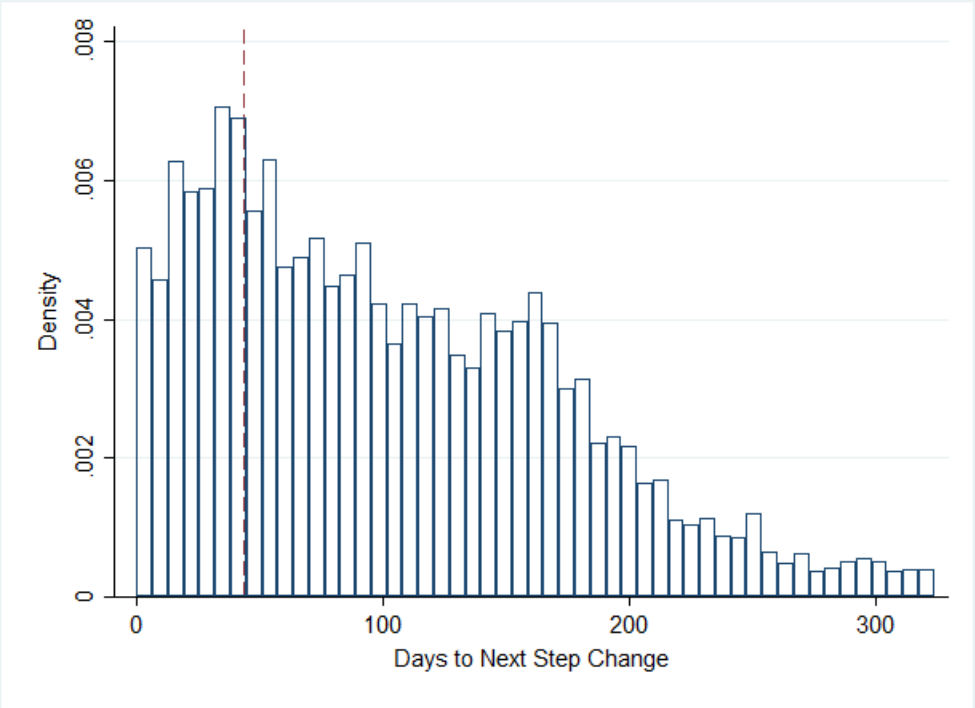


Figure A.5: Density of Days to Next Incentive Step Change

Notes: This graph plots the distribution of days to next incentive step change. The vertical dashed red line represents 47 days, the cutoff used in the heterogeneity analysis.

A.2 Additional Tables

Table A.1: Variable Descriptions, Continued

Variable	Description	Source
A. Application Characteristics		
CSI rating	AC output of photovoltaic module	CSI
total cost ($)	Total cost of eligible portion of the project	CSI
incentive amount ($)	CSI incentive amount for the system	CSI
net unit cost ($)	Net out-of-pocket cost per kW of capacity	CSI
third-party ownership (TPO)	Indicates whether the system owner is a different entity than the host customer	CSI
year	Application begin year	CSI
B. Weather		
Note: The term "post-period" below refers to the 113 days following application submission.		
solar insolation	Solar insolation, post-period average	POWER/NASA
max temp (F)	Maximum temperature, post-period average	GSOD/NOAA
#days(max $\leq$ 40)	Number of days with max temp below 40F in the post-period	GSOD/NOAA
#days(max $>$ 100)	Number of days with max temp above 100F in the post-period	GSOD/NOAA
wind speed (knots)	Wind speed, post-period average	GSOD/NOAA
#(precipitation$>$ 0)	Number of days with positive precipitation in the post-period	GSOD/NOAA
production index	Index of solar panel productivity, calculated using solar insolation, max temp, and wind speed, post-period average	Calculated
cooling degree days	Total cooling degree days in the post-period	Calculated
heating degree days	Total heating degree days in the post-period	Calculated
total degree days	Sum of cooling and heating degree days in the post-period	Calculated
C. Economic Conditions		
leading index	The leading index for California, predictor of six-month growth	FRED
unemployment	The unemployment rate in California	FRED
prime interest rate	The U.S. (Fed) prime rate	
index of consumer sentiment	Index of consumer sentiment in the West	Survey of Consumers
index of buying condition	Index of buying conditions in the West	Survey of Consumers
D. Zip Code Demographics and Interconnection Characteristics		
white (%)	Percent of white population in the zip code	2011 ACS
bachelor degree (%)	Percent of population with bachelor's degree in the zip code	2011 ACS
median income ($)	Median income in the zip code	2011 ACS
monthly housing cost ($)	Average monthly housing cost in the zip code	2011 ACS
mean household size	Average household size	2010 Census
urban (%)	Percent of urban households	2010 Census
current installed base	Number of interconnected solar systems in	NEM Interconnection

Continued on next page

Table A.1 Variable Descriptions, Continued

Variable	Description	Source
	the zip code	
installed penetration	Number of interconnected solar systems per thousand households in the zip code	NEM Interconnection

Table A.2: Cancellation and Solar Insolation (113 Days After Sign-up)

Cancel = 1	(1)	(2)	(3)	(4)	(5)
solar insolation	−0.146	−0.203	−0.110	−0.121	−0.143
	[0.055]**	[0.045]***	[0.051]**	[0.057]**	[0.085]*
wind speed	0.262	0.799	−0.065	0.314	0.105
	[0.241]	[0.271]***	[0.204]	[0.256]	[0.243]
#days(tmax $<$ 40)	1.055	1.406	0.658	0.757	1.071
	[0.383]***	[0.520]***	[0.348]*	[0.304]**	[0.394]***
#days(tmax$\geq$100)	−0.020	−0.024	−0.019	−0.024	−0.006
	[0.023]	[0.021]	[0.024]	[0.022]	[0.025]
CSI Rating	0.590	0.592	0.589	0.578	0.588
	[0.091]***	[0.091]***	[0.087]***	[0.086]***	[0.091]***
unit cost	0.000	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
third-party ownership	0.833	0.830	0.962	0.962	0.840
	[0.520]	[0.514]	[0.510]*	[0.508]*	[0.520]
CA leading index	0.457	0.506	0.435	1.860	0.378
	[0.485]	[0.481]	[0.499]	[0.552]***	[0.515]
CA unemployment rate	−0.782	−0.735	−0.990	0.746	−0.711
	[0.460]*	[0.543]	[0.453]**	[0.833]	[0.529]
prime interest rate	−2.084	−2.155	−2.289	−2.103	−2.345
	[1.102]*	[1.061]**	[1.111]**	[1.199]*	[1.354]*
index of consumer sentiment	0.188	0.205	0.202	0.125	0.149
	[0.058]***	[0.053]***	[0.052]***	[0.070]*	[0.054]***
index of buying conditions	−0.134	−0.141	−0.143	−0.119	−0.092
	[0.040]***	[0.039]***	[0.038]***	[0.045]**	[0.040]**
D.V. Mean	12.06	12.06	12.06	12.06	12.06
R^2 (within)	0.033	0.035	0.042	0.051	0.034
N	154,519	154,518	154,513	154,465	154,519
Fixed Effects					
Zip Code	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes			Yes
Quarter	Yes		Yes		
Month					Yes
County-quarter		Yes			
County-year			Yes		
County-quarter-year				Yes	

Notes: This table reports estimates based on a time frame of 113 days following the application date. Columns (1)-(2) show full results of regressions (1)-(2) in Table 1.2. Columns (3)-(5) shows results with other fixed-effect specifications. Coefficients are multiplied by 100 for legibility. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.3: Cancellation and Solar Insolation (60 Days Before Completion/Cancellation)

Cancel = 1	(1)	(2)	(3)	(4)	(5)
solar insolation	−0.186	−0.192	−0.181	−0.185	−0.090
	[0.045]***	[0.045]***	[0.048]***	[0.051]***	[0.113]
wind speed	0.815	0.754	0.728	0.719	0.787
	[0.173]***	[0.162]***	[0.185]***	[0.208]***	[0.169]***
#days(tmax $<$ 40)	−1.278	−1.572	−1.482	−2.069	−1.206
	[0.489]**	[0.361]***	[0.417]***	[0.391]***	[0.494]**
#days(tmax$\geq$100)	−0.256	−0.330	−0.261	−0.336	−0.250
	[0.030]***	[0.065]***	[0.032]***	[0.069]***	[0.031]***
CSI Rating	0.590	0.592	0.589	0.578	0.588
	[0.091]***	[0.091]***	[0.087]***	[0.086]***	[0.091]***
unit cost	0.000	0.000	0.000	0.000	0.000
	[0.000]**	[0.000]**	[0.000]**	[0.000]***	[0.000]**
third-party ownership	0.766	0.767	0.691	0.638	0.766
	[0.362]**	[0.349]**	[0.366]*	[0.372]*	[0.364]**
CA leading index	−0.100	−0.173	−0.139	1.071	−0.321
	[0.269]	[0.269]	[0.274]	[0.516]**	[0.289]
CA unemployment rate	−0.780	−0.736	−0.811	−0.142	−0.436
	[0.374]**	[.378]*	[0.397]**	[0.930]	[0.473]
prime interest rate	−0.556	−0.324	−0.576	−0.315	0.026
	[0.924]	[0.983]	[0.939]	[1.136]	[1.118]
index of consumer sentiment	0.266	0.253	0.285	0.244	0.242
	[0.090]***	[0.097]**	[0.092]***	[0.118]**	[0.093]**
index of buying conditions	−0.086	−0.082	−0.099	−0.121	−0.064
	[0.037]**	[0.039]**	[0.037]**	[0.056]**	[0.040]
D.V. Mean	6.70	6.70	6.70	6.70	6.70
R^2 (within)	0.072	0.076	0.081	0.098	0.073
N	62,069	62,062	62,053	61,951	62,069
Fixed Effects					
Zip Code	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes			Yes
Quarter	Yes		Yes		
Month					Yes
County-quarter		Yes			
County-year			Yes		
County-quarter-year				Yes	

Notes: This table reports estimates based on a time frame of 60 days prior to the completion/cancellation date. Columns (1)-(2) show full results of regressions (3)-(4) in Table 1.2. Columns (3)-(5) shows results with other fixed-effect specifications. Coefficients are multiplied by 100 for legibility. Standard errors (in squared brackets) are clustered by county.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.4: Robustness Check: Varying Controls

Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
solar insolation	−0.129	−0.151	−0.219	−0.254	−0.115	−0.094
	[0.041]***	[0.039]***	[0.044]***	[0.053]***	[0.045]**	[0.035]***
wind speed			0.797	0.791		
			[0.259]***	[0.168]***		
#days(tmax $<$ 40)					1.460	−1.622
					[0.550]**	[0.366]***
#days(tmax $\geq$ 100)					−0.021	−0.333
					[0.020]	[0.066]***
Sample						
Time Frame	First 113	Last 60	First 113	Last 60	First 113	Last 60
Duration	All	0-100	All	0-100	All	0-100
D.V. Mean	12.06	6.70	12.06	6.70	12.06	6.70
R^2 (within)	0.035	0.073	0.035	0.074	0.035	0.076
N	154,518	62,062	154,518	62,062	154,518	62,062
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
County-quarter	Yes	Yes	Yes	Yes	Yes	Yes

Notes: All coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.5: Robustness Check: Zip-Code-Specific Time Trends

Cancel = 1	(1)	(2)	(3)	(4)
solar insolation	−0.104	−0.171	−0.119	−0.184
	[0.051]**	[0.045]***	[0.047]**	[0.045]***
wind speed	−0.061	0.745	0.178	0.746
	[0.215]	[0.189]***	[0.238]	[0.201]***
#days(tmax $<$ 40)	0.569	−1.634	0.672	−1.806
	[0.362]	[0.451]***	[0.384]*	[0.376]***
#days(tmax$\geq$100)	−0.022	−0.265	−0.023	−0.271
	[0.022]	[0.030]***	[0.023]	[0.033]***
Sample				
Time Frame	First 113	Last 60	First 113	Last 60
Duration	All	0-100	All	0-100
D.V. Mean	12.06	6.70	12.06	6.70
R^2 (within)	0.053	0.103	0.053	0.104
N	154,513	62,053	154,513	62,053
Time Trend				
Zip-year	Yes	Yes		
Zip-quarter-year			Yes	Yes
Fixed Effects				
Zip Code	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
County-quarter	Yes	Yes	Yes	Yes

Notes: This table shows the results from adding two types of zip-code-specific time trends to the two main specifications. All coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.6: Robustness Check: Logit Model

Cancel = 1	(1)	(2)	(3)	(4)
solar insolation	−1.071	−1.540	−1.206	−1.656
	(0.293)***	(0.425)***	(0.436)***	(0.785)**
wind speed	4.038	2.567	2.187	0.911
	(1.609)**	(1.746)	(1.795)	(1.810)
#days(tmax $<$ 40)	9.284	9.843	10.734	10.011
	(4.222)**	(4.221)**	(4.166)***	(4.204)**
#days(tmax$\geq$100)	−0.294	−0.190	−0.166	−0.050
	(.125)**	(0.131)	(0.132)	(0.140)
N	153,452	153,452	153,452	153,452
Fixed Effects				
Zip Code	Yes	Yes	Yes	Yes
Year	Yes		Yes	
Quarter	Yes		Yes	
Quarter-year		Yes		Yes
Month				Yes

Notes: This table reports estimates based on a time frame of 113 days after the sign-up. All coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are always controlled for. Marginal effect calculations are not feasible due to the fixed effects, and specifications with county-interacted fixed effects are omitted due to computation problems. Standard errors (in parentheses) are clustered by zip code.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.7: Robustness Check: Varying Period Length

Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
Period length	70 days	70 days	90 days	90 days	113 days	113 days
solar insolation	−0.147	−0.185	−0.157	−0.206	−0.146	−0.203
	[0.045]***	[0.043]***	[0.048]***	[0.043]***	[0.055]**	[0.045]***
wind speed	0.319	0.652	0.317	0.751	0.262	0.799
	[0.210]	[0.248]**	[0.220]	[0.249]***	[0.241]	[0.271]***
#days(tmax < 40)	0.358	0.511	0.379	0.569	1.055	1.406
	[0.351]	[0.400]	[0.352]	[0.421]	[0.383]***	[0.521]***
#days(tmax ≥ 100)	−0.042	−0.052	−0.030	−0.036	−0.020	−0.024
	[0.024]*	[0.022]**	[0.025]	[0.023]	[0.023]	[0.021]
D.V. Mean	12.06	12.06	12.06	12.06	12.06	12.06
R^2 (within)	0.033	0.035	0.033	0.035	0.033	0.035
N	154,519	154,518	154,519	154,518	154,519	154,518
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Notes: Columns (1)-(2), (3)-(4), and (5)-(6) report estimates based on a time frame of 70, 90, and 113 days following the application date. All coefficients are multiplied by 100 for better display. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.8: Robustness Check: Different Subsamples

Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
solar insolation	−0.185	−0.208	−0.113	−0.129	−0.104	−0.117
	[0.055]***	[0.064]***	[0.047]**	[0.054]**	[0.038]***	[0.044]**
wind speed	0.478	0.470	0.229	0.254	0.214	0.258
	[0.292]	[0.391]	[0.208]	[0.273]	[0.178]	[0.219]
#days(tamx<40)	0.513	0.895	0.360	0.499	0.434	0.484
	[0.226]**	[0.331]***	[0.176]**	[0.240]**	[0.173]**	[0.228]**
#days(tmax≥100)	0.019	0.012	0.016	0.014	0.016	0.019
	[0.022]	[0.013]	[0.014]	[0.010]	[0.009]*	[0.009]**
Sample						
Time Frame	Last 60	Last 60	Last 60	Last 60	Last 60	Last 60
Duration Restriction	0-100	0-100	0-150	0-150	0-200	0-200
D.V. Mean	6.70	6.70	4.97	4.97	4.49	4.49
R^2 (within)	0.070	0.073	0.047	0.049	0.039	0.041
N	62,069	62,062	93,061	93,055	111,654	111,650
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Notes: Columns (1)-(2), (3)-(4), and (5)-(6) restrict the sample to those with reported duration within 100, 150, and 200 days, respectively. All coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.9: Channel: Solar Production vs. Energy Demand (Alternative Time Frame)

Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
$ProdIndex_{post}$	−1.343	−1.486	−1.529	−1.639	−1.208	−1.277
	[0.300]***	[0.311]***	[0.300]***	[0.351]***	[0.359]***	[0.360]***
CDD			0.376	0.312		
			[0.261]	[0.248]		
$TDD(=CDD+HDD)$					0.287	0.506
					[0.243]	[0.251]**
D.V. Mean	6.70	6.70	6.70	6.70	6.70	6.70
R^2 (within)	0.071	0.074	0.071	0.074	0.071	0.074
N	62,069	62,062	62,069	62,062	62,069	62,062
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Notes: This table reports estimates based on a time frame of 60 days prior to the completion/cancellation date. All indices are normalized. All coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.10: Attenuation Bias with Different Period Lengths

Cancel = 1	(1) 10 days	(2) 30 days	(3) 50 days	(4) 70 days	(5) 90 days	(6) 113 days
ProdIndex	0.057	−0.249	−0.443	−0.542	−0.692	−0.761
	[0.141]	[0.141]*	[0.144]***	[0.154]***	[0.172]***	[0.191]***
Test: Equality of Coefficients with Column (6)						
$\chi^2 : \beta_{(i)} = \beta_{(6)}$	25.38***	13.70***	6.90***	5.00**	1.89	-
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The length of the post-contract period used to calculate the production index is indicated in the header. All indices are normalized. All coefficients are multiplied by 100 for legibility. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered at county level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.11: Potential Selection by Pre-contract Weather (System Size)

System Size	(1) 1 month	(2) 1 month	(3) 2 months	(4) 2 months
$ProdIndex_{pre}$	0.024 [0.014]*	0.020 [0.014]	0.006 [0.013]	0.004 [0.012]
D.V. Mean	4.69	4.69	4.69	4.69
R^2 (within zip code)	0.204	0.206	0.204	0.206
N	153,108	153,108	153,108	153,108
Fixed Effects				
Zip Code	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Quarter	Yes		Yes	
County-quarter		Yes		Yes

Notes: All indices are normalized. Other system characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.12: Responses to Weather Updates

Cancel = 1	(1)	(2)	(3)	(4)	(5)
$ProdIndex_{pre}$	−0.039	−0.854	−0.591		
	[0.246]	[0.374]**	[0.354]		
$Update$		−0.769	1.170	0.472	0.588
		[0.264]***	[0.653]*	[0.252]*	[0.352]
$Indicator =$					
$1(Update < 0)$			−0.000		
			[0.005]		
$1(ProdIndex_{pre} > \overline{Index}_z)$				−0.014	
				[0.004]***	
$1(ProdIndex_{post} < \overline{Index}_z)$					0.006
					[0.004]
$Update \times Indicator$			−2.779	−1.705	−1.053
			[0.600]***	[0.332]***	[0.280]***
D.V. Mean	12.06	12.06	12.06	12.06	12.06
R^2 (within)	0.033	0.033	0.034	0.033	0.033
N	154,519	154,519	154,519	154,519	154,519
Fixed Effects					
Zip Code	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes
Quarter	Yes	Yes	Yes	Yes	Yes

Notes: $ProdIndex_{pre}$ is the normalized average production index for a 60-day pre-contract period. $Update$ is the average post-period index minus the pre-period one. $1(Update < 0)$ indicates whether $Update$ is negative. $1(ProdIndex_{pre} > \overline{Index}_z)$ is an indicator of whether the pre-period index is higher than the zip-code average, and $1(ProdIndex_{post} < \overline{Index}_z)$ is whether the post-period one is lower. All coefficients are multiplied by 100 for legibility. *Indicator* in the interaction term in the last rows refers to the indicator variable that is included in the same regression. System characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.13: Heterogeneous Effects by Area Characteristics

	Panel A. Local Demographics			
Cancel = 1	Median Income Above Median (1)	% White Above Median (2)	% College Above Median (3)	Household Size Above Median (4)
$\beta_1 : ProdIndex_{post}$	-0.664	-0.802	-0.715	-0.731
	[0.386]*	[0.292]***	[0.321]**	[0.337]**
$\beta_2 : ProdIndex_{post} \times 1(Character)$	-0.125	0.079	-0.064	-0.045
	[0.283]	[0.231]	[0.239]	[0.247]
$\beta_1 + \beta_2$	-0.789	-0.723	-0.779	-0.776
p-value: $\beta_1 + \beta_2 = 0$	0.003***	0.013**	0.006***	0.004***
	Panel B. Housing and Geographic Characteristics			
Cancel = 1	% Urban > 0.5 (5)	Housing Cost Above Median (6)	Northern California (7)	Non-coastal Counties (8)
$\beta_1 : ProdIndex_{post}$	0.381	-0.575	-0.875	-0.840
	[0.442]	[0.335]*	[0.313]***	[0.261]***
$\beta_2 : ProdIndex_{post} \times 1(Character)$	-1.211	-0.249	0.177	0.314
	[0.362]***	[0.233]	[0.211]	[0.260]
$\beta_1 + \beta_2$	-0.830	-0.824	-0.698	-0.526
p-value: $\beta_1 + \beta_2 = 0$	0.003***	0.003***	0.010***	0.137
Fixed Effects				
Zip Code	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Quarter	Yes	Yes	Yes	Yes

Notes: The area characteristic of interest is indicated in the header. The last two rows in each panel show the value of $\beta_1 + \beta_2$, which is the weather effect in zip codes with the characteristic of interest, and the corresponding p-value. The production index is normalized. All coefficients are multiplied by 100 for legibility. Other system characteristics and monthly economic conditions are always controlled for. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.14: Does Current Weather Predict the Near Future?

Future Index	(1)	(2)	(3)	(4)
ProdIndex	−0.001	−0.008	0.033	−0.114
	[0.026]	[0.025]	[0.027]	[0.136]
R^2	0.740	0.881	0.882	0.953
N	154,104	154,104	154,103	154,054
Fixed Effects				
Zip Code	Yes	Yes	Yes	Yes
Year		Yes	Yes	
Quarter		Yes		
County-quarter			Yes	
County-quarter-year				Yes

Notes: The dependent variable is mean daily production index in the one-year period starting right after the 113-day post-period. Standard errors (in squared brackets) are clustered at county level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.15: Weather and Complete Duration

Complete Duration	(1)	(2)	(3)	(4)	(5)	(6)
$Index_{post}$	−3.402	−3.170	−3.568	−3.430		
	[2.024]*	[1.951]	[2.235]	[2.200]		
$Index_{pre}$			−1.710	−2.729		
			[2.719]	[3.429]		
#days(prcp $>$ 0)					−0.093	−0.121
					[0.133]	[0.190]
wind speed					−2.186	−2.420
					[1.067]**	[1.595]
#days(tmax $<$ 40)					3.790	5.451
					[3.282]	[3.961]
#days(tmax $>$ 100)					0.238	−0.001
					[0.122]*	[0.085]
D.V. Mean	128.11	128.11	128.11	128.11	128.11	128.11
R^2	0.115	0.123	0.115	0.123	0.115	0.123
N	130,162	130,161	130,162	130,161	130,166	130,165
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Notes: The post-period production index is calculated based on the reported duration for each observation. Standard errors are clustered at county level. System characteristics and monthly economic conditions are always controlled for. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

A.3 Kleissl (2013) PV Performance Model

The PV performance model used in this paper is developed by Kleissl (2013), which is funded by the CSI Research, Development, Demonstration and Deployment (RD&D) Program. It uses satellite solar resource data to simulate PV power output while accounting for local weather conditions and system characteristics. It explicitly modeled the following components: irradiance on tilted surface, panel temperature effect on efficiency, DC-AC conversion efficiency, and maximum power point efficiency. The model parameters are then calibrated using CSI PV performance data.

I use this model to construct the solar production index, which is intended to measure how favorable the weather is for solar generation. The main weather inputs are daily solar insolation, maximum temperature, and wind speed. I use the same tilt, azimuth angle and system size for all applications so that the variations in the index only reflect differences in weather conditions. The predicted output for each application is then averaged over the relevant period and normalized to have mean zero and standard deviation one. Therefore, it is not necessary for the modeled output to completely match actual solar generation.

Kleissl (2013) has validated the 30-minute output of this model against measured power output at 192 PV sites over SDG&E, SCE, and PG&E territories. Fig 7-1 in Kleissl (2013) shows that the predictions are unbiased with typical errors of 4-9%. Because this measure is averaged over a much longer period in the current paper, such measurement errors should be significantly reduced.

The performance of this model is also compared against the PVWATTS model at three representative sites. PVWATTS is a standard calculator of PV performance in the industry. Table 6-1 in Kleissl (2013) shows that they have similar performance in terms of the relative mean bias error and relative mean absolute error, and the current model is slightly better calibrated for California.

A.4 Late Cancellation Reporting

As discussed in Section 1.3.1, the majority of cancellations have a reported cancellation period beyond one year, which is the program deadline for installation. These observations are referred to as "late reports" henceforth. In this section, we examine these late reports more closely to shed light on the nature of this data issue. We will first look at patterns in the timing of these reports and then move on to understanding the potential roles of program administrators and contractors.

Figure A.6 plots the distribution of reported cancellations by calendar month. Between 2007 and 2016, the number of reported cancellations varies substantially across calendar months. However, there is no discernible seasonal pattern, which suggests the reports are not likely to be driven by seasonal production cycles. Additional analysis also shows no evidence of concentrated reporting right before or after incentive step changes in the program.[1] Figure A.7 breaks down the sample by whether the observation is reported before or after the deadline. Again, there is no seasonal pattern in either of the distributions. [2]

Figure A.8 plots the full distribution of cancellation duration. Here, we find clear evidence of bunching right after the one-year deadline. More than half of the late reports are within 30 days from the deadline and 92.8% are within 180 days. Such bunching strongly suggests that the timing of these reports are related to the program deadline rather than the true time frame of cancellation. There is no incentive for the contractor or the customer to report a cancellation in time. As a result, a likely scenario is that the cancellation would remain unknown to the program administrator until the deadline when a report on the project status is required. Moreover, some cancellations reported before the deadline might also be subject to delay due to the same incentive problem.

[1] Results available upon request.

[2] October 2012 appears to be an outlier with an unusually high number of cancellations reported before the deadline. On close scrutiny, this is mainly due to October 11 and 12. Moreover, over 95% of cancellations reported on these two days are associated with Vertigo Solar. This suggests at least some cancellations reported before the deadline are subject to contractor delay.

The accuracy of the reported dates might be related to practices of the program administrator (PA). There is a separate PA for each of the three investor-owned utilities: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and Center for Sustainable Energy (for San Diego Gas and Electric (SDG&E) territory). In Figure A.9, I plot the distribution of reported cancellation duration separately for each of them. The pattern for each PA is qualitatively similar to the aggregate: most of the cancellations are reported late with clear bunching immediate after the deadline. This suggests that the administrative process is similar across PAs.

Next, we turn our attention to the contractors. As many contractors handle the CSI application for their customers, they might play a major role in determining how promptly the cancellations are reported. Table A.16 below lists the top 30 contractors by size (as indicated by the number of applications). There are a total of 2,398 contractors in the data, but the size distribution is highly right-skewed. For example, SolarCity is the largest contractor in the market with 15.1% of applications. It also has the largest number of cancellations and late reports. Together, the top 30 contractors account for about 60% of all applications, cancellations, and late reports. Among them, late reporting appears to be universal. The probability is substantial across the board and varies around the sample average of 0.62.

To sum up, we have three main findings. First, there is no seasonal pattern regarding the report timing of cancellations. Second, there is strong evidence of bunching right after the individual deadlines, suggesting the reported dates reflect more on program rules rather than the actual cancellation time frame. Last but not least, the problem is pervasive in the program regardless of the associated program administrator and the contractor.

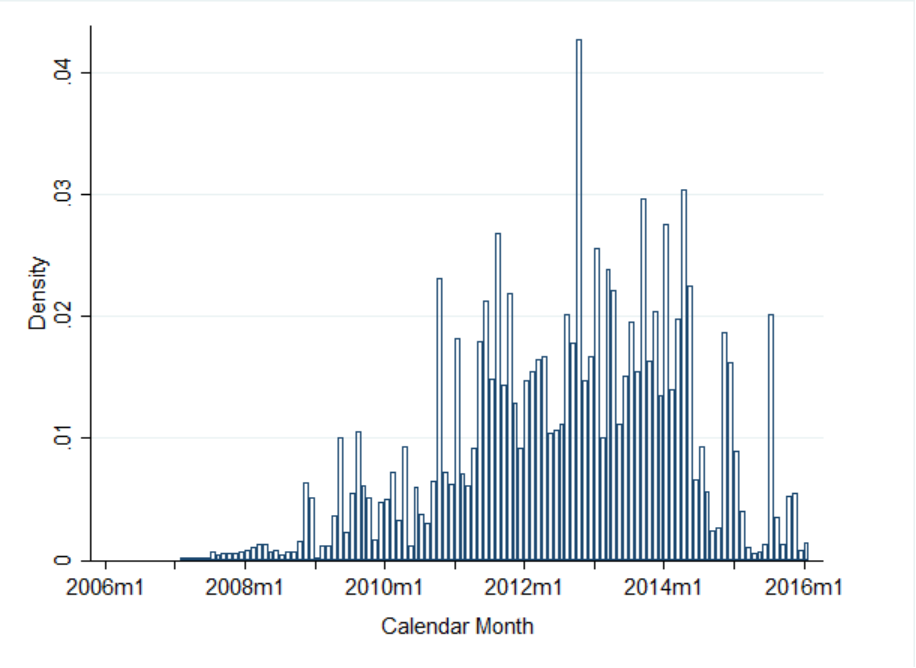


Figure A.6: Monthly Distribution of Reported Cancellations

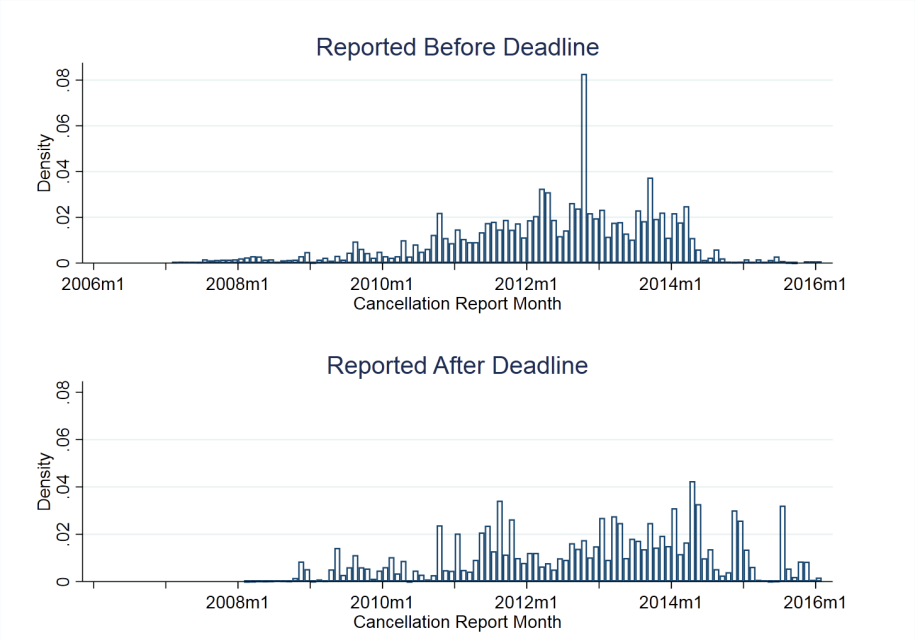


Figure A.7: Monthly Distribution of Reported Cancellations by Late Status

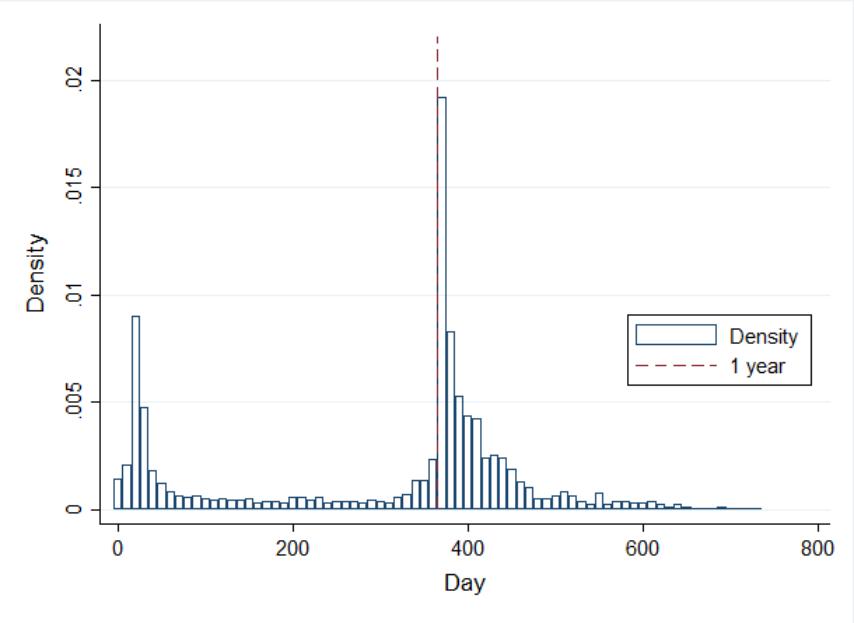


Figure A.8: Full Distribution of Reported Cancellation Duration

Notes: The red vertical line indicates the deadline at one year.

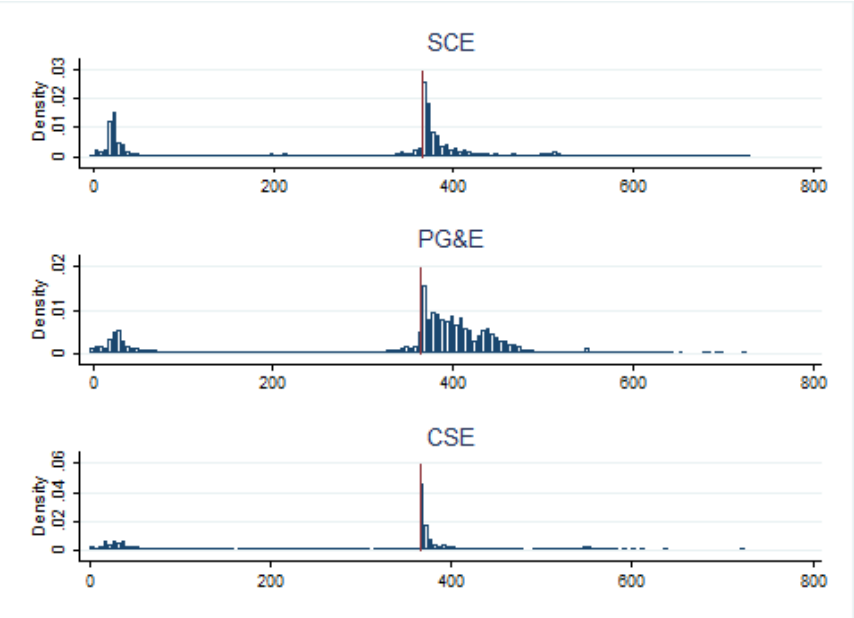


Figure A.9: Distribution of Reported Cancellation Duration by Program Administrator

Notes: The red vertial line indicates the deadline at one year. SCE accounts for 53.69% of all cancellations, PG&E 37.09%, and CSE 9.23%.

Table A.16: Top 30 Contractors by Number of Applications

Contractor	Rank	Apps %	Cancel %	Late %	Duration mean	Pr(Late)
SolarCity	1	15.10	15.82	17.83	347.7	0.697
Verengo Solar	2	7.748	11.98	9.081	282.3	0.469
Rec Solar	3	5.608	4.743	5.808	351.4	0.758
Real Goods Energy	4	3.827	2.289	2.686	330.5	0.726
Petersen-Dean	5	3.085	5.348	4.723	287.9	0.546
Sungevity	6	2.672	2.008	2.170	334.2	0.668
American Solar Direct	7	2.029	1.188	0.925	234.9	0.481
Akeena Solar	8	1.372	0.820	0.889	337.8	0.671
Galkos Construction	9	1.252	1.271	0.916	221.3	0.446
Burke Electric	10	1.252	2.058	2.855	378.9	0.858
Future Energy Corporation	11	1.074	0.473	0.107	98.87	0.140
HelioPower	12	1.072	1.150	1.388	331.4	0.746
Sullivan Solar Power	13	1.028	0.688	0.694	289.6	0.624
Solar Service Center	14	0.982	0.583	0.596	291.1	0.632
The Solar Company	15	0.906	0.677	0.720	333.6	0.659
Smart Energy Solar	16	0.803	0.930	0.694	221.1	0.462
Energy Efficiency Solar	17	0.747	0.556	0.614	309.4	0.683
Baker Electric Solar	18	0.736	0.341	0.231	222.9	0.419
Stout & Burg Electric	19	0.704	0.226	0.0800	213.9	0.220
Solar West Electric	20	0.581	0.138	0.0978	202.4	0.440
Vivint Solar Developer Llc	21	0.567	0.286	0.382	362.3	0.827
Advanced Solar Electric	22	0.565	0.501	0.569	377.9	0.703
Solar Network	23	0.553	0.374	0.382	281.9	0.632
Natural Energy	24	0.549	0.253	0.222	283.1	0.543
Sunlogic	25	0.545	0.462	0.543	349.8	0.726
Elite Electric	26	0.541	0.792	0.756	278.2	0.590
Mohr Power Solar	27	0.516	0.743	0.720	301.6	0.600
Sierra Pacific Home & Comfort	28	0.504	0.880	1.245	382.9	0.875
Sungate Energy Solutions	29	0.477	0.347	0.240	209.5	0.429
Skytech Solar	30	0.447	1.370	2.081	393.0	0.940
All Others	–	42.16	40.70	39.75	281.5	0.604
Total	1-2398	100	100	100	281.83	0.619

Notes: This table lists top 30 contractors by the number of applications. Columns 1-2 shows contractor name and rank, respectively. Column 3 shows, among all applications, the percentage associated with each contractor. Similarly, columns 4 and 5 show these percentages for cancellations and late cancellation reports, respectively. Column 6 shows the average duration as reported. Column 7 shows the probability of late report, which is the number of late reports divided by the number of cancellations. The second to last row shows aggregate statistics for the remaining contractors, and the last row shows aggregate statistics for all contractors.

Appendix B

Chapter 2 Appendix

B.1 Additional Figures

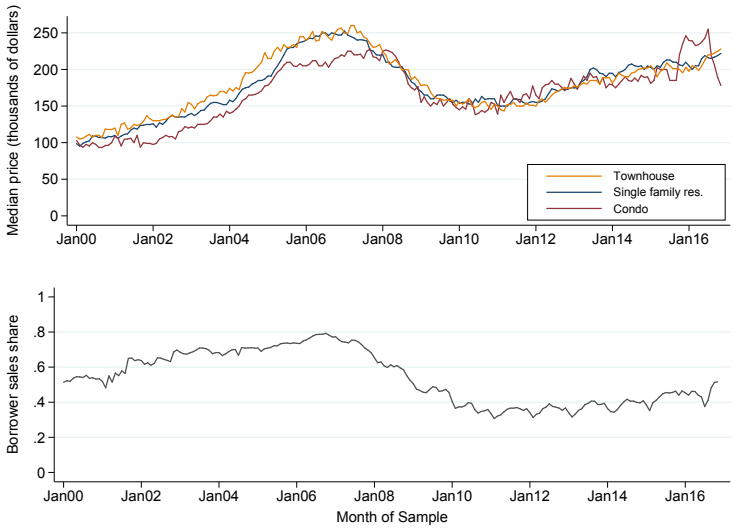
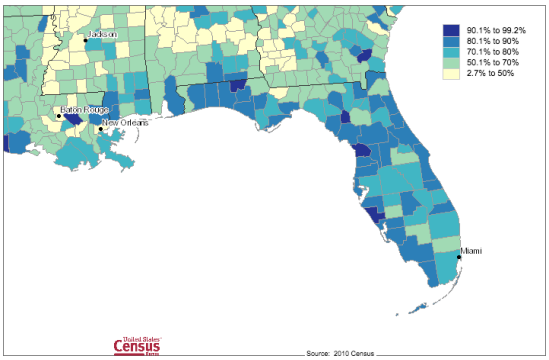
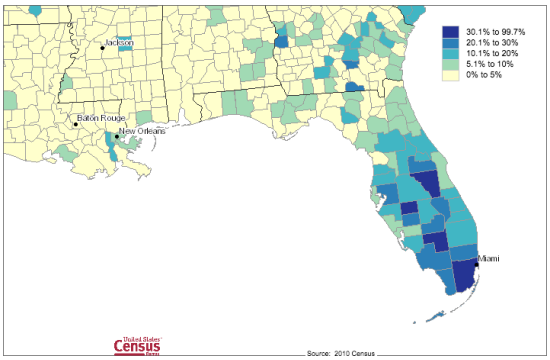


Figure B.1: Florida Borrowers Market Price and Sales Share

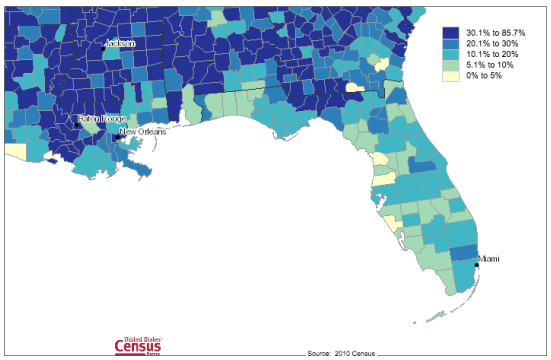
Source: Authors' calculation based on ZTRAX data.
Notes: The top panel plots the median price in the borrower subsample. The bottom panel plots the share of borrower sales among all sales. All time series are monthly.



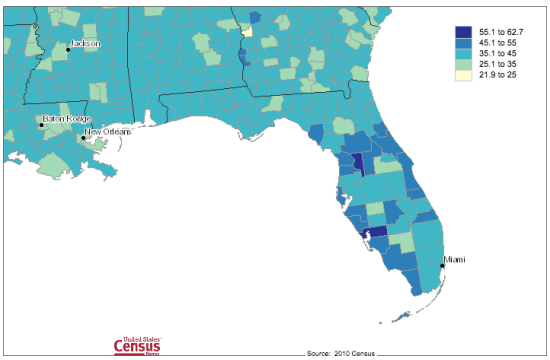
(a) Percent White



(b) Percent Hispanic or Latino



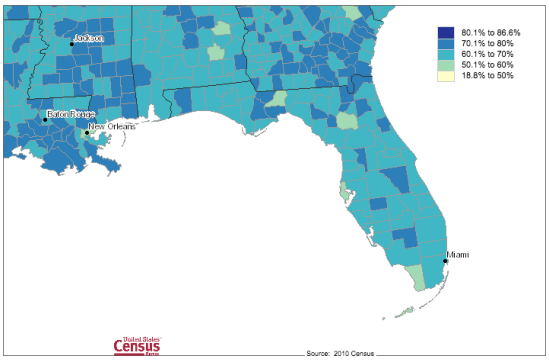
(c) Percent Black



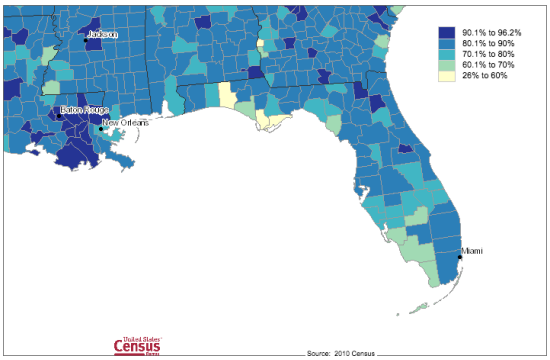
(d) Median Age

Figure B.2: Demographics in Florida Counties

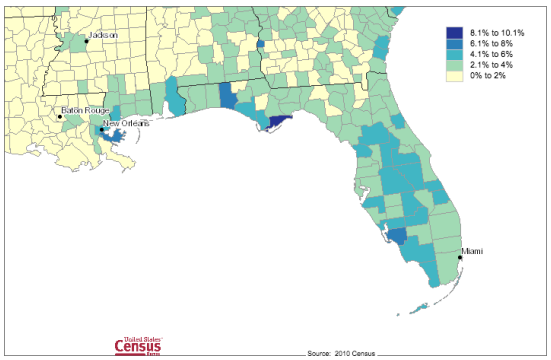
Source: Census Data Mapper based on data from the 2010 Census.



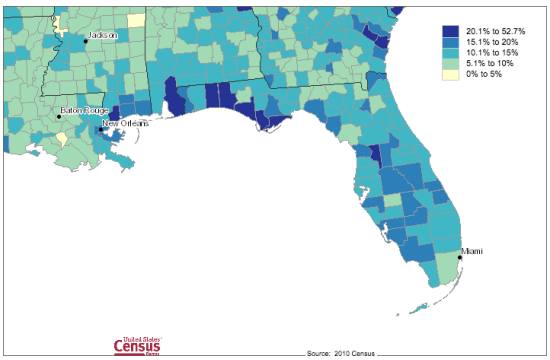
(a) Percent Family Households



(b) Percent Occupied Housing Units



(c) Homeowner Vacancy Rate



(d) Rental Vacancy Rate

Figure B.4: Family and Housing Statistics in Florida Counties

Source: Census Data Mapper based on data from the 2010 Census.

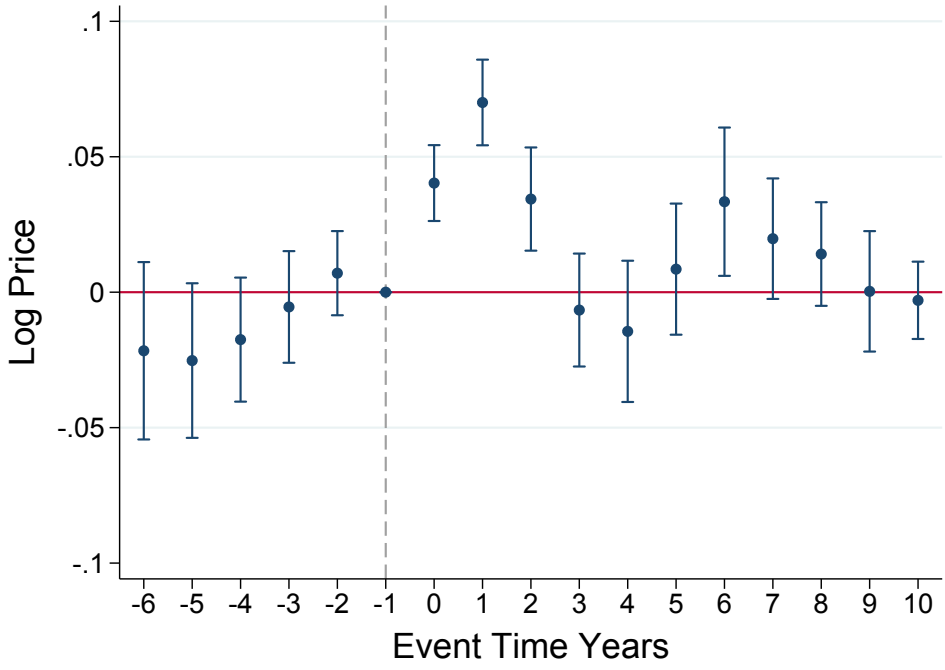


Figure B.6: Hurricane Effects on House Prices – Borrower Sample

Notes: Estimates from Equation (2.1) are plotted with their 95% confidence intervals. The model controls for standard hedonic variables and census tract, month, and county-year fixed effects. Standard errors are clustered by county. The results are based on the sub-sample of borrowers (N = 4,177,090).

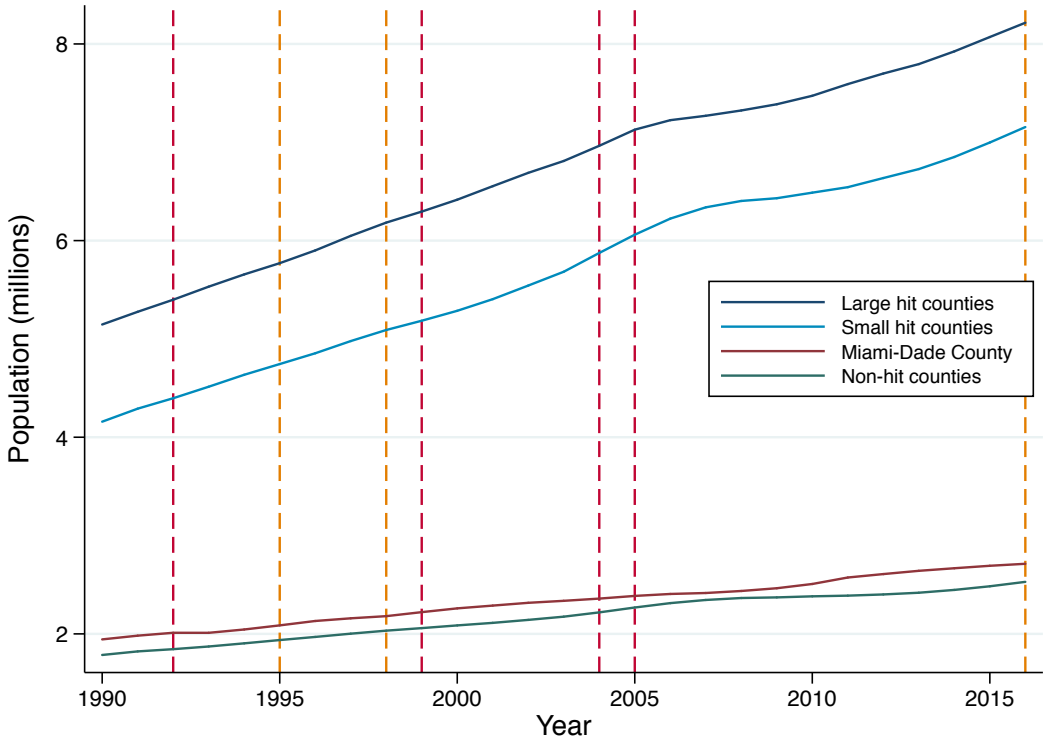


Figure B.7: Florida Population by County Aggregation

Source: Authors' calculation based on data from US Census Bureau.
Notes: Red dashed lines denote hurricane years in which at least 1/3 of hit county populations were affected (including Miami-Dade) according to the census tract hit definition, with the exception of Andrew in 1992, which affected all of Miami-Dade, but respectively only 17% and 5% of the other large and small counties. Orange dashed lines represent all other hurricane years. The 1995 and 1998 hurricanes only affected populations in the group of small counties (around 18% in both years).

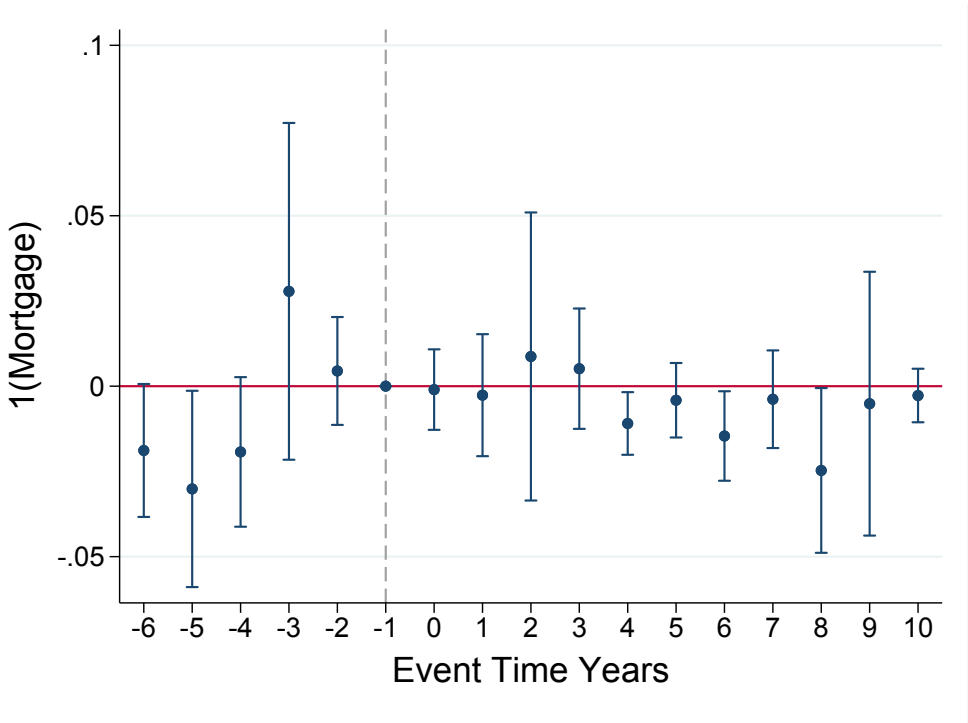


Figure B.8: Hurricane Effects on the Share of Transactions Involving a Mortgage

Notes: The dependent variable is an indicator of whether the transaction involves a mortgage. The estimates are from a linear probability model which controls for standard hedonic variables and census tract, month, and county-year fixed effects. Standard errors are clustered by county. The results are based on the full sample (N = 7,408,259).

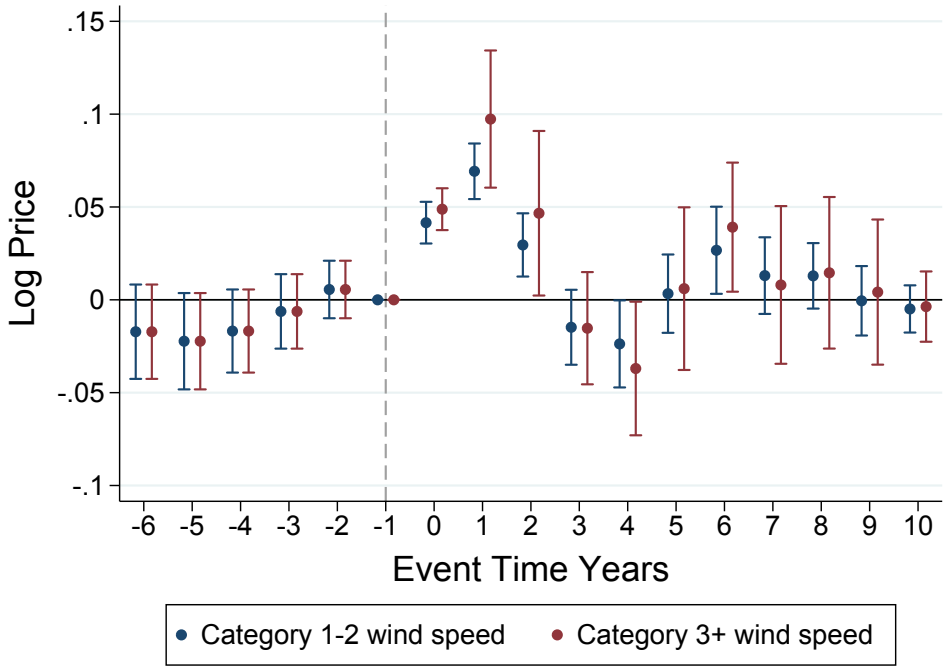


Figure B.9: Heterogeneous Price Effects by Hurricane Intensity – HMDA

Notes: Estimates for separate wind categories are plotted with their 95% confidence intervals. The dependent variable is log price. The model controls for standard hedonic variables and census tract, month, and county-year fixed effects. Standard errors are clustered by county. The results are based on the HMDA sub-sample. (N = 1,967,334).

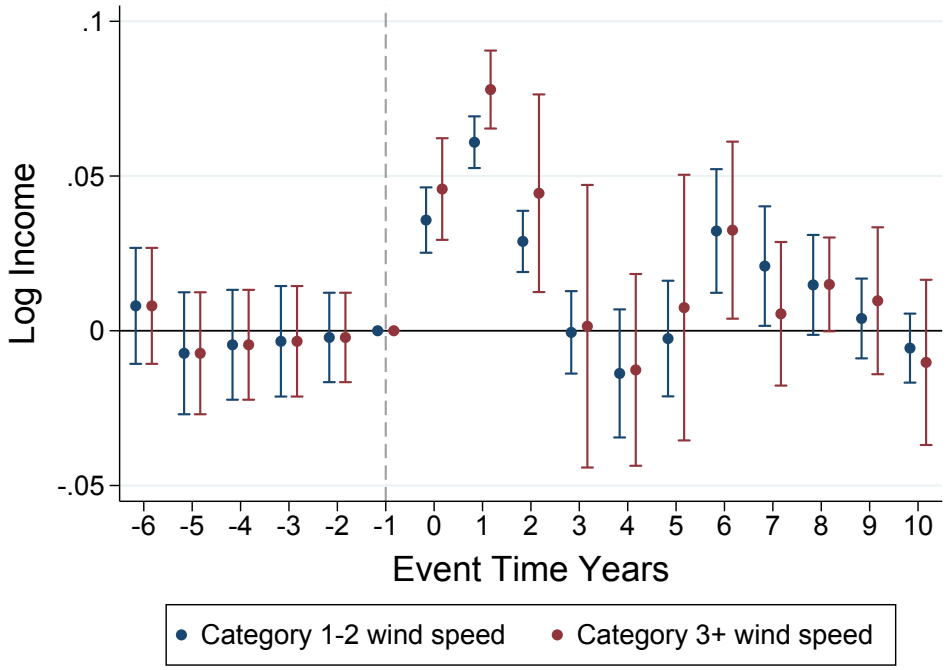


Figure B.10: Heterogeneous Income Effects by Hurricane Intensity – HMDA

Notes: Estimates for separate wind categories are plotted with their 95% confidence intervals. The dependent variable is log income. The model controls for standard hedonic variables and census tract, month, and county-year fixed effects. Standard errors are clustered by county. The results are based on the HMDA sub-sample. (N = 1,881,208).

B.2 Additional Tables

Table B.1: Summary Statistics for Housing Transactions

	Full Sample		Borrower Sample		HMDA Sample	
Variable	Mean	Median	Mean	Median	Mean	Median
Price	274793.9	155000	239283.2	178000	252847.4	187500
Assessed Value	150314	103208	149673.1	110516	159508.5	117750
No. of Buildings	0.955	1	0.960	1	0.951	1
No. of Stories	1.254	1	1.234	1	1.235	1
No. of Bathrooms	1.812	2	1.870	2	1.867	2
Lot Size	15094.95	8676	11730.01	8725	14624.76	8999
House Age	22.512	19	20.043	16	21.204	18
Effec. House Age	17.993	15	15.549	12	16.513	14
Loan Amount	–	–	196927.1	149809	198629.7	157102
Buyer Income	–	–	–	–	113683.2	73000
N	7,606,673		4,300,986		2,011,933	
% Single Family	70.52		76.62		77.35	
% Condo	23.49		16.96		16.56	
% Townhouse	6.00		6.42		6.08	

Notes: The unit of observation is a transaction. The full sample contains all transactions. The borrower sample contains all transactions with a mortgage. The HMDA sample contains all transactions that have a valid match to the HMDA records.

Table B.2: Summary Statistics for Parcel-Year Panel

	Statistics	Full Sample	Single Family	Condominium	Townhouse
Transacted[1]	Mean	0.107	0.110	0.098	0.099
Assessed Value	Mean	149359.6	154694.6	136878.4	131665
	Median	105360	110480	84223.5	106373
No. of Buildings	Mean	0.960	0.989	0.878	0.968
	Median	1	1	1	1
No. of Stories	Mean	1.223	1.107	2.059	1.258
	Median	1	1	1	1
No. of Bathrooms	Mean	1.665	1.743	1.371	1.972
	Median	2	2	2	2
Lot Size	Mean	16908.74	16255.06	29942.06	4180.165
	Median	9017	9496	8800	3485
House Age	Mean	30.171	31.859	27.700	21.324
	Median	28	29	28	21
Effec. House Age	Mean	23.842	23.884	25.076	18.139
	Median	22	21	26	18
N		49,523,080	33,709,921	12,836,476	2,976,683
Share (%)		100	68.07	25.92	6.01

Notes: The unit of observation is a parcel-year. The second column indicates the type of statistics reported.
[1] "Transacted" is a dummy variable indicating whether the parcel has been involved in a transaction in the given year.

Table B.3: Variations in Hurricane Indicators (Housing Price Model)

	Pre-Hurricane Event Year									
Hurricane	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
Andrew92	.	.	.	.	.	.	.	.	.	.
Opal95	.	.	.	.	.	.	.	.	.	.
Earl98	.	.	.	.	.	.	.	.	.	.
Georges98	.	.	.	.	.	.	.	.	.	.
Irene99	.	.	.	.	.	.	.	.	.	.
Charley04	.	.	.	.	.	1.3	2.0	2.1	2.3	2.5
Frances04	.	.	.	.	.	1.0	1.4	1.6	1.8	2.1
Jeanne04	.	.	.	.	.	1.1	1.4	1.6	1.8	1.9
Ivan04	.	.	.	.	.	0.1	0.1	0.2	0.2	0.2
Dennis05	.	.	.	.	0.1	0.1	0.1	0.1	0.2	0.2
Katrina05	.	.	.	.	0.5	0.8	0.8	0.8	0.9	1.0
Wilma05	.	.	.	.	2.6	3.1	3.4	3.8	4.1	4.6
Hermine16	0.1	0.1	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1
Matthew16	0.6	0.5	0.5	0.6	0.6	0.7	0.9	1.0	1.0	0.7
Irma17	1.1	1.3	1.6	1.7	1.7	2.0	2.1	2.3	1.9	0.3
Total	1.8	1.8	2.2	2.3	5.0	8.5	10.2	11.3	11.8	10.9

	Post-Hurricane Event Year										
	0	1	2	3	4	5	6	7	8	9	10
Andrew92	.	.	.	.	.	.	.	1.0	1.5	1.4	1.4
Opal95	.	.	.	.	0.2	0.3	0.3	0.4	0.4	0.4	0.3
Earl98	.	0.2	0.3	0.3	0.4	0.4	0.4	0.3	0.2	0.2	0.2
Georges98	.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Irene99	2.8	3.4	3.6	4.0	4.2	4.6	3.7	2.4	1.6	1.8	2.5
Charley04	2.5	2.1	1.4	0.9	0.9	1.4	1.5	1.5	1.8	1.8	1.9
Frances04	2.3	2.0	1.3	0.8	0.9	1.2	1.1	1.2	1.5	1.6	1.7
Jeanne04	2.0	1.5	1.0	0.7	0.7	1.0	1.1	1.1	1.3	1.4	1.4
Ivan04	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Dennis05	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Katrina05	0.9	0.6	0.4	0.3	0.6	0.6	0.6	0.7	0.7	0.7	0.6
Wilma05	3.4	2.2	1.5	1.8	2.5	2.6	2.6	3.0	3.1	3.1	2.3
Hermine16	0.0	.	.	.	.	.	.	.	.	.	.
Matthew16	0.1	.	.	.	.	.	.	.	.	.	.
Irma17	.	.	.	.	.	.	.	.	.	.	.
Total	11.7	10.3	8.4	8.1	9.3	10.7	9.9	10.2	10.5	10.8	10.7

Notes: House hit definition defined in Section 2.2.2. Dots denote hurricane event years not covered by the 2000-2016 sample.

Table B.4: Variations in Post-Hurricane Indicators (Transaction Probability Model)

	Pre-Hurricane Event Year									
Hurricane	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
Andrew92	.	.	.	.	.	.	.	.	.	.
Opal95	.	.	.	.	.	.	.	.	.	.
Earl98	.	.	.	.	.	.	.	.	.	.
Georges98	.	.	.	.	.	.	.	.	.	.
Irene99	.	.	.	.	.	.	.	.	.	.
Charley04	.	.	.	.	.	.	1.9	1.9	1.9	1.9
Frances04	.	.	.	.	.	.	1.3	1.3	1.3	1.3
Jeanne04	.	.	.	.	.	.	1.4	1.4	1.4	1.4
Ivan04	.	.	.	.	.	.	0.1	0.1	0.1	0.1
Dennis05	.	.	.	.	.	0.1	0.1	0.1	0.1	0.1
Katrina05	.	.	.	.	.	0.7	0.7	0.7	0.7	0.7
Wilma05	.	.	.	.	.	3.0	3.0	3.0	3.0	3.0
Hermine16	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Matthew16	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
Irene17	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9
Total	2.7	2.7	2.7	2.7	2.7	5.8	9.2	9.2	9.2	9.2

	Post-Hurricane Event Year										
	0	1	2	3	4	5	6	7	8	9	10
Andrew92	.	.	.	.	.	.	.	.	1.4	1.4	1.4
Opal95	.	.	.	.	.	0.3	0.3	0.3	0.3	0.3	0.3
Earl98	.	.	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Georges98	.	.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Irene99	.	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2	3.2
Charley04	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9	1.9
Frances04	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
Jeanne04	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Ivan04	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Dennis05	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Katrina05	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Wilma05	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
Hermine16	0.1	.	.	.	.	.	.	.	.	.	.
Matthew16	0.8	.	.	.	.	.	.	.	.	.	.
Irene17	.	.	.	.	.	.	.	.	.	.	.
Total	7.3	9.6	9.9	9.9	9.9	10.2	10.2	10.2	11.6	11.6	11.6

Notes: Parcel hit definition described in Section 2.2.2. Dots denote hurricane event years not covered by our 2000-2016 sample.

Table B.5: Variations in Cat-3 Hurricane Indicators (Housing Price Model)

	Pre-Hurricane Event Year									
Hurricane	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
Andrew92	.	.	.	.	.	.	.	.	.	.
Opal95	.	.	.	.	.	.	.	.	.	.
Earl98	.	.	.	.	.	.	.	.	.	.
Georges98	.	.	.	.	.	.	.	.	.	.
Irene99	.	.	.	.	.	.	.	.	.	.
Charley04	.	.	.	.	.	.	.	.	.	.
Frances04	.	.	.	.	.	0.0	0.1	0.1	0.1	0.1
Jeanne04	.	.	.	.	.	0.1	0.1	0.2	0.3	0.2
Ivan04	.	.	.	.	.	0.0	0.0	0.0	0.0	0.0
Dennis05	.	.	.	.	.	.	.	.	.	.
Katrina05	.	.	.	.	.	.	.	.	.	.
Wilma05	.	.	.	.	0.3	0.3	0.4	0.4	0.4	0.4
Hermine16	.	.	.	.	.	.	.	.	.	.
Matthew16	.	.	.	.	.	.	.	.	.	.
Irma17	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.0
Total	0.1	0.1	0.2	0.2	0.5	0.6	0.8	0.9	0.9	0.8

	Post-Hurricane Event Year										
	0	1	2	3	4	5	6	7	8	9	10
Andrew92	.	.	.	.	.	.	.	0.3	0.5	0.4	0.4
Opal95	.	.	.	.	.	.	.	.	.	.	.
Earl98	.	.	.	.	.	.	.	.	.	.	.
Georges98	.	.	.	.	.	.	.	.	.	.	.
Irene99	.	.	.	.	.	.	.	.	.	.	.
Charley04	.	.	.	.	.	.	.	.	.	.	.
Frances04	0.1	0.1	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1
Jeanne04	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1
Ivan04	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Dennis05	.	.	.	.	.	.	.	.	.	.	.
Katrina05	.	.	.	.	.	.	.	.	.	.	.
Wilma05	0.3	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.3	0.2
Hermine16	.	.	.	.	.	.	.	.	.	.	.
Matthew16	.	.	.	.	.	.	.	.	.	.	.
Irma17	.	.	.	.	.	.	.	.	.	.	.
Total	0.7	0.5	0.3	0.3	0.4	0.4	0.4	0.8	1.0	1.0	0.9

Notes: House hit definition described in Section 2.2.2. Dots denote hurricanes not reaching category 3 speeds in Florida or event years not covered by sample.

Table B.6: The Impacts of Hurricanes on Prices and Transaction Probability (Event Study)

	(1) Price – Full Sample	(2) Price – Repeat Sales	(3) Probability – Full Sample
Event Time Year -6	-0.004	-0.02	-0.003
	(0.04)	(0.04)	(0.002)
Event Time Year -5	0.002	-0.02	-0.004*
	(0.05)	(0.05)	(0.002)
Event Time Year -4	-0.009	-0.02	-0.003
	(0.03)	(0.05)	(0.002)
Event Time Year -3	-0.1	-0.03	-0.002
	(0.1)	(0.04)	(0.002)
Event Time Year -2	-0.008	-0.03	-0.003
	(0.03)	(0.03)	(0.002)
Event Time Year 0	0.08***	0.05	-0.009**
	(0.02)	(0.04)	(0.004)
Event Time Year 1	0.1***	0.1***	-0.004*
	(0.03)	(0.03)	(0.002)
Event Time Year 2	0.03	0.08***	-0.005*
	(0.03)	(0.03)	(0.003)
Event Time Year 3	0.02	0.03	-0.003*
	(0.02)	(0.03)	(0.002)
Event Time Year 4	-0.03	0.01	-0.001
	(0.03)	(0.03)	(0.002)
Event Time Year 5	-0.02	0.04	-0.0004
	(0.02)	(0.03)	(0.001)
Event Time Year 6	-0.03	0.06*	0.004
	(0.02)	(0.04)	(0.002)
Event Time Year 7	-0.01	0.05**	-0.0006
	(0.02)	(0.02)	(0.002)
Event Time Year 8	0.02	0.06***	-0.0005
	(0.04)	(0.01)	(0.002)
Event Time Year 9	-0.04	0.05***	0.0007
	(0.03)	(0.02)	(0.001)
Event Time Year 10	-0.005	0.05**	0.0008
	(0.01)	(0.02)	(0.001)
County-Year-Type FEs	Yes	Yes	Yes
Month-Type FEs	Yes	Yes	Yes
Tract FEs	Yes		
Parcel FEs		Yes	Yes
Hedonic Variables	Yes	Yes	Yes
N	1,741,743	1,338,384	49,120,139
R^2	0.61	0.78	0.086

Notes: Standard errors (in parentheses) are clustered at the county level. The unit of analysis is a transaction in columns (1)-(2), and a parcel-year in column (3). The hedonic variables include bins of number of stories and number of bathrooms, lot size, house age, and effective age. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.7: The Impact of Hurricanes on Housing Prices (Diff-in-Diff)

Log(Price)	(1) Full Sample	(2) Single Family	(3) Condominium	(4) Townhouse
Event Time Year 0-2	0.05***	0.04**	0.08***	0.02
	(0.02)	(0.02)	(0.008)	(0.02)
Age	-0.005***	-0.004***	-0.02***	-0.008***
	(0.0008)	(0.0004)	(0.005)	(0.001)
Effective Age	-0.008***	-0.007***	-0.003	-0.01***
	(0.0009)	(0.0008)	(0.004)	(0.005)
Lot Size	0.003***	0.003***	-0.0003	0.03**
	(0.0008)	(0.0005)	(0.0003)	(0.01)
County-Year-Type FEs	Yes	Yes	Yes	Yes
Month-Type FEs	Yes	Yes	Yes	Yes
Tract FEs	Yes	Yes	Yes	Yes
Hedonic Variables	Yes	Yes	Yes	Yes
N	7,408,259	5,221,707	1,741,743	444,809
R^2	0.56	0.55	0.61	0.60

Notes: Standard errors (in parentheses) are clustered at the county level. The unit of analysis is a transaction. The dependent variable is log transaction price. Other hedonic variables include the number of stories and the number of bathrooms in bins. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.8: The Impact of Hurricanes on Transaction Probability (Diff-in-Diff)

Transaction Probability	(1) Single Family	(2) Condominium	(3) Townhouse
Event Time Year 0-2	-0.003	-0.02**	-0.005*
	(0.003)	(0.01)	(0.002)
Effective Age	-0.0006	-0.0009	-0.002**
	(0.0005)	(0.0006)	(0.001)
Lot Size	0.00007**	-0.00002	-0.0003
	(0.00003)	(0.00003)	(0.0004)
County-Year-Type FEs	Yes	Yes	Yes
Parcel FEs	Yes	Yes	Yes
Hedonic Variables	Yes	Yes	Yes
N	33,651,628	12,834,214	2,976,067
R^2	0.087	0.083	0.097

Notes: Standard errors (in parentheses) are clustered at the county level. The unit of analysis is a parcel-year. The dependent variable is an indicator of whether a transaction takes place that involves the given parcel in a given year.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table B.9: Differential Capitalization

Log(Price)	Main Effect	Hurricane Interaction
Event Time Year 0-2	0.0698*	
	(0.0403)	
Home Characteristics		
House Age	-0.0061***	0.0007
	(0.0008)	(0.0007)
Effec. House Age	-0.0089***	0.0022**
	(0.0009)	(0.001)
One-Story	-0.0849***	0.0165
	(0.0193)	(0.0159)
Lot Size (1000 sq. ft.)	0.0028***	0.0013**
	(0.0008)	(0.0006)
No. of Bathrooms	0.1692***	-0.0355***
	(0.0294)	(0.0128)
N	7,408,259	
Within R^2	0.11	

Notes: This table reports the estimates from a single regression. The "Main Effect" column reports the main effects of home characteristics on the price, and the "Hurricane Interaction" column reports the effects of their interaction with the simplified DID treatment indicator. The model controls for tract-type, month-type, and county-year-type FEs. Standard errors are clustered by county. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Determining Hurricane Exposure

B.3.1 Imputing Maximal Reach Radius of 96 Knots Wind Speed

In this section, we describe how we calculate the maximal reach radius associated with a wind speed of 96 nautical miles per hour (kn). For each hurricane track point, we observe the radii associated with a wind speed of 34, 50, and 64 kn. We estimate the relationship between the maximal reach radii and wind speed using the following model:

$$log(Maxradius_{sht}) = \alpha_{ht} + \beta_1 Speed + \beta_2 Speed^2 + \varepsilon_{sht} \tag{B.1}$$

where α_{ht} are hurricane-track-point fixed effects and $Speed$ (s) takes one of the three speed values available. Note that instead of specifying the minimum pressure and maximum wind speed, we choose to employ a set of fixed effects which absorb their variations. The relationship between a wind speed threshold and its associated maximal radius is very well captured by this model as suggested by its estimation's R^2 of 0.93 (0.90 within track-point fixed effects). The estimated function is negative and concave, suggesting the radius decreases at an increasing rate when the wind speed increases. Full results are reported in Table B.10.

Table B.10: Wind Speed and Maximal Reach Radius Model

	log(Maxradius)
Speed	-0.0224***
	(0.004)
Speed^2	-0.0002***
	(0.00004)
Hurricane-track-point FEs	Yes
N	1188
R^2	0.93
R^2 (within)	0.90

Notes: Standard errors in parentheses (clustered at the hurricane-track-point level).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We use the estimates (including the track-point fixed effect) to predict the maximum radii

for 96kn wind speeds at track points where the maximum sustained wind speed is actually above 96 kn. This procedure raises the typical concerns regarding out-of-sample predictions since 96 kn is not within the support of wind speeds used in the estimation. To address this concern, we check the validity of our predictions by comparing them to the observed radius of the maximum speed, which is provided in the Extended Best Track dataset. In particular, because the radius is strictly decreasing in wind speed, our imputed radius should always be greater than (or equal to) the radius of maximum speed in these cases. Our prediction satisfies this condition for over 90% of imputations and is within rounding error (all radii measurements in the dataset are rounded to the nearest 5 nautical miles) for 10 of the 13 extrapolations for which it fails. In these 13 cases, we replace the model's predicted maximal radius for 96 kn speeds with the wider observed radius of the maximum speed reached.

B.3.2 Hurricane Track Interpolation and Exposure Calculation

In this section, we describe the procedure used to assign the exposure status of a house to a hurricane, with the geometric relationships illustrated in Figure B.11. Suppose we want to determine the exposure status of a house H_1 to a hurricane, whose track is recorded in three observations, P_1, P_2, and P_3. At each track point, we observe the maximal reach radius $r(\cdot)$ associated with 64 kn wind speed and the maximum wind speed $v(\cdot)$ at the center.

We take the following steps to calculate exposure:

1. We interpolate linearly between neighboring track points ($P_1 \to P_2$, $P_2 \to P_3$), and calculate the distance from the house H_1 to each linear segment. For instance, the distance between H_1 and P_1–P_2 is $d(H_1, I_1)$, or the distance between H_1 and its projection point I_1.

2. We calculate key variables for each projection point by interpolating between the observed track points. These variables are the 64 kn wind speed radius (denoted $r(\cdot)$) and the center wind speed (denoted $v(\cdot)$). They are calculated as a weighted average of the observed values associated with the two end points of the segment, weighted by inverse distances.

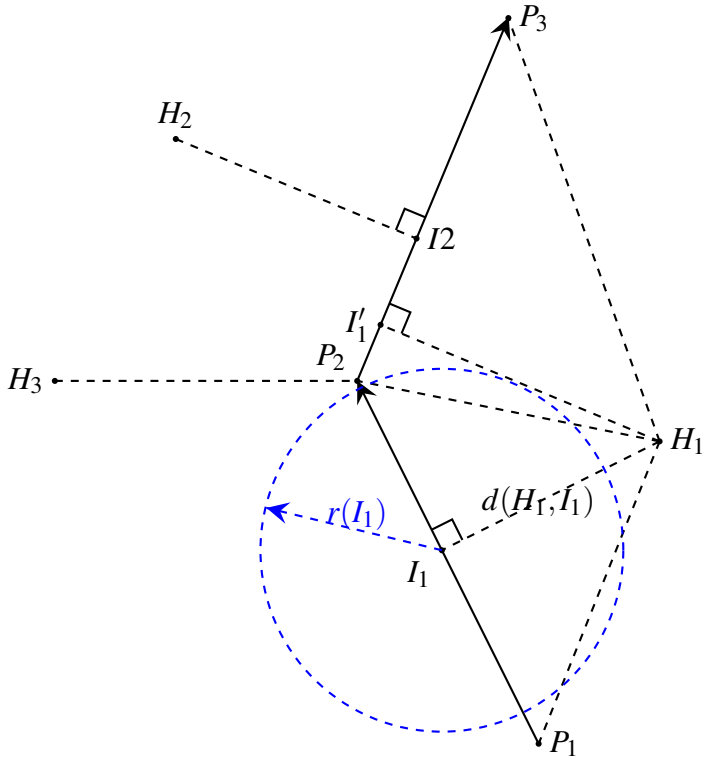


Figure B.11: Hurricane Track Interpolation

For example, the 64 kn wind speed for I_1 is calculated using observed radii of P_1 and P_2 as follows:

$$r(I_1) = \frac{d(I_1,P_2)}{d(P_1,P_2)} r(P_1) + \frac{d(I_1,P_1)}{d(P_1,P_2)} r(P_2).$$

Similarly, $v(I_1)$ is calculated using $v(P_1)$ and $v(P_2)$:

$$v(I_1) = \frac{d(I_1,P_2)}{d(P_1,P_2)} v(P_1) + \frac{d(I_1,P_1)}{d(P_1,P_2)} v(P_2)$$

3. We also calculate the distance of the house to each track point: $d(H_1,P_1)$, $d(H_1,P_2)$, and $d(H_1,P_3)$. We thus have a collection of points $\mathscr{S}_{H_1} = \{P_1, I_1, P_2, I_1', P_3\}$ with their corresponding distances to H_1, 64 kn wind speed radius, and center wind speed. $\mathscr{S}_{H_1}$ contains all possible points that might be the closest to H_1 on the interpolated hurricane track. We call it the potential exposure set.

4. For each point in $\mathscr{S}_{H_1}$, we check two conditions: (1) whether H_1 is within its 64 kn wind speed radius (*e.g.* $d(H_1, I_1) < r(I_1)$); (2) whether the center wind speed is above 64 kn (*e.g.* $v(I_1 \geq 64)$). If both are satisfied for any point in $\mathscr{S}_{H_1}$, H_1 is considered to be exposed to this hurricane.[1]

Our approach takes care of two general cases where the house does not lie "inside" the curve like H_1: the potential exposure set of H_2 consists of P_1, P_2, I_2, and P_3, while that of H_3 consists of P_1, P_2, and P_3. In practice, a hurricane track is observed in many more segments. We pre-select the four segments closest to the house and perform the above procedure to all points involved in those segments.

B.4 Zillow-HMDA Matching Procedure

We use the following procedure to match the HMDA data to transactions:

1. We first select the subset of HMDA loan applications that are (1) successful and (2) whose purpose is purchase.

2. We create all possible pairs of observations from Zillow and HMDA with the same year, census tract, and loan amount (in 1000s) using the "joinby" command in Stata. A small percentage of lenders make multiple loans of the same amount in a single census tract every year. If there also exist multiple Zillow records with these same characteristics, we drop all such matches because we cannot infer the exact mapping between the multiple observations on the two sides.

3. Lender names may be recorded differently across and even within Zillow and HMDA. Extensive manual inspection has revealed some general patterns of mismatch. We apply

[1] In general, the radius will be zero or close to zero when the center wind speed does not reach 64 kn. The small discrepancy comes from the linear change assumption. We take the conservative approach by requiring both conditions.

corresponding corrections to both datasets. For example, we replace acronyms such as "FCU" and "NB" with their full forms ("Federal Credit Union" and "National Bank").

4. We calculate the Jaccard similarity index for every pair of HMDA-Zillow observations using the "matchit" command in Stata. This index indicates the extent of the overlap between the strings containing the lenders' names from both datasets.

5. We keep all pairs with their Jaccard similarity index above an acceptable threshold. The threshold is chosen so that we observe it to produce the correct pairing in large random subsamples chosen from every year in our data.

6. Finally, we drop all cases when a single HMDA record is joined to more than one transaction in Zillow.

B.5 A Model on Income and Location Valuation

Our empirical results show that the average income of home buyers increases when the housing stock is exogenously reduced by hurricanes. In this section, we examine the theoretical implication of this finding for the underlying relationship between income and location preference in a simple model. We also explore the role of credit constraint.

Suppose a location j has a fixed and homogeneous housing stock S_j.[2] Consider a population of households with different income levels who might potentially live in location j. Suppose the amenity value of location j to household i with income Y_i takes the form

$$V_{ij} = \alpha_j Y_i + \theta_{ij}. \tag{B.2}$$

which consists of two components. The first component is income-related: α_j is the correlation between the amenity value of location j and household income. It can be positive, negative, or

[2] We assume a homogeneous housing stock for ease of notation. It is not essential for deriving our results and can be relaxed.

zero, depending on the characteristics of location j. The second component, θ_{ij}, is household-specific and orthogonal to income. We also assume it is bounded: $\theta_{ij} \in [\underline{\theta}, \overline{\theta}]$.

Denote the equilibrium housing price as p_j^*. Household i would buy a house here if and only if $V_{ij} \geq p_j^*$. The equilibrium is characterized by demand being exactly equal to the fixed housing stock:

$$S_j = \sum_i 1(V_{ij} \geq p_j^*). \tag{B.3}$$

Equation (B.3) implies that $p_j^*(S_j) < 0$ is a decreasing function: the larger the housing stock, the lower the price. In addition, the average income of buyers in the equilibrium is given by

$$\overline{Y_j}^* \equiv E\left[Y_i \middle| V_{ij} \geq p_j^*\right]. \tag{B.4}$$

Consider an exogenous shock that decreases the stock of housing, such as a hurricane. We have the following proposition:

Proposition 1. *(without credit constraint) If $\overline{Y_j}^*$ increases when S_j decreases, then $\alpha_j > 0$.*

Proof. We prove by contradiction. Suppose $\alpha_j < 0$, then

$$\begin{aligned}
\overline{Y_j}^*(S_j) &\equiv E\left[Y_i \middle| V_{ij} \geq p_j^*(S_j)\right] \\
&= E\left[Y_i \middle| \alpha_j Y_i + \theta_{ij} \geq p_j^*(S_j)\right], \\
&= E\left[Y_i \middle| Y_i \leq \frac{\theta_{ij} - p_j^*(S_j)}{-\alpha_j}\right]
\end{aligned}$$

where the second equality follows by substituting in Equation (B.2). Suppose the housing stock is decreased by ΔS. In order to maintain the relationship depicted in Equation (B.3), we have $p_j^*(S_j - \Delta S) > p_j^*(S_j)$. It follows that

$$\overline{Y_j}^*(S_j - \Delta S) = E\left[Y_i \middle| Y_i \leq \frac{\theta_{ij} - p_j^*(S_j - \Delta S)}{-\alpha_j}\right] \leq E\left[Y_i \middle| Y_i \leq \frac{\theta_{ij} - p_j^*(S_j)}{-\alpha_j}\right] = \overline{Y_j}^*(S_j),$$

which contradicts the assumption.

Suppose $\alpha_j = 0$, then Equation (B.4) is reduced to

$$\overline{Y_j}^* \equiv E\left[Y_i \middle| V_{ij} \geq p_j^*(S_j)\right] = E\left[Y_i \middle| \theta_{ij} \geq p_j^*(S_j)\right] = E\left[Y_i\right]$$

because θ_{ij} is orthogonal to Y_i. In this case, market adjustment will be associated with no change in average buyer income. Contradiction. □

We now consider the case with credit constraints, where household i are only able to buy a house worth up to cY_i and c is a positive constant. This is consistent with the two main forms of credit constraints often studied in the literature, where the mortgage limit is proportional to the household income, and up to a percentage of the house value (Greenwald, 2018). In this case, the equilibrium price is characterized by

$$S_j = \sum_i 1(p_j^* \leq \min(V_{ij}, cY_i)) \tag{B.5}$$

We then have the following proposition that relaxes the requirement of $\alpha_j > 0$ and is still able to generate the negative relationship between housing stock and average buyer income:

Proposition 2. *(with credit constraint) Suppose households are subject to a credit constraint so that they can only buy a house worth up to cY_i, where $c > 0$.*

1. *If $\alpha_j \geq 0$, then $\overline{Y_i}^*$ increases as S_j decreases.*
2. *If $-c < \alpha_j < 0$, there exist $\underline{S}$ and $\overline{S}$ such that, when $S_j \in (\underline{S}, \overline{S})$, $\overline{Y_i}^*$ increases as S_j decreases. (When $S_j < \underline{S}$, no transaction is possible. When $S_j > \overline{S}$, the credit constraint is not binding for anyone.)*

Proof. Suppose $\alpha_j > 0$. The average buyer income is given by

$$\begin{aligned}
\overline{Y_j}^*(S_j) &\equiv E\left[Y_i \middle| p_j^*(S_j) \leq \min(V_{ij}, cY_i)\right] \\
&= E\left[Y_i \middle| Y_i \geq \max\left(\frac{p_j^*(S_j) - \theta_{ij}}{\alpha_j}, \frac{p_j^*(S_j)}{c}\right)\right] \\
&< E\left[Y_i \middle| Y_i \geq \max\left(\frac{p_j^*(S_j - \Delta S) - \theta_{ij}}{\alpha_j}, \frac{p_j^*(S_j - \Delta S)}{c}\right)\right] \\
&= \overline{Y_j}^*(S_j - \Delta S).
\end{aligned}$$

The strict inequality follows from

$$\max\left(\frac{p_j^*(S_j) - \theta_{ij}}{\alpha_j}, \frac{p_j^*(S_j)}{c}\right) < \max\left(\frac{p_j^*(S_j - \Delta S) - \theta_{ij}}{\alpha_j}, \frac{p_j^*(S_j - \Delta S)}{c}\right).$$

Suppose $\alpha_j = 0$. Then we have

$$\begin{aligned}
\overline{Y_j}^*(S_j) &\equiv E\Big[Y_i \Big| p_j^*(S_j) \leq \min(\theta_{ij}, cY_i)\Big] \\
&= E\Big[E\Big[Y_i | p_j^*(S_j) \leq cY_i\Big] \Big| p_j^*(S_j) \leq \theta_{ij}\Big] \\
&= E\Big[Y_i \Big| Y_i \geq \frac{p_j^*(S_j)}{c}\Big] \\
&< E\Big[Y_i \Big| Y_i \geq \frac{p_j^*(S_j - \Delta S)}{c}\Big] \\
&= \overline{Y_j}^*(S_j - \Delta S).
\end{aligned}$$

The equality in the third row holds because θ_{ij} is orthogonal to Y_i.

Suppose $-c < \alpha_j < 0$. We define $\underline{S}$ and $\overline{S}$ as follows:

$$p_j^*(\underline{S}) = \frac{c\overline{\theta}}{c - \alpha_j}; \quad p_j^*(\overline{S}) = \frac{c\underline{\theta}}{c - \alpha_j}.$$

For any $S_j \in (\underline{S}, \overline{S})$, marginal buyers at equilibrium price $p_j^*(S_j)$ have average income defined by:

$$Y_j|_{p_j^*(S_j)} = E\Big[Y_i \Big| \frac{p_j^*(S_j)}{c} \leq Y_i \leq \frac{\overline{\theta} - p_j^*(S_j)}{-\alpha_j}\Big].$$

Then we have

$$\frac{dY_j|_{p_j^*(S_j)}}{dS_j} = \frac{p_j^{*'}}{c} f\Big(\frac{p_j^*}{c}\Big) - \frac{p_j^{*'}}{-\alpha_j} f\Big(\frac{\overline{\theta} - p_j^*}{-\alpha_j}\Big) > 0,$$

where $f(\cdot)$ is the density function of income assumed to follow a uniform distribution here. We can replace this assumption with $f(\cdot)$ being bounded on both sides while also change the range of negative values that α_j can take.

□

Appendix C

Chapter 3 Appendix

C.1 Additional Figures

Figure C.1: Example of an ActBlue Donation

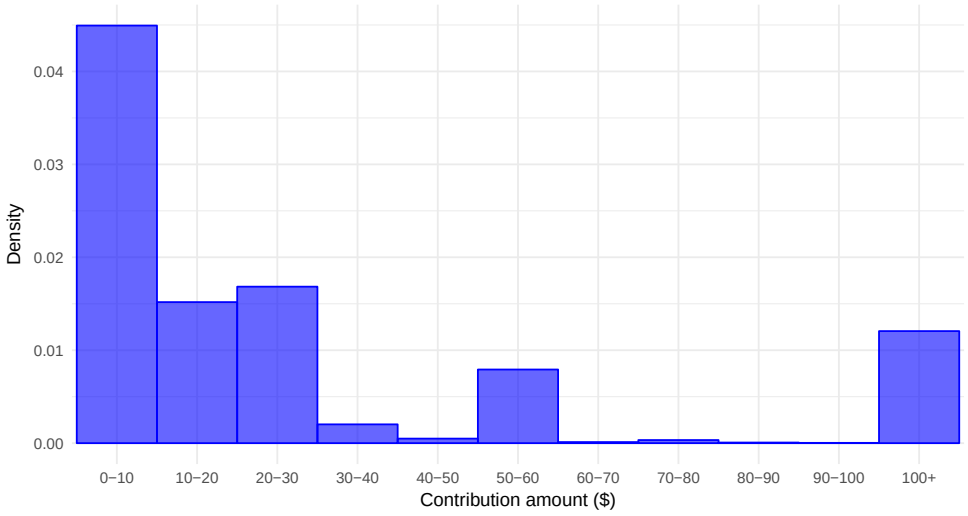


Figure C.2: Distribution of Individual ActBlue Donation Amounts

C.2 Additional Tables

Table C.1: FEMA Disaster Declarations, 1990-2012

Type	# Disasters	# County-Year Obs.
Coastal Storm	18	402
Drought	5	178
Fire	785	2,535
Flood	178	3,198
Freezing	6	127
Hurricane	233	7,680
Severe Ice Storm	44	1,566
Severe Storm(s)	738	13,187
Snow	126	2,745
Tornado	41	480
Tsunami	3	9
Typhoon	29	69
Total	2,206	32,176

Table C.2: Predicting Total Democratic Donations Using ActBlue Donations

Dep. Var.	Amount		Number	
	(1)	(2)	(3)	(4)
ActBlue	85.33***		14.67***	
	(5.63)		(1.30)	
ActBlue × 2006		209.93***		32.58***
		(24.98)		(8.14)
ActBlue × 2008		99.09***		21.95***
		(6.14)		(2.97)
ActBlue × 2010		57.51***		12.15***
		(6.32)		(2.44)
ActBlue × 2012		111.23***		14.63***
		(6.55)		(1.00)
Observations	200	200	200	200
R^2	0.74	0.86	0.83	0.88

Notes: the above table includes point estimates from various O.L.S. regressions of the amount and number of donations to Democrats from all sources, to the amount and donations from ActBlue sources. All regressions include an intercept term. Standard errors are clustered at the state level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.3: Positive and Negative Temperature Shocks on ActBlue Contributions

Dep. Var.	(1) Count/1M pop	(2) Avg. amount
$TmaxDev^{+}_{c,w}$	0.274^{***}	-0.070
	(0.098)	(0.052)
$TmaxDev^{+}_{c,w-1}$	0.110^{***}	-0.075
	(0.029)	(0.059)
$TmaxDev^{+}_{c,w-2}$	0.165^{***}	0.072
	(0.041)	(0.059)
$TmaxDev^{+}_{c,w-3}$	0.173^{***}	-0.096^{**}
	(0.037)	(0.042)
$TmaxDev^{+}_{c,w-4}$	0.124^{***}	-0.035
	(0.033)	(0.039)
$TmaxDev^{-}_{c,w}$	0.015	-0.096
	(0.031)	(0.059)
$TmaxDev^{-}_{c,w-1}$	-0.095	0.003
	(0.074)	(0.039)
$TmaxDev^{-}_{c,w-2}$	0.070^{*}	-0.041
	(0.041)	(0.051)
$TmaxDev^{-}_{c,w-3}$	0.036	-0.175^{**}
	(0.032)	(0.075)
$TmaxDev^{-}_{c,w-4}$	0.029	0.031
	(0.031)	(0.060)
N	935,201	935,201
R^2	0.204	0.054
Mean D.V.	15.40	13.19
County F.E.	Yes	Yes
Week F.E.	Yes	Yes
State-cycle F.E.	Yes	Yes

Notes: point estimates from equation (C.1) are shown. The dependent variable in column (1) is the number of ActBlue contributions per 1 million population. The dependent variable in column (2) is the average amount per ActBlue contribution. The sample consists of ActBlue contributions by week and county. Standard errors (in parenthesis) are clustered by county.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table C.4: The Effect of Extreme Temperature Events on ActBlue Contributions

Dep. Var.	(1) Count/1M pop	(2) Avg. amount
$TmaxHigh_{c,w}$	0.353^{*}	-0.043
	(0.183)	(0.138)
$TmaxHigh_{c,w-1}$	0.211^{**}	0.058
	(0.092)	(0.187)
$TmaxHigh_{c,w-2}$	0.224^{***}	0.215
	(0.082)	(0.211)
$TmaxHigh_{c,w-3}$	0.277^{**}	-0.058
	(0.135)	(0.126)
$TmaxHigh_{c,w-4}$	0.010	-0.032
	(0.159)	(0.125)
$TmaxLow_{c,w}$	-1.020^{***}	-0.318
	(0.368)	(0.229)
$TmaxLow_{c,w-1}$	-0.641^{*}	0.176
	(0.336)	(0.183)
$TmaxLow_{c,w-2}$	-0.083	-0.063
	(0.188)	(0.211)
$TmaxLow_{c,w-3}$	-0.315^{*}	-0.140
	(0.187)	(0.220)
$TmaxLow_{c,w-4}$	-0.372^{**}	0.442
	(0.170)	(0.368)
N	936,954	936,954
R^2	0.203	0.054
Mean D.V.	15.40	13.19
County F.E.	Yes	Yes
Week F.E.	Yes	Yes
State-cycle F.E.	Yes	Yes

Notes: point estimates from equation (C.2) are shown. The dependent variable in column (1) is the number of ActBlue contributions per 1 million population. The dependent variable in column (2) is the average amount per ActBlue contribution. The sample consists of ActBlue contributions by week and county. Standard errors (in parenthesis) are clustered by county.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table C.5: Heterogeneous Effects Across Quarters in Election Cycle

Dep. Var.	(1) Count/1M pop	(2) Avg. amount	(3) Amount/1M pop
$\overline{TmaxDev}^{2W}_{c,w} \times$			
Q1 (Dec-Feb)	–0.114**	0.401*	–1.901
	(0.049)	(0.241)	(4.431)
Q2 (Mar-May)	0.183**	0.086	–4.271
	(0.072)	(0.158)	(4.233)
Q3 (Jun-Aug)	0.681***	–0.273***	–3.407
	(0.093)	(0.078)	(5.062)
Q4 (Sep-Nov)	0.200***	0.054	2.314
	(0.051)	(0.062)	(4.076)
Q5 (Dec-Feb)	0.324***	–0.077 *	14.826*
	(0.073)	(0.046)	(7.697)
Q6 (Mar-May)	0.010	0.142**	–2.968
	(0.063)	(0.055)	(4.634)
Q7 (Jun-Aug)	0.023	0.223**	8.640
	(0.131)	(0.095)	(7.540)
Q8 (Sep-Nov)	1.256**	–0.795***	31.664*
	(0.596)	(0.126)	(18.835)
N	941,672	941,672	941,672
R^2	0.209	0.054	0.073
Mean D.V.	15.40	13.19	639.65
County F.E.	Yes	Yes	Yes
Week F.E.	Yes	Yes	Yes
State-cycle F.E.	Yes	Yes	Yes

Notes: point estimates from equation (C.4) are shown, which allows the estimates to differ by quarter-in-cycle. The dependent variable in column (1) is the number of ActBlue contributions per 1 million population. The dependent variable in column (2) is the average amount per ActBlue contribution. The dependent variable in column (3) is the ActBlue contribution amount per 1 million population. The sample consists of ActBlue contributions by week and county. Standard errors (in parenthesis) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table C.6: The Effects of Extreme Temperature on Amount Raised ($1,000)

Dep. Var.	(1) Total	(2) Total (PAC)	(3) Total (Ind.)	(4) Challenger	(5) Incumbent	(6) Share (C)
Hot	-106.4	-6.998	-56.97	-28.24	-78.15	0.00160
	(137.8)	(29.85)	(92.21)	(78.99)	(83.71)	(0.0187)
LCV	-75.00	-21.50	37.19	-121.5	46.53	-0.0269
	(257.1)	(40.37)	(210.4)	(97.15)	(192.8)	(0.0232)
Hot × LCV	123.5	0.871	65.81	12.90	110.6	0.00538
	(231.0)	(52.98)	(157.6)	(108.9)	(153.7)	(0.0277)
Cold	-43.86	0.333	-18.95	-28.16	-15.70	-0.00420
	(117.2)	(22.41)	(91.79)	(78.13)	(61.89)	(0.0136)
Cold × LCV	92.25	-23.10	85.53	22.17	70.08	-0.0122
	(154.3)	(34.22)	(129.0)	(93.32)	(91.45)	(0.0178)
Observations	3299	3299	3299	3299	3299	3299
R^2	0.411	0.531	0.386	0.259	0.456	0.285
Mean D.V.	1535.8	535.7	823.8	405.6	1130.2	0.196
Cycle F.E.	Yes	Yes	Yes	Yes	Yes	Yes
C.D. F.E.	Yes	Yes	Yes	Yes	Yes	Yes

Notes: point estimates from equation (3.4) are shown. The dependent variable in columns (1)-(5) is the amount of money raised in an election cycle from different sources in a given district, expressed in thousands of dollars. The dependent variable in column (6) is the share of total funds raised by the challengers. The sample includes competitive elections. Standard errors are clustered at the State level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.7: The Effects of Extreme Temperature on the Number of Donors

Dep. Var.	(1) Total	(2) Challenger	(3) Incumbent	(4) Share (C)
Hot	-204.2	-170.8**	-33.37	-0.0114
	(125.7)	(76.48)	(94.47)	(0.0219)
LCV	-342.7	116.7	-459.4	0.0225
	(371.5)	(171.5)	(285.3)	(0.0250)
Hot × LCV	476.0	230.6	245.4	0.00832
	(311.7)	(197.1)	(188.0)	(0.0369)
Cold	28.87	-33.83	62.70	0.00152
	(242.4)	(91.73)	(179.7)	(0.0171)
Cold × LCV	-16.09	7.949	-24.03	-0.0145
	(351.2)	(137.7)	(254.6)	(0.0344)
Observations	3299	3299	3299	3299
R^2	0.296	0.217	0.320	0.282
Mean D.V.	862.6	288.2	574.4	0.208
Cycle F.E.	Yes	Yes	Yes	Yes
C.D. F.E.	Yes	Yes	Yes	Yes

Notes: point estimates from equation (3.4) are shown. The dependent variable in columns (1)-(3) is the number of donors in an election cycle from different sources in a given district. The dependent variable in column (4) is the share of total donors corresponding to the challengers. The sample includes competitive elections. Standard errors are clustered at the State level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.8: The Effects of Extreme Temperature on Elections

	(1) Competitive	(2) Unopposed	(3) Open Seat	(4) Incumbent victory
Hot	-0.0574*	0.0516*	0.00582	0.00733
	(0.0289)	(0.0294)	(0.0181)	(0.0138)
LCV	-0.0144	-0.0220	0.0364	0.0732*
	(0.0467)	(0.0492)	(0.0310)	(0.0385)
Hot × LCV	0.0568	-0.0608	0.00397	-0.0197
	(0.0411)	(0.0374)	(0.0309)	(0.0278)
Cold	0.0245	-0.00792	-0.0166	0.0124
	(0.0292)	(0.0260)	(0.0167)	(0.0145)
Cold × LCV	-0.0722	0.0622	0.0101	-0.00588
	(0.0545)	(0.0448)	(0.0314)	(0.0208)
Observations	4478	4478	4478	4095
R^2	0.166	0.231	0.103	0.169
Mean D.V.	0.737	0.178	0.0855	0.948
Cycle F.E.	Yes	Yes	Yes	Yes
C.D. F.E.	Yes	Yes	Yes	Yes

Notes: point estimates from equation (3.4) are shown. The dependent variable in columns (1)-(3) is the probability of a congressional race being of a certain type (competitive, unopposed, open seat). The dependent variable in column (4) is the probability that the incumbent is re-elected. Columns (1)-(3) include all elections and column (4) excludes open seat elections. Standard errors are clustered at the State level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.3 Additional Specifications in the Short-Run Analysis

As shown in Section 3.3.1, our main estimating equation takes the form

$$Y_{cw} = \gamma' Weather_{cw} + \delta_w + \delta_c + \delta_{se} + \varepsilon_{cw}.$$

Specifications for Table 3.2:

– Columns (1) and (4):

$$Weather_{cw} = [TmaxDev_{cw}, PrcpDev_{cw}]^T.$$

– Columns (2) and (5):

$$Weather_{cw} = [TmaxDev_{cw}, ..., TmaxDev_{c,w-4}, PrcpDev_{cw}, ..., PrcpDev_{c,w-4}]^T.$$

– Columns (3) and (6):

$$Weather_{cw} = [\overline{TmaxDev}^{2W}_{c,w}, \overline{PrcpDev}^{2W}_{c,w}]^T$$

where

$$\overline{TmaxDev}^{2W}_{c,w} = \frac{1}{2}(TmaxDev_{cw} + TmaxDev_{c,w-1})$$
$$\overline{PrcpDev}^{2W}_{c,w} = \frac{1}{2}(PrcpDev_{cw} + PrcpDev_{c,w-1}).$$

Specification for Table C.3:

$$\begin{aligned} Weather_{cw} = [&TmaxDev^+_{cw}, ..., TmaxDev^+_{c,w-4}, TmaxDev^-_{cw}, ..., TmaxDev^-_{c,w-4}, \\ &PrcpDev_{cw}, ..., PrcpDev_{c,w-4}]^T \end{aligned} \tag{C.1}$$

where $TmaxDev^+ = TmaxDev \times (TmaxDev > 0)$ and $TmaxDev^- = |TmaxDev| \times (TmaxDev < 0)$. This specification allows us to estimate the effects of positive and negative deviations separately.

Specification for Table C.4:

$$Weather_{cw} = [TmaxHigh_{cw}, ..., TmaxHigh_{c,w-4}, TmaxLow_{cw}, ..., TmaxLow_{c,w-4}, PrcpDev_{cw}, ..., PrcpDev_{c,w-4}]^T, \tag{C.2}$$

where $TmaxHigh_{cw}$ is the total number of days in week w when the maximum temperature exceeds 95th percentile of the historical distribution in the month, and $TmaxLow_{cw}$ counts days with temperature below the 5th percentile.

Specification for Table 3.3:

$$Y_{cw} = \beta_1 \overline{TmaxDev}^{2W}_{c,w} + \beta_2 IncChar + \beta_3 \overline{TmaxDev}^{2W}_{c,w} \times IncChar + \gamma \overline{PrcpDev}^{2W}_{c,w} + \delta_w + \delta_c + \delta_{se} + \varepsilon_{cw}, \tag{C.3}$$

where $\overline{TmaxDev}^{2W}_{c,w}$ and $\overline{PrcpDev}^{2W}_{c,w}$ are defined as above. *IncChar* is a measure that characterizes the environmental stance of the incumbents associated with a county. Our coefficient of interest is β_3, which shows how the effects of weather shocks vary according to the incumbent characteristic.

Specification for Figure 3.2 and Table C.5:

$$Y_{cw} = \sum_{t=1}^{8} \beta_t \overline{TmaxDev}^{2W}_{c,w} \times Q_t + \gamma \overline{PrcpDev}^{2W}_{c,w} + \delta_w + \delta_c + \delta_{se} + \varepsilon_{cw}, \tag{C.4}$$

where $\overline{TmaxDev}^{2W}_{c,w}$ and $\overline{PrcpDev}^{2W}_{c,w}$ are defined as above. Q_t is a set of eight indicators for quarters in the election cycle. This specification allows us to obtain a separate estimate for each quarter-in-cycle.