Housing Dynamics in the Face of Shocks

by

Sebastian Sandoval Olascoaga

B.Sc., Centro de Investigación y Docencia Económicas, 2012
M.Sc., Paris School of Economics/École Normale Supérieure, 2015

Submitted to the Department of Urban Studies and Planning
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Urban Economics and Policy

at the

Massachusetts Institute of Technology

September 2023

Authored by: Sebastian Sandoval Olascoaga
Department of Urban Studies and Planning
August 16th, 2023

Certified by: Professor Mariana C Arcaya, Sc.D.
Department of Urban Studies and Planning
Thesis Supervisor

Accepted by: Professor Jinhua Zhao, Ph.D.
Department of Urban Studies and Planning
Co-Chair, Ph.D. Committee

Housing Dynamics in the Face of Shocks

by

Sebastian Sandoval Olascoaga

Submitted to the Department of Urban Studies and Planning
on August 16^{th}, 2023, in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy in Urban Economics and Policy

Abstract

This thesis explores housing dynamics in the context of shocks and understanding its impact on the housing sector, the well-being of communities, and the development of its citizens. It investigates the localized effect of extreme weather events on communities and individuals, the repercussions of the COVID-19 pandemic-induced eviction moratoria expiration on public health, and the influence of house flipping practices on neighborhood stability and housing affordability. This study sheds light on the critical role of housing stability in overall quality of life and societal progress, highlighting the pressing need for informed decision-making and policy formulation in the face of evolving challenges. The findings present implications for public health, climate resilience, neighborhood stability, and housing outcomes, contributing to the existing knowledge and paving the way for comprehensive housing systems that foster individual and societal well-being and prosperity.

Thesis Supervisor: Professor Mariana C Arcaya, Sc.D.
Title: Professor of Urban Planning and Public Health

Acknowledgments

How to begin? Embarking on this Ph.D. journey has been a team effort. While this thesis carries my name, it is a product of countless hours of guidance, support, and encouragement from many. I owe a debt of gratitude to all who believed in me, stood by my side, and kept me going. This work is as much theirs as it is mine, and for that, I am deeply thankful.

I don't have enough words to thank my mentor and brilliant advisor, Mariana Arcaya. Her guidance and unwavering belief in me have been the backbone of my academic journey for the last five years. Every challenge I encountered was met with her patient advice, and every achievement was celebrated with her genuine enthusiasm. Mariana's dedication to mentoring is not just about academic rigor; it's also about treating her advisees as equals and real human beings. I was immensely fortunate for her to trust my judgment and potential, and I owe my accomplishments to her invaluable insights, wisdom, and steady support.

My advisor, Atheendar Venkataramani, embodies humility, brilliance, and empathy in a manner that has made him a role model for me from the very start. Atheen, I am profoundly grateful for your decision to join my committee and our joint research without questions asked. I also had the luck to have Albert Saiz and devin michelle bunten serve as advisors. Albert, thank you for showing me how to remain resilient and patient during complex and volatile times. Being advised by you has been a privilege. devin, I am profoundly grateful for the endless hours of discussions, feedback, and encouragement. Thank you for keep challenging me and encouraging me to improve my work.

I wish to extend my heartfelt gratitude to several institutions, organizations, and departments for their generous financial support during my doctoral journey. First and foremost, I would like to thank the Massachusetts Institute of Technology for generously financing my studies for the last five years. From generously giving me a Presidential Fellowship to supporting me through annual research and teaching assistantships, they allowed me to fulfill a lifelong dream. The Boston Area Research Initiative generously gave me a research seed grant that kicked off my analysis, data collection, and writing of the third chapter. The Priscilla King Gray Public Service Center at MIT also generously supported me in 2021 with two fellowships during January and the summer to further develop and implement my thesis. The School of Architecture and Planning and the Department of Urban Studies and Planning not only helped me with teaching and research assistantship positions but also extended me a Lloyd and Nadine Rodwin International Fellowship and a Harold Horowitz Research Fund, which allowed me to write my first chapter of the dissertation and my first-year paper.

I also wish to convey my sincere appreciation to Schery, Daryl, Taylor, Chen, Ben, Reginald, Lily, Dana, Glenn, and the entire PR, Press, and Products Teams at Redfin. During my two enriching summers at Redfin, I was embraced with open arms and given opportunities that profoundly influenced both my research and career trajectory. The knowledge and hands-on experience I amassed during that time have been pivotal, leaving a long-lasting mark on my academic and professional journey. Interning at Redfin changed the way how I think about the relationship between extreme weather events, climate change, migration, and housing. The supportive and intellectually stimulating environment over there not only equipped me with a broader understanding of the field but also underscored the importance of bridging the gap between academia, industry, and the public. I am profoundly thankful for the mentorship and opportunities I received, and for the lasting impact it has had on my academic journey.

A big acknowledgment to OptumLabs. Their guidance and unwavering support, especially during my visiting fellowship, proved instrumental. Their generosity in sharing data amidst the COVID-19 pandemic was pivotal in our efforts to understand the disparities the pandemic inflicted upon minority populations. I am particularly indebted to Pamela Morin, Emily Carbonara, and Alia Rawji for their steadfast patience and support. The ethos at OptumLabs, emphasizing the importance of a health equity perspective in all research, has left a profound impact on my academic trajectory.

I want to extend my deepest appreciation to each member—especially to the Resident Researchers—of the Healthy Neighborhood Research Consortium. Their commitment and dedication to understanding community dynamics and health outcomes provided an invaluable foundation for this thesis. Special thanks to Vedette Gavin, Reann Gibson, Andrew Seeder, Yael Nidam, Gabriela Zayas del Rio, Samara Ford, Rebecca Houston-Read, Julia Curbera, and César Garcia Lopez. Their unwavering dedication, combined with their exceptional support and relentless hard work, has been the backbone of this endeavor. I am genuinely grateful for the opportunity to have worked alongside such a passionate and knowledgeable group.

I owe a deep debt of gratitude to Justin Steil. His insights into the relationship between law, housing, and inequality have been invaluable. Beyond sharing his knowledge with me, his guidance and time generosity have fostered both my professional development and personal growth. Siqi Zheng deserves big acknowledgment. Her passion for urban and environmental economics was pivotal during my qualifying exams and further discussions. Thank you for always being available to discuss research ideas and provide guidance. I am also grateful to Christopher Zegras, Lawrence Vale, Eran Ben-Joseph, and David Shu for their unwavering kindness, support, and guidance throughout my MIT journey. I am also indebted to Ezra Glenn, Takeo Kuwabara, Sandy Wellford, Karen Yegian, Peggy Bryan, Sandra Elliot, Janine Marchese, and Ellen Rushman. Their consistent support, hard work, and dedication ensure the seamless execution of our academic missions. To a group of esteemed professors and researchers whose collective wisdom and academic insight have enriched and shaped my academic journey: David Autor, Amy Finkelstein, William Wheaton, Josh Angrist, Esther Duflo, Abhijit Banerjee, Benjamin Olken, James Poterba, Joseph Newhouse, Alberto Abadie, and Iván Werning.

I am thankful to have met such an intelligent, friendly, and engaged community at MIT. Madeleine Daepp, Andrew Binet, Shin Bin Tan, Chenab Navalkha, Priyanka DeSouza, Dorothy Tang, Gokul Sampath, Lidia Cano, Margaret Haltom, Mohit Karnani, Evan Soltas, Charlie Rafkin, Ari Bronsoler, and Noémie Sportiche provided invaluable support, insights, and motivation throughout my Ph.D. journey. Thank you to my friends Nick Allen, Denise Tellbach, and Vaishnavi Ramaswamy. My time at the Graduate Student Council with you was full of good memories, and fantastic achievements, and it gave me a sense of community during the pandemic. You are all genuinely good human beings that deeply care about the well-being of students and their families. Thank you, my friends! To Carmelo, Daniel, Darien, Anna, Kevin, Binzhe, and Silvia. Thank you all for being an amazing Ph.D. cohort!

A mi querida familia. Kira, mi primera perrita, y el sol que unió a nuestra familia durante muchos años. Gracias por amarnos sin pedir nada a cambio. Te extraño cada día y agradezco todas las bienvenidas (y despedidas) que me diste sin importar la hora del día o el clima. Lucas, el más apuesto (y fotogénico) perrito, gracias por cuidar de mis padres en nuestra ausencia y por ser tan cariñoso con todos. A mi hermano Carlos: fue un regalo y un privilegio tenerte a mi lado en Cambridge mientras realizaba mi doctorado. Tu apoyo fue esencial para ayudarme a comprender y adaptarme a un entorno completamente nuevo, marcando la diferencia entre el éxito y el fracaso. Observarte enfrentar desafíos, superarte día a día y tu incansable búsqueda de soluciones para problemas complejos, me inspira y me recuerda la importancia de la perseverancia y la pasión. Gracias por ser no solo un hermano, sino también un amigo y un faro de inspiración a lo largo de todas estas décadas. A mi papá, Francisco (Paco). Muchas gracias por quererme y apoyarme como soy. Tu entusiasmo por el progreso del país y la sociedad me llena de esperanza. Tu cariño, amor, y lealtad a nuestra familia me da un excelente modelo a seguir cuando tenga mis hijos. Gracias por tu apoyo y amor. A mi hermosa mamá, Lilia. Desde el primer aliento que tomé hasta este mismo momento, has sido el faro constante en mi vida. Tu amor incondicional, tu paciencia infinita y tu sabiduría inquebrantable han guiado cada paso que he dado. Gracias, mamá, por tus sacrificios, por tu apoyo incondicional, por tus enseñanzas y, sobre todo, por ese amor eterno.

To Leia, my Binbin. You have brought so much joy and love to our home. Smile, smile, smile (bark, bark) is your anthem, and I cannot believe we have such a novel and beautiful dog. Thanks for being there during my thesis writing and showing us unconditional love.

From the moment our paths crossed at Lamplighter, my life transformed in ways I never imagined. Your presence has been a beacon of love, support, and unwavering faith throughout my Ph.D. journey. Every day, I stand in awe of your brilliance, your compassion for those you hold dear, and your indomitable spirit. You possess an innate courage to seek the truth and an unwavering passion to defend those you cherish. Thank you for the countless runs near the Arboretum, the shared meals, the infectious laughter, and the dances that make time stand still. My gratitude for you deepens with each passing day. My love for you grows immeasurably. You have been, and continue to be, the very foundation of my academic voyage. I love you more each day.

Time to celebrate!

Contents

List of Figures

List of Tables

Chapter I

Introduction

Housing is fundamental to human well-being, closely connected to various social, economic, and health outcomes. Access to safe, stable, and affordable housing is not only a basic human need but also a key determinant of overall quality of life. Beyond providing shelter, housing is pivotal in shaping individual opportunities, community cohesion, and societal progress. It serves as a platform for personal and family development, enabling individuals to thrive, pursue education, and engage in productive employment. Moreover, housing forms the bedrock of vibrant and resilient communities, fostering social interactions, promoting social cohesion, and contributing to the overall fabric of society.

However, the relationship between housing and societal outcomes can be complex. We live in a world of constant, abrupt, and intense non-linear changes—or *shocks*—that pose challenges to housing dynamics and community well-being. Earth's climate is changing faster than at any point in the history of modern civilization (USGCRP, 2018), increasing the probability of someone experiencing an extreme weather event. Rising temperatures, changing precipitation patterns, and increased frequency and intensity of natural disasters have significant implications for housing resilience and adaptability. Tornadoes, hurricanes, floods, and wildfires can cause widespread destruction, displacing individuals and disrupting communities. Understanding the localized impacts of these events on housing markets, economic activity, educational outcomes, and social dynamics is crucial for effective disaster management, resilience planning, and equitable recovery.

Furthermore, recent events and emerging health trends have brought the importance of housing to the forefront, underscoring the urgent need to understand the dynamics at play. The COVID-19 pandemic, for instance, has highlighted the critical importance of secure housing in promoting public health. As individuals and families faced job losses, financial strains, and heightened vulnerability, the significance of stable housing became starkly apparent. To protect individuals from losing their homes and mitigate the potential for further virus spread, eviction moratoria were enacted by local, state, and federal governments. Examining the implications of lifting these moratoria can provide valuable insights into the associations between housing instability and public health outcomes. Understanding these shocks and the potential policies to mitigate and adapt to them will become more valuable as recent estimates suggest an increase in the likelihood of experiencing another one-in-a-lifetime epidemic in the coming years (Marani et al., 2021).

Moreover, the rise of certain housing market practices, such as house flipping, has raised concerns about the equitable distribution of resources, neighborhood stability, and housing affordability. House flipping, characterized by the purchase of housing by an investor who attempts to profit from buying low and selling high rather than for occupation, can contribute to rapidly changing neighborhoods, often leading to gentrification, displacement, and affordability challenges. Exploring the mechanisms and effects of house flipping in specific contexts provides insights into the potential harms and benefits associated with this practice. It could also inform policy interventions to foster inclusive and sustainable housing markets.

Considering the interconnected nature of these housing-related challenges, it becomes clear that a comprehensive understanding of the implications of housing dynamics is essential. By examining the social, economic, and health outcomes associated with housing instability, extreme weather events, and housing market practices, this dissertation seeks to contribute to the existing body of knowledge and provide fresh insights for informed decision-making and policy formulation. The findings of this research have important implications for public health, climate resilience, neighborhood stability, and equitable housing outcomes. By addressing these challenges, policymakers, researchers, and communities can work together to create housing systems that support the well-being and prosperity of individuals and societies as a whole.

This thesis begins by exploring the effects of localized climate shocks on places and people. It continues by analyzing how the eviction moratoria expiration in 2020 affected the COVID-19 infection risk and spread. Finally, it investigates the housing practice of house flipping within rapidly changing neighborhoods in the Greater Boston Area. The following is a summary of each chapter.

I The Effects of Localized Climate Shocks on Places and People

Extreme weather events, known as climate shocks, are increasingly frequent and severe, impacting economic and social development. My first chapter focuses on tornadoes as a specific type of climate shock and examines their heterogeneous impacts on people and places. By leveraging the exogenous direction and width of tornado paths in the United States between 1996 and 2019, the longitudinal effects on various aspects, including the real estate sector, local governments, school outcomes, household behavior, and business activity, are analyzed through a generalized difference-in-differences event study estimator.

The results demonstrate that tornadoes lead to a sustained decrease in property values in affected regions, accompanied by an increase in foreclosures and a decline in investor interest. Suburban areas experience a continuous downward price adjustment, while the central business district exhibits more variation. Recovery funds and tornado intensity have no significant impact on property prices. Tornadoes also lead to a permanent decrease in the tax base of localities, reduced school test scores, immediate displacement of people, and a decline in the percentage of small businesses.

Additionally, tornadoes generate spillover effects on adjacent areas. The findings contribute to the understanding of the economic and social implications of climate shocks and highlight the need for targeted policies to minimize post-disaster inequalities and promote resilience.

II Eviction Moratoria Expiration and COVID-19 Infection Risk

The second chapter focuses on understanding whether lifting a state-level eviction moratorium impacted the risk of individuals being diagnosed with COVID-19. Here, we use a cohort of 509,694 individuals living in the United States and a difference-in-differences survival analysis. The findings reveal that residents in states that decided to lift eviction moratoria were at an increased risk of receiving a diagnosis of COVID-19. The increased risk becomes more pronounced 12 weeks post the ending of the eviction moratorium when compared to residents in states where the eviction moratoria were kept in place.

As time progressed, the impact strengthened, pointing to a potential cumulative effect. This increase over time was particularly pronounced among individuals with a higher number of comorbidities and those belonging to lower socioeconomic status groups. These findings suggest that eviction-led housing insecurity could potentially be a significant contributing factor to the exacerbation of the COVID-19 pandemic. This amplification may be due to the resultant increase in population mobility, the disruption of social distancing efforts, and the forced congregation of evicted individuals in shared or public spaces, all of which are well-established risk factors for the transmission of COVID-19.

Future policy measures to mitigate the spread of *health shocks* should consider housing stability as a key element, particularly in communities with a higher prevalence of comorbidities and lower socioeconomic resources. The results of this research contribute to a broader understanding of how socioeconomic factors interplay in the context of a global *health shock* and provide evidence for policymakers to make informed decisions.

III House Flipping Within Rapidly Changing Neighborhoods

The rise of house flipping, the practice of buying a property with the intention of quickly renovating and reselling it for a profit, has taken the real estate market by storm, with investors and speculators alike seeking to turn a quick profit. However, as the practice gains popularity, concerns about its impact on rapidly changing neighborhoods have emerged.

This chapter delves into the mechanisms and effects of house flipping in such neighborhoods in the Greater Boston area, employing a mixed-methods approach that incorporates various data sources and analytical techniques, including digitized scans of HOLC maps, transaction records from the state of Massachusetts, and a PAR collaboration with the Healthy Neighborhoods Research Consortium (HNS). I find that a significant proportion of affordable housing sales (23%) were flipped between 2008 and 2021 compared to only 7% of expensive housing sales, flipped properties generated higher financial gains on average than non-flipped sales, and homes that were later flipped were more likely to be bought by investors and more likely to be bought with cash. Furthermore, properties within neighborhoods that faced historical disinvestment are 12.2% more likely to be flipped than those not in these areas.

Through a partnership with resident researchers of the HNS consortium, we identified the biggest harms and benefits of house flipping hot spots within their neighborhoods and the potential policies targeted to increasing access to information, supporting home repairs, revising zoning laws, creating cooperative and land trusts with limited-equity covenants, and providing financial assistance to first-time homebuyers to prepare neighborhoods better to house flipping.

IRB Approvals

Each chapter of my dissertation received approval from MIT's Committee On the Use of Humans as Experimental Subjects (COUHES). Chapter 1 went through a comprehensive review with approval number #2110000488. Chapter 2 also went through a comprehensive review with approval number E-3391. Finally, Chapter 3 underwent a comprehensive review with approval number #2302000888.

Chapter II

The Effects of Localized Climate Shocks on Places and People

I Introduction

Extreme weather events, i.e., *climate shocks*, are becoming more frequent, severe, and expensive, placing them at the forefront of economic and social development. They affect society's macroeconomic structures, such as the real estate market, local and state government finances, and the insurance industry, while also impacting an individual's well-being, behavior, preferences for living locations, and business development. Natural disasters, including tornadoes which are the focus of this study, test the flexibility and resilience of economic and political institutions. They pose the question of how an area or a neighborhood will recover, regain its ability to mobilize resources and move from a situation of lesser wealth to greater wealth. How a place rebuilds (*or not*) reflects an area's economic outlook and social composition, and it shows how politics and policy succeed and fail in protecting the most vulnerable populations.

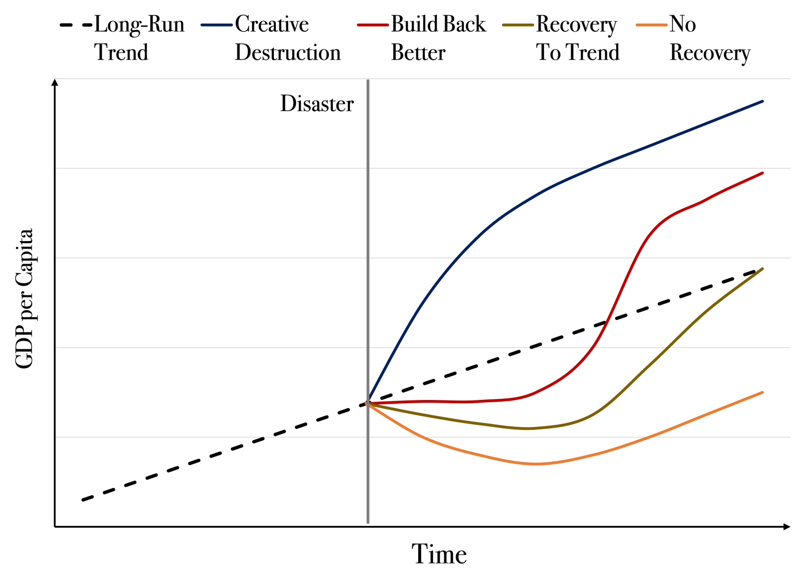
Given that *climate shocks* are predicted to intensify and affect new regions in the coming decades (Stott, 2016), it remains crucial to identify the heterogeneous impacts of these events on people, businesses, and places and to understand what conditions and policies maximize (*or not*) resiliency. This chapter leverages thousands of local quasi-natural experiments (i.e., tornadoes) to understand how people and places respond to *climate shocks* and how they recover from them. It focuses on answering (a) what are the impacts of an extreme geographic local-random weather event on a geographic area's built environment? (b) what are the impacts of extreme geographic local-random weather events on the well-being of its residents and the prosperity of its businesses? (c) what are the spillovers of local *climate shocks* on adjacent locations? and (d) what could policymakers and local governments do and *shouldn't do* to minimize post-disaster inequalities, foster resilience, and reduce ex-ante vulnerability?

To shed light on these questions, I leverage the exogenous direction and width of tornado paths in the United States between 1996 and 2019 to understand the longitudinal impact of an unpredictable natural hazard on places and people. I use rich granular data and Sun and Abraham (2021)'s generalized difference-in-differences event study estimator to quantify the impacts of local *climate shocks* on the real-estate sector, local governments, school outcomes, households, and businesses behavior.

I have the following set of results. First, following a tornado, properties in the regions affected by it sell at a discounted price relative to those barely touched by it. This effect persists for several years with increasing sales price discounts as time passes, suggesting a post-disaster no-recovery growth path. Second, the decrease in sale prices following a tornado coincides with an increase in the supply of listings, an increase in the number of properties foreclosed, and a decrease in the probability that an investor buys a property within the affected regions. Third, properties affected in the suburbs experience a continuous down-permanent adjustment in their sale prices after a tornado occurs, whereas those within the central business district experience considerable variation in their adjustment. In contrast, those in rural areas do not suffer a decline in prices, suggesting that tornadoes act as a risk adjustment to investors and homebuyers within the city. Fourth, neither allowing affected regions to access recovery funds released by a FEMA presidential declaration nor the intensity of a tornado affects the behavior of property sale prices within the affected regions. Fifth, following a tornado, localities experience an abrupt and permanent tax base decrease, affecting the amount of property taxes collected per property. Sixth, I find evidence that math, reading, and language test scores for grades 3 to 8 decreased for affected school districts following a tornado. Seventh, I find that tornadoes tend to displace people almost immediately. Eighth, following a tornado, there is an abrupt reduction in the percentage of small businesses within the affected regions. Ninth, tornadoes not only affect areas directly impacted by them, but they also generate spillover effects on adjacent ones.

This chapter builds on and contributes to the following pieces of literature. First, the literature on the economics of natural disasters remains inconclusive on the causal effects of environmental shocks on long-run economic growth. The four hypotheses shown in Figure II-1 have been proposed to describe the potential trajectories an economy follows after a *climate shock* (Hsiang and Jina, 2014)."Creative destruction" argues that disasters stimulate economies to grow faster because demand for goods and services rises as populations replace lost capital, inflowing international aid promotes growth, or environmental disruption fosters innovation (Skidmore and Toya, 2002). "Build back better" argues that while growth may suffer initially due to lost lives and productive capital, replacing lost assets with modern units has a positive net effect on long-run growth (Crespo Cuaresma, Hlouskova and Obersteiner, 2008; Hallegatte and Dumas, 2009). "Recovery to trend" argues that growth will slow but eventually return to its pre-disaster level. This happens because when capital and labor become scarce after a disaster, the marginal product of capital rises, encouraging people and capital to relocate to places where these scarcities exist. Strobl (2011) is an example of this. "No recovery" theory says climate shocks restrict growth by destroying productive capital or durable consumer goods (such as real estate), which are then replaced with capital that would otherwise go to productive investments. Because of the intensity of the immediate negative effect, no recovery occurs (Field et al., 2012).

Figure II-1: Potential GDP Growth Trajectories following a Climate Shock



Note: Figure replicated from Hsiang and Jina (2014).

In this context, recent work still finds puzzling results without a definite conclusion of which trajectory happens under which circumstances, primarily due to empirical challenges related to omitted variables bias and endogenous disaster location (Field et al., 2012; Hsiang and Jina, 2014). Given a recent call to conduct "more research on the long-term impacts of natural disasters (Botzen, Deschenes and Sanders, 2019)," my chapter contributes to this literature by exploiting a spatially and temporally unpredictable *climate shock* alongside granular data, which aids in solving endogeneity issues.

Second, the majority of disaster-related research has focused on national economic indicators (e.g., (Botzen, Deschenes and Sanders, 2019)) or aggregated spatial units, such as counties or block groups (e.g., Boustan et al. (2020); Raker (2020). The limitations in a data aggregation approach do not allow exploring the differential effects of climate shocks on places and people based on individual-level socioeconomic and house-level characteristics. This may lead to the risk of committing an ecological fallacy, especially when attempting to extrapolate aggregate results to individual circumstances. Given a set of ex-ante local and individual conditions, the effect and the recovery from an extreme weather event can differ between individuals and geographic areas. For example, economically advantaged populations move out after hurricanes and leave socially vulnerable groups stuck in place (Logan, Issar and Xu, 2016). In contrast, the impact on the built and business environment after the same disaster might invigorate (and change) or depress local economies (Akao and Sakamoto, 2018). Understanding whether extreme weather events are particularly dangerous for people and local geographies that have experienced high levels of social vulnerability or economic decline could help inform how post-recovery interventions could be deployed with an equity focus. Furthermore, by not following individual units across time, we cannot understand the impact of individuals staying in place or migrating to other destinations.

Third, my research takes advantage of an unpredictable extreme weather event, i.e., a tornado, at the spatial and time level, which allows me to deal with spatial endogeneity, sorting, and boundary issues within the disaster-related research literature (Fussell, Sastry and VanLandingham, 2010; Heckman and Smith, 1999; Raker, 2020). A key issue with identifying the variation in the impact and recovery from disasters is the idea that vulnerable populations tend to move into climate-vulnerable locations, making it challenging to isolate the direct effects of the natural disaster from the impact on living in a place where a disaster occurred (Arcaya, Raker and Waters, 2020). Given a tornado's unpredictability, individuals and potential home buyers cannot sort into places less likely to experience a tornado or leave places that will experience it. My research design allows for isolating the direct effects of the natural disaster from the impact on living in a place where a disaster occurred.

This chapter is organized in the following way. Section I describes the data, methodology, and empirical strategy. Section II presents the results for the real estate sector, local tax collection, school outcomes, household location, and business creation. Section IV discusses and concludes.

II Data and Empirical Strategy

This paper combines data from several sources to understand the impact of local shocks on the real estate sector, local tax collection, school outcomes, household location, and business creation through a generalized difference-in-differences event study estimator (Sun and Abraham, 2021).

II.I Data

The data sets used in this paper come from multiple sources: (a) tornado widths and paths provided by NOAA's Severe Weather GIS database, (b) multiple listing services, property transactions, and tax records from CoreLogic and Redfin, (c) household and business data from Axle, (d) school outcomes from Stanford's Educational Opportunity Project, and (e) FEMA's Presidential Disaster Declaration Database.

Tornadoes

Tornado polygons come from NOAA's Severe Weather GIS database containing tornado occurrences between 1996 and 2019.[1] Each tornado polygon contains (1) year, month, day, and time of occurrence; (2) Fujita-scale (0, 1, 2, 3, 4, or 5)[2]; (3) number of injuries and fatalities; (4) estimated property loss (in millions of dollars); (5) starting latitude/longitude and ending latitude/longitude; (6) length of the tornado in miles; and (7) width of the tornado in yards.

[1] NOAA's database began in 1950; however, I decided to use 1996 as a starting point given the introduction of the Doppler radar at the end of the 1980s, which considerably increased tornado spotting relative to previous years. Thus, reducing Type II errors at the expense of longitudinal quality. Furthermore, prior to 1996, NOAA's damage categorization was not available across events.

[2] Enhanced Fujita-scale after January 2007.

Real-Estate: Property Transactions, Multiple Listing Service, and Tax Records

Real-estate information comes from three sources. First, arm's length property *sale* transactions originated from county records and were extracted by Redfin from CoreLogic. Each transaction contains a unique property and transaction identifiers, sale date, sale price (in USD dollars), the type of buyer[3], and geographic and basic characteristics of the property, among other variables. Second, multiple listing service data comes from Redfin and covers those regions where the brokerage operates.[4]. Each listing contains unique listing and property identifiers, sale date, sale price, listing added and end date, listing price, number of bedrooms, year built of the property, approximate square feet of the property, number of bedrooms, whether the listing is new construction, geographic characteristics of the property, among other variables. Third, a yearly property tax history information between 2006 and 2020. Each data point represents fiscal year tax information for each available property. Each data point contains the yearly taxable land value, improvement values, and taxes due for each property over time.

FEMA's Presidential Disaster Declaration

Upon the request of a state's governor, the President of the United States can issue a major disaster declaration, which releases federal funds from FEMA in the aftermath of an extreme weather event (Raker, 2020). The FEMA Presidential Disaster Declaration Database is a centralized repository for tracking and documenting these declared disasters. Each data point contains a single FEMA Presidential Disaster Declaration with the type of incident, the type of declaration, the state where the incident occurred, and the dates when it occurred. For tornado incidences, a presidential disaster declaration can span multiple days. For these cases, every tornado occurring within the temporal and spatial time frame was considered part of a Presidential Disaster Declaration.

[3]Either an Institutional, Investor, or Home buyer

[4]For a full list of Redfin's market access, see the following link.

School Outcomes

School data comes from Stanford's Educational Opportunity Project from 2009 to 2018 at the geographic school district level in the United States. The data contains yearly geographic school district achievement estimates, long by grade-year-subject. Its estimates are scaled relative to national student-level grade- and subject-specific test-score distributions. It contains a measurement of how school cohorts perform across time.

Households and Businesses Panels

Household and business data come from Data Axle, formerly Infogroup, Historical Business, and Historical Residential files. The Historical Business data is a year-end calendar snapshot of local business data between 1997 and 2020 in the United States. It contains longitudinal business IDs, location information, latitude and longitude information of the business, industry codes, corporation hierarchy information, employment variables, sales, and year established, among other variables. The Historical Residential file contains a year-end snapshot of publicly available household information between 2006 and 2020. It has information such as longitudinal household IDs, location information, latitude, and longitude information of the household, residential length, real estate property characteristics of the home the household resides in, characteristics of the head of household, number of children in the household, renter or owner status, and estimated values of the household, such as household wealth, income, and property values, among others.

II.II Empirical Strategy

Treatment Assignment

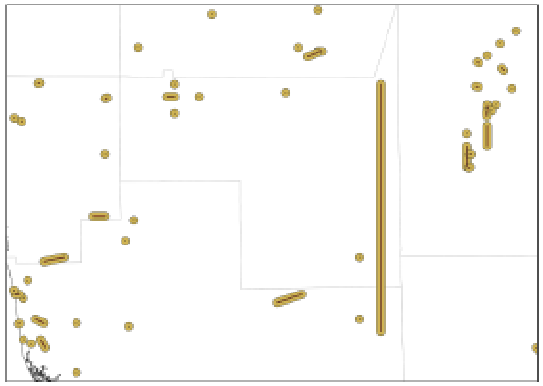
Tornadoes with positive property damage and width were chosen from NOAA's Severe Weather GIS database, representing 49.5% of all spotted tornadoes since 1996.[5] This approach diverges from previous studies focusing on the social and market consequences of extreme weather events. Including tornadoes with lower property damages or shorter paths allows me to evaluate whether smaller but more common hazards may influence dynamics (Howell and Elliott, 2019).

Within the raw database, each tornado path is presented as a line. I used a tornado's width as its buffer area to convert each path to a polygon. This new polygon represents the treatment area in my empirical analyses. The control region begins at the border of the treatment area, and its length is one kilometer. Relative to previous analyses estimating the impact of natural disasters on places and people, my research uses precise latitude and longitude coordinates of each unit of analysis to merge it with the treatment and control areas.[6] Figure II-2 provides an example of the treatment (in black) and control (in yellow) catchment zones of several tornadoes within a specific geographic zone. As one can notice, each tornado has a unique treatment and control area, allowing it to conduct within tornado analyses. In Subsection III.II, I present four other ways to identify treatment and control areas that test for spillover effects and more detailed impact zones.

[5]The year cut-off was chosen to consider the introduction of the Doppler radar at the end of the 1980s, which considerably increased tornado spotting relative to previous years. After 1996, NOAA's database also provides consistent property loss estimates across time. The positive cut-off allows the understanding of local-random shocks on people and places.

[6]Previous studies assign entire census tracts or counties to treated units, even though the natural disaster barely touched those areas. This strategy produces the following biases. First, the unit of analysis does not correspond with the locations in which the observation units are damaged, creating Type I and II errors. Second, it assumes that every unit within the chosen area is equally affected by natural disasters. Third, it severely amplifies the effect of small natural events by making them appear as if they affect large spaces.

Figure II-2: Example of Treatment and Control Catchment Areas of a Tornado Path
Treatment = Black, Control = Yellow



Spatial Merge

I do the following to assign each real-estate property, household, and business to a unique treatment area. For real estate information, I used each property's latitude and longitude at any point in time to spatially merge it to a treatment or control polygon. For each household and business, I used the household's or business's latitude and longitude at the time of the tornado occurrence to spatially merge each data point to a tornado treatment or control polygon. After getting the subset of properties, households, and businesses within a control or treatment area, I created a panel of each observation through time. For school districts, I follow an empirical strategy similar to previous studies, where a school district is considered treated if a tornado goes through it.

Target Estimand

I seek to identify the average treatment on the treated (ATT) at *any* of the post-treatment time points, for cohort, g, at event time, $e \equiv t - g$, which is defined as:

$$\mathrm{ATT}_{g,e} \equiv \mathbb{E}\left[Y^1_{i,t+e}(g) - Y^0_{i,t+e}(g) \mid A = 1, G_i = g\right] \tag{II.1}$$

Where [ ] represents the expected value and ( ) ( ) are the potential outcomes of outcome, , at event time, , for individual, , when the treatment, , was first administered. Note that ATT identifies the average treatment effect on the treated cohort, , at event time, . I am interested in the ATT across treated cohorts for a given event time, i.e., ATT , which I obtain as the weighted mean of each ATT :

ATT ATT Where: 1 = / 1 = (II.2)

Where, 1 = , represents the number of individuals in the cohort, , whereas, 1 = , is the number of individuals that received treatment at least at one point in time.

Estimator

To estimate the impact of localized climate shocks, i.e., tornadoes, on places and people, I use Sun and Abraham (2021)'s generalized difference-in-differences event study estimator. This strategy has as a unit of observation the property listing, property tax collection, school district, individual, or business during different periods, and the estimator takes the following empirical form for each yearly cohort:

= + + + + + + (II.3)

Where represents an outcome of interest for observation, , located in the catchment area, , during the period, . = 1 Cohort = is an event study dummy that is 1 if the unit is, , periods away from the treatment and 0 otherwise. Period = 1 is the baseline. and are the lowest and highest number of leads and lags surrounding the treatment period, respectively. is the total number of cohorts; in this analysis, I use the year a tornado hit a catchment area as the cohort. are characteristics of unit, , in catchment area, , during period, . , , and are years, tornado, and property fixed-effects, respectively. are the residuals. Standard errors are clustered at the tornado level. To go from to , I use the weight, , previously defined for event time, .

The causal identifying assumption is that outcomes in areas impacted by the tornado would have continued along the same trajectory without exposure within catchment areas. To formally test this assumption, I conduct a joint test of the null hypothesis: $= \quad = \quad = 0$.

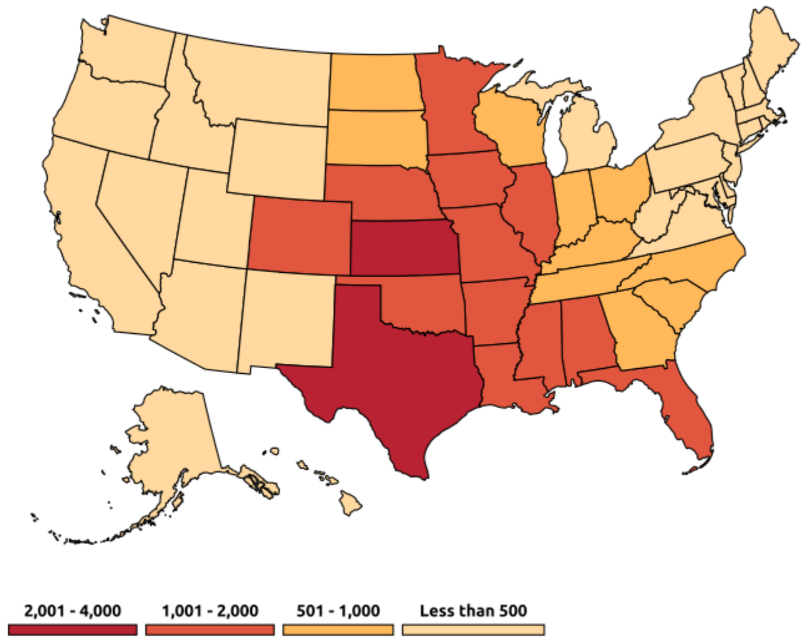
III Results

III.I Descriptive Statistics of Tornadoes

Tornado instances frequently occur in the United States. From 1996 to 2019, there were 21 174 tornadoes with positive property damage passing through residential property. On average, there were 882 instances per year (max: 1 384, min: 620) and 74 per month (max: 180, min: 28). The months of May (180), April (146), and June (121) had, on average, the highest number of tornadoes per month, while January (37), February (35), and December (28) had the lowest. In these tornado instances, there were, on average, 1 2 (min: 0, max: 1 500) and 0 8 (min: 0, max: 158) injuries and deaths, in that order. Finally, a tornado's average length and width were 6 1 kilometers and 139 meters, respectively.

As seen in Figure II-3, the spatial distribution of tornadoes across US states from 1996 to 2019 shows varying occurrence levels across states. A concentration of tornado activity is observed in the central part of the country, also known as Tornado Alley, encompassing states such as Texas, Oklahoma, Kansas, Nebraska, and Iowa. Additionally, states in the Southeast, such as Alabama, Mississippi, Tennessee, and Florida, exhibited high tornado occurrences between 1996 and 2019.

Figure II-3: Cumulative Number of Tornadoes by State between 1996 and 2019



III.II Real-Estate Sector

Pre-Trends Analysis

The real estate sector analysis presented here is based on 1,996,930 listings, which encompass both *treated* and *control* areas. Tables A.1 and A.2 in the Appendix provide estimates of the average difference in selected outcomes before a tornado occurred between treatment and control areas. I found no statistically significant differences in sale prices, property size, year built, bathroom count, walk score, transit score, and bike score between treatment and control areas before a tornado. This result suggests that treated and control areas followed similar trends before the tornado occurred, providing evidence for the parallel trends assumption. As additional evidence for the parallel trends assumption, Figure II-4 provides estimates of the pre-tornado (i.e., from to 2) 's of equation II.3. Here, I cannot reject the null hypothesis of a joint test of coefficients = = = 0 (p-value = 0 61).

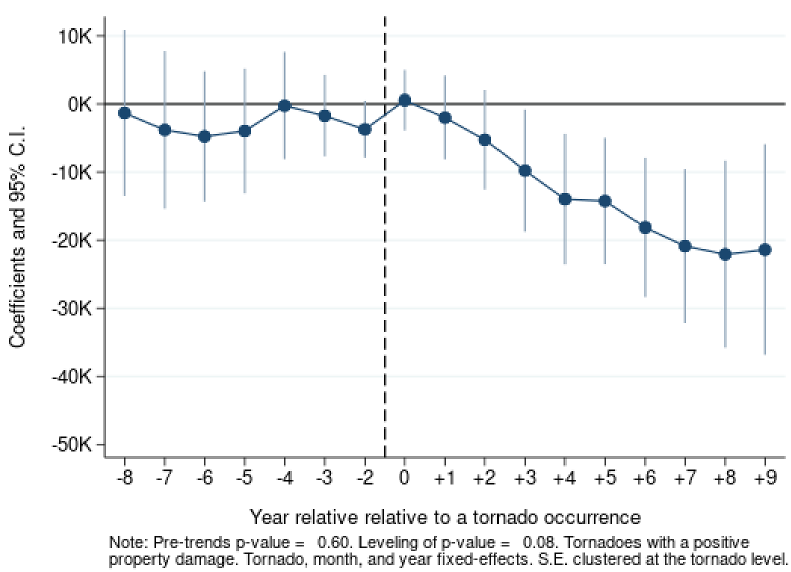
Average Treatment Effects

Using estimator II.3, I find evidence that properties within the tornado path (i.e., the *treated* regions) sell at a discounted price relative to the ones surrounding them (i.e., the *controlled* regions) during the aftermath of the event. Figure II-4 shows that two years after a tornado occurs, a property within the affected region sells for $8 055 less (p-value = 0 025) relative to properties within control regions.

As the years passed, this discount kept increasing, reaching almost a sale price of $23 412 (p-value = 0 002) less during year eight. In this analysis, I can reject the leveling of the coefficients post-treatment with a p-value of 0 01. This result provides evidence of a "no recovery" growth trajectory in which a tornado affects the real-estate sector in the short and long term.

Figure II-4: Impact of a Tornado on Sale Prices, by Years Since a Tornado Occurs

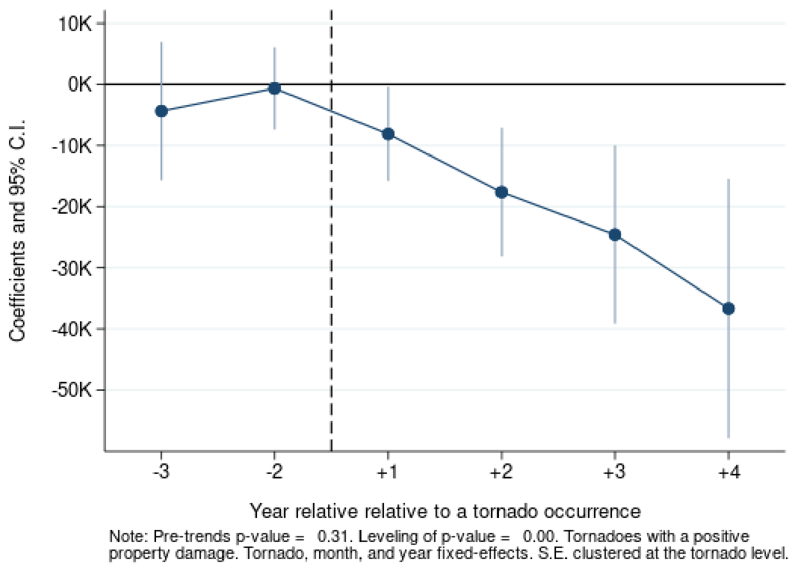
Years since Occurrence, US Dollars ($1,000)



On average, it takes 4 3 years (s.d. = 2 1) for a property to be sold after a tornado occurs. This explains why the coefficients of Figure II-4 become statistically significant until the third year. Figure II-5 orders the event time study as the transaction number within a property relative to the tornado occurrence. I find similar results to the ones from the previous analysis. On average, the first transaction after a tornado occurs is sold for $9 837 (p-value = 0 008) relative to control properties within a tornado area, whereas the fourth transaction is sold for $30 347 (p-value 0 001)

Figure II-5: Impact of a Tornado on Sale Prices, by Transactions Since a Tornado Occurs

Number of Transactions since Occurrence, US Dollars ($1,000)



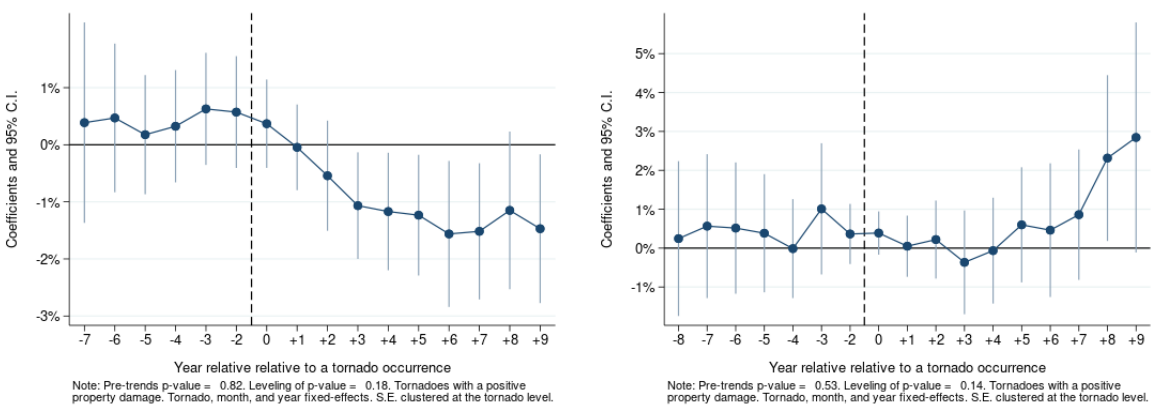
The decrease in sale prices presented above coincides with changes in the demand and supply of housing within catchment areas. First, the composition of buyers within catchment areas changed after a tornado occurrence. As seen in Figure II-6 (a), the probability that an institutional investor bought a property declines. On average, two years after a tornado occurs, an affected property has approximately 1% (p-value = 0 045) lower probability of being purchased by an institutional investor. As time passes, this effect keeps declining, reaching almost 2% (p-value = 0 001) after six years. This effect happens despite an increase in the role of institutional investors since the 80s within these catchment areas. Figure A.1 in the Appendix shows that in 2021, the probability of a housing transaction being bought by an institutional investor grew by 15 3%, relative to the year 1986.

Second, the percentage of bank-owned property transactions that were foreclosed and repossessed increases steeply years after a tornado occurs in affected areas (see Figure II-6 (b)). Seven years after a tornado occurs, the probability that a bank sold a housing transaction is 1% higher in affected areas, reaching almost 2 5% nine years after.

Figure II-6: Impact of a Tornado on the Probability a Property was Bought by an Institutional Investor and Sold by a Bank

(a) Bought by an Institutional Investor

(b) Sold by a Bank

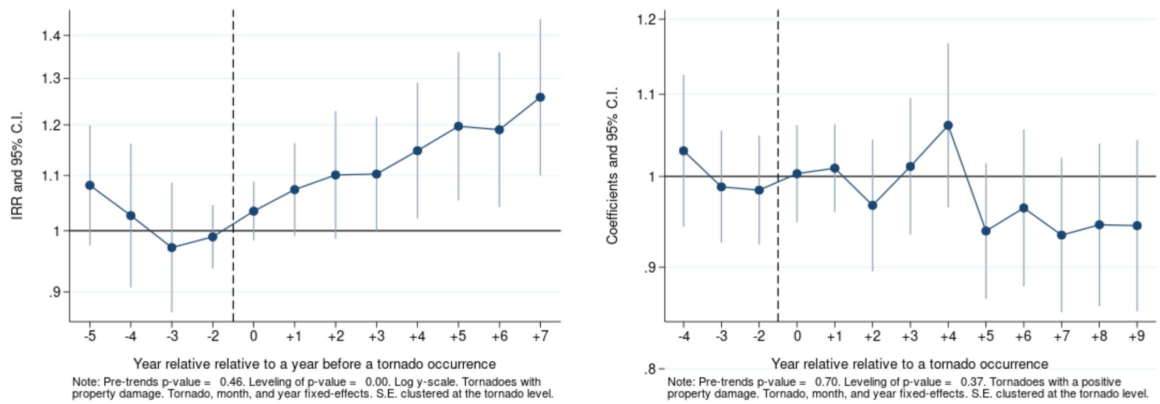


Third, potentially helping explain the decrease in sale prices previously shown, I found that the affected regions have a higher rate of listings (Figure II-7 (a)) relative to control regions. The coefficients begin to be statistically significant in year number four (IRR = 1 15, p-value = 032) and reach a maximum in year number nine (IRR = 1 40, p-value 001). In other words, the ratio of new listings within a catchment area between affected and not affected areas broadened since the tornado occurred.

Finally, the time a property stays on the market remains the same between affected and non-affected areas. As shown in Figure II-7 (b), the difference in the number of days a listing stays on the market before being sold is not statistically different between areas impacted by a tornado and those that are not.

Figure II-7: Impact of a Tornado on the Number of Listings and Days on the Market

(a) Number of Listings (b) Days on the Market



FEMA Presidential Disaster Declarations and Dosage Effects. FEMA Presidential Disaster Declarations may play a vital role in accelerating the recovery of areas affected by disasters. With access to FEMA funds released because of a Presidential Declaration, areas impacted by a tornado may have the opportunity to bounce back more swiftly compared to those without such resources. To explore the effect of these declarations on the subsequent recovery of an impacted region, I multiply in estimator II.3 by a dummy variable, taking a value of one if the tornado received a Presidential Disaster Declaration and zero otherwise. The evidence presented in Figure A.2 in the Appendix indicates that there is no statistically significant difference in sale prices of listings between areas that received FEMA funding from a Presidential Disaster Declaration and areas that did not receive such funding.

I also explore the dosage effects of a tornado on the local real estate market.[7] I employ a similar empirical approach as in the FEMA's Disaster Declaration analysis, but in this case, I use a categorical variable to capture the impact of each tornado's magnitude. To incorporate this factor, I multiply in estimator II.3 by a categorical variable representing the intensity level of the tornado according to the Fujita-Scale. As seen

[7] Dosage effects, in the context of studying the impacts of *climate shocks*, refer to the relationship between the severity or magnitude of a *climate shocks* and its subsequent impacts. These impacts can vary depending on the dosage or intensity of the disaster, which can be measured by factors such as the scale, magnitude, duration, and frequency of the *climate shock*.

in Figures A.3 (a) to (d), the analysis shows that there is no statistically significant difference between the magnitude of each tornado and the behavior of sale prices of local real estate within the affected regions, providing evidence against the dosage effect hypothesis.

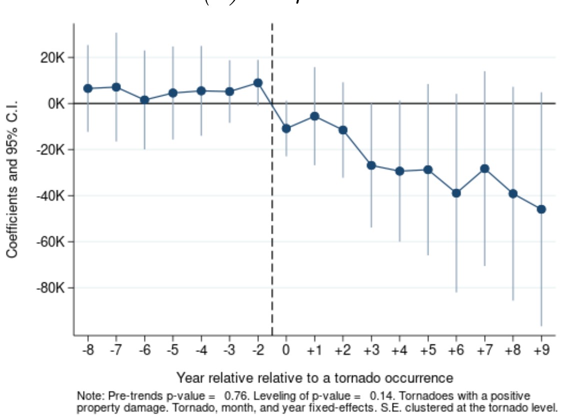
Spatial Effects

To understand whether the spatial distribution of economic or demographic units mediates the impact of a tornado, Figure II-8 presents estimates of the impact of a tornado on sale prices, stratified by the distance of the property to the closest central business district (CBD). The effects of a tornado on properties happen primarily outside of the central business district (Figure II-8 (b)). Properties affected by a tornado located within 4 kilometers of the central business district have a negative impact but quickly go to sale prices before the tornado occurs (Figure II-8 (a)). In contrast, properties outside the central business district do not have a sale penalty after a tornado, suggesting that these properties are sold at the price of the land in these locations (Figure II-8 (c)).

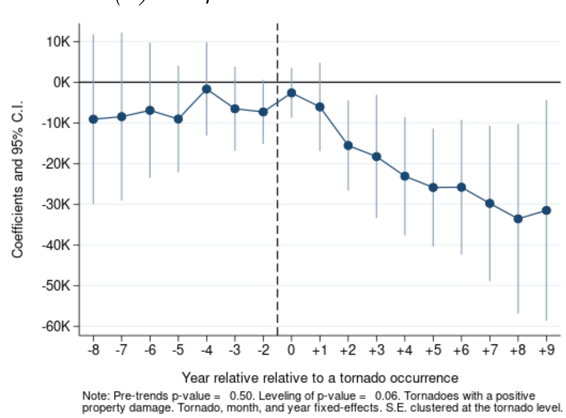
Figure II-8: Impact of a tornado on sale prices, by distance to the nearest CBD

US Dollars ($1,000)

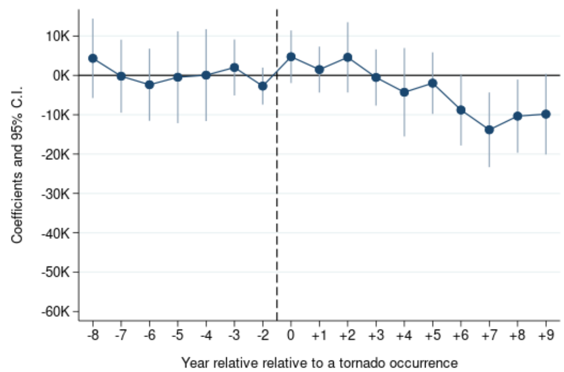
(a) 4 km



(b) 4 km and 18 km



(c) 18 km



Note: Pre-trends p-value = 0.60. Leveling of p-value = 0.21. Tornadoes with a positive property damage. Tornado, month, and year fixed-effects. S.E. clustered at the tornado level.

Spillover and Direct Effects

Tornadoes can generate spillover effects outside the boundaries of the directly affected regions. For example, the destruction caused by these *climate shocks* may not only result in the loss of housing within the directly affected areas but also could trigger a decrease in property values in nearby regions, as the perception of increased risk and lower desirability of the region could potentially reduce demand for housing. Additionally, tornado-damaged properties may experience significant value depreciation, indirectly impacting surrounding areas. Potential buyers may be hesitant to pay higher prices for homes with a higher risk of tornadoes or a history of severe damage that might indirectly affect insurance costs. Moreover, the disruption of infrastructure, including utilities and transportation networks, can impede housing availability and accessibility, creating a negative premium for adjacent regions. Furthermore, decreased school quality, amenities, or services resulting from tornado damage can lead to negative spillovers in neighboring areas.

To test for potential spillovers of tornadoes in nearby regions, I define the treatment and control areas in the following new ways:

1. *Spillover Effects*: In this case, treated regions are defined as areas falling within the width of the tornado path. Conversely, control regions are designated as those

areas located at a distance equivalent to one tornado width away from the treated region's border.

2. *Direct Effects*: For this scenario, treated regions are those areas encompassed within 80% of the tornado path width.[8]. Control regions, as used consistently throughout this study, include all areas except for those control regions identified as spillover areas under the *spillover effects* scenario.

3. *Spillover + Direct Effects*: Here, treated regions include areas that fall within the treated regions defined under the *direct effects* scenario, as well as the spillover control regions outlined in the *spillover effects* scenario. The control regions retain the same definition as in the *direct effects* scenario.

4. *Both Spillover and Direct Effects*: Lastly, for this scenario, the treatment has two arms, which are the same as those previously delineated under the *spillover* and *direct effects* scenarios. The control regions remain identical to those defined in the *direct effects* scenario.

In Figure II-9 (a), one notices that tornadoes have spillover effects on adjacent areas, as the coefficients are not statistically significant from zero; thus, I'm not able to reject the null hypothesis regarding the difference in sale prices between treated and spillover areas during the aftermath of a tornado. This suggests that both treated and spillover regions experience a negative decline in sale prices, indicating that tornadoes impact not only treated areas but also the surrounding regions.

Moreover, Figure II-9 (b) illustrates the results when refining the control areas by excluding spillover areas. In this case, the decline in sale prices is even steeper compared

[8]Tornadoes do not typically follow a strictly linear path creating uncertainty about the center (or non-circularity) of a tornado path. Their trajectories can be influenced by various environmental factors, including the movement and structure of the thunderstorm from which they are generated, the local topography, and other meteorological conditions (Brooks, Doswell III and Kay, 2003). I use an 80% width to balance uncertainty and a non-linear path when identifying direct hits.

to the original analysis. This finding suggests that the initial catchment areas' intention-to-treat analysis provided a conservative estimate, serving as a lower bound for the true effect of tornadoes on the sale prices.

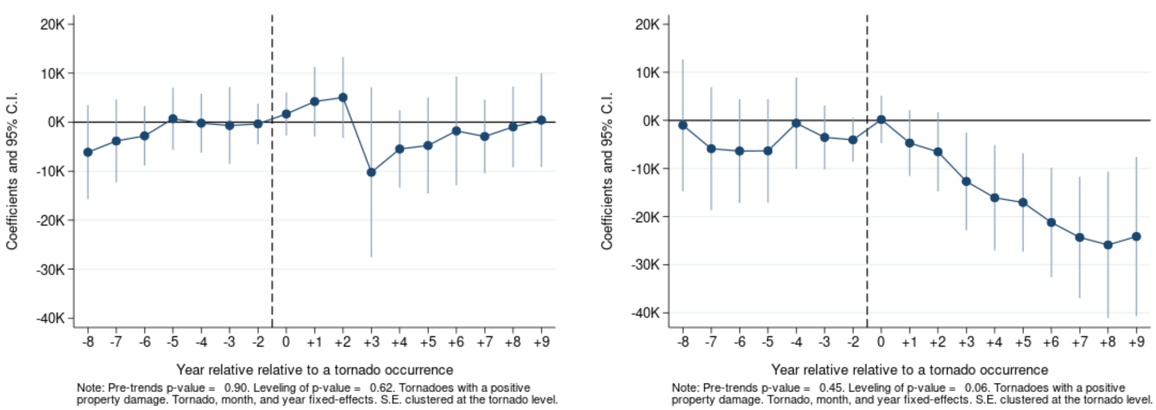
Figure II-9 (c) demonstrates that even when spillover areas are included within the treatment group, negative effects on sale prices are observed after a tornado affects an area. This further reinforces the notion that tornadoes have significant adverse impacts on sale prices in the affected regions, extending to both treated and spillover areas.

Finally, Figure II-9 (d) shows in the same graph the average treatment effect for both the direct (depicted in red) and spillover (illustrated in blue) treatment groups relative to the control group. After four years, the point estimates from the *direct* treatment arm consistently appear lower than those from the *spillover* arm. However, one cannot reject the null hypothesis between each treatment arm.

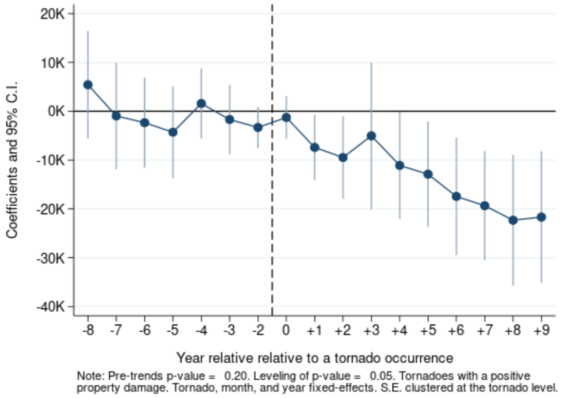
Figure II-9: Impact of a tornado on sale prices, by Treatment Design

US Dollars ($1,000)

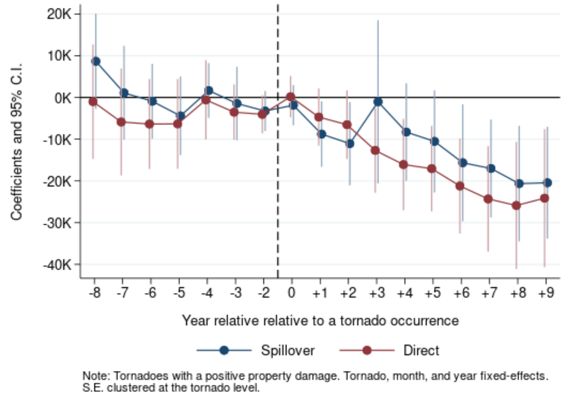
(a) Spillover Effects | *(b) Direct Effects*



(c) Spillover + Direct Effects



(d) Both Spillover and Direct Effects



I find similar results as the ones shown above when exploring the impact of a tornado on the probability of a property being purchased by an institutional investor (as depicted in Figures A.4 (a) to (d)), the likelihood of a property being sold by a bank (shown in Figures A.5 (a) to (d)), the sale prices by distance to the nearest Central Business District for distances ranging from 4 km to 18 km (illustrated in Figures A.9 (a) to (d)), and for distances greater than or equal to 18 km (seen in Figures A.10 (a) to (c)), as well as the tax due per property (highlighted in Figures A.11 (a) to (d)). That is, I found evidence of spillover effects impacting adjacent areas, with the effects becoming more pronounced upon excluding spillover areas. Moreover, these negative impacts persist even when considering spillover areas as part of the treatment.

However, upon further refining the treatment group, I found that the sale prices by distance to the nearest Central Business District for properties located less than 4 km away (as portrayed in Figure A.8 (b)) are not only statistically significant but also substantial. Lastly, a comparison of the number of listings between treated and spillover areas, shown in Figure A.6 (a), revealed a higher count in the former. This discrepancy in listing numbers might explain the more pronounced decline in sale prices when spillover areas are excluded from the analysis, as visualized in Figure II-9 (b). The increased volume of listings in treated areas could induce greater competition among home sellers, thereby precipitating a steeper plunge in sale prices.

III.III Property Taxes

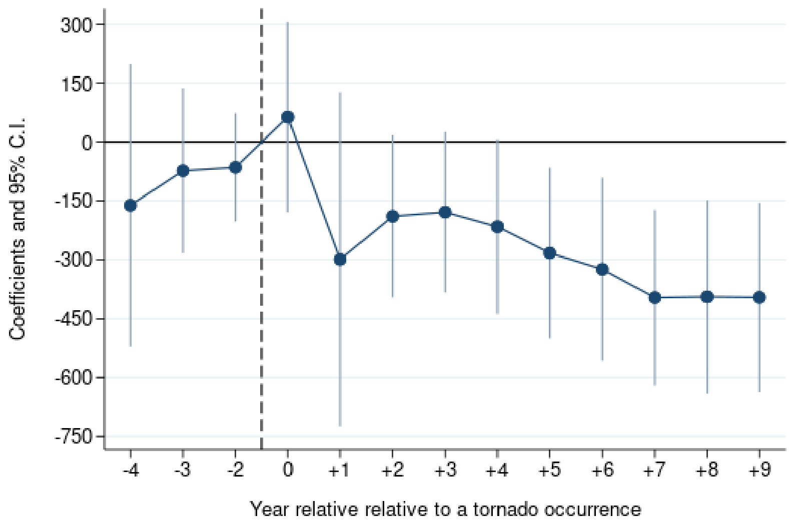
Property taxes play a crucial role in financing local communities by providing a stable source of revenue for essential public services such as schools, infrastructure development, and emergency services that could potentially be impacted by a *climate shock*. If an area is affected by a natural disaster and market values decrease, tax collection should decrease since it is linked to a property's assessed value. At the extreme, if a property is destroyed, the owner would only need to pay taxes on the value of the land.

A loss in property taxes can severely impact a communities' well-being and development. As previously described, local governments provide services primarily financed via local taxes. Elementary and secondary education is a far larger share of direct local government spending among these services. In 2019, 40% of direct local government spending went to elementary and secondary education (Boddupalli and Rueben, 2021).

To pay for these services, local taxes heavily depend on property taxes. In 2019, 30% of local's government total revenue (Boddupalli and Rueben, 2021) and an abrupt decrease in tax collection could leave affected localities vulnerable. A persistent decrease in tax collection can have far-reaching implications for the local development of towns, potentially resulting in adverse spillover effects on unaffected properties. This negative feedback loop offers a plausible explanation for the decline of certain areas and a subsequent decrease in intergenerational mobility following a localized climate shock. Such a decline in tax revenues can hinder the necessary investments in critical infrastructure, public services, and community development, perpetuating a cycle of economic decline and impeding upward mobility for future generations.

Property Tax Loss. As seen in Figure II-11, I find a permanent and abrupt reduction in tax collection within affected areas during the aftermath of a tornado. Specifically, on average, taxes due per property in affected areas drop ($300) a year after a tornado occurs. The impact of a tornado on tax collection increases as time passes, and it reaches almost an average decrease of $400 (p-value = 002) per property nine years later.

Figure II-11: Impact of a Tornado on Taxes Due Per Property
US Dollars



Note: Tornadoes with a positive property damage. Tornado and year fixed-effects.
Pre-trends p-value = 0.02. Leveling of p-value = 0.00. S.E. clustered at the tornado level.

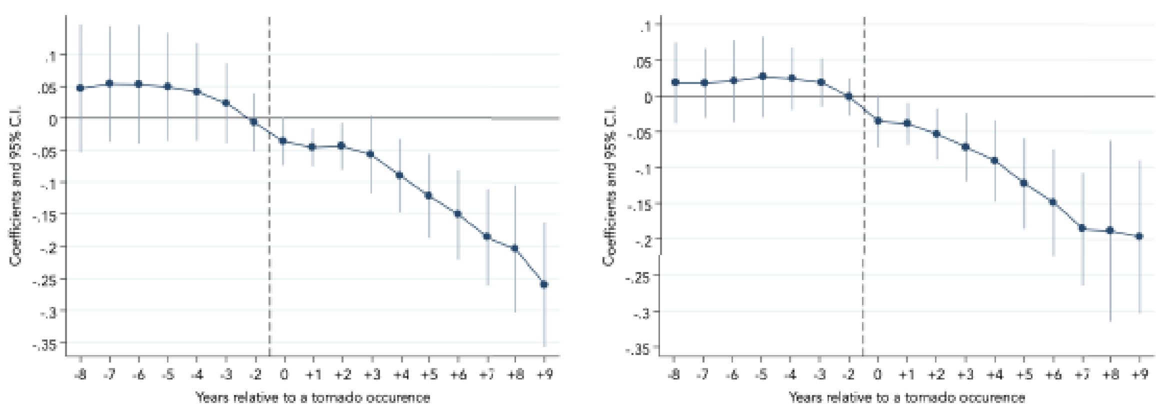
III.IV School Outcomes

Tornadoes may have broader impacts beyond the real estate and local tax collection, such as on the school outcomes of affected regions immediately in the following way. The aftermath of a tornado could, for example, affect the learning environment, generate emotional distress, hinder displacement, and decrease school resources and funding. These challenges could affect the academic performance, the test scores, and the educational experience of students. Additionally, negative school outcomes can create a ripple effect, leading to long-term consequences for students' educational attainment, career prospects, and socioeconomic well-being. Lower test scores can impact college admissions, scholarships, and future employment, perpetuating a cycle of disadvantage across generations.

As seen in Figures II-12 (a) and (b), there is a decline in math, reading, language, and art test scores over time for students in grades 3 to 8 within tornado-affected school districts compared to those in districts unaffected by tornadoes. After a tornado impacts a school district, the average test scores show a statistically significant decline over time. One year following the tornado, there was an average decrease of -.05 (p-value = .001) in math scores and -.04 (p-value = .001) in reading, language, and art scores. These average test scores continue to decline in subsequent years. By the sixth year, math scores are, on average, -.15 lower in tornado-affected districts than non-affected areas, with the largest difference observed nine years later, where scores are -.26 lower. Reading, language, and art test scores experienced a significant decline in tornado-affected school districts. These scores were found to be, on average, -.15 lower, reaching their lowest point of -.19 after nine years.

Figure II-12: Impact of a Tornado on the Change in Test Scores

(a) Math (b) Reading, Language, and Art



Note: Estimates include school district area, cohort, and year fixed-effects. Test Scores for students in grades 3 - 8.

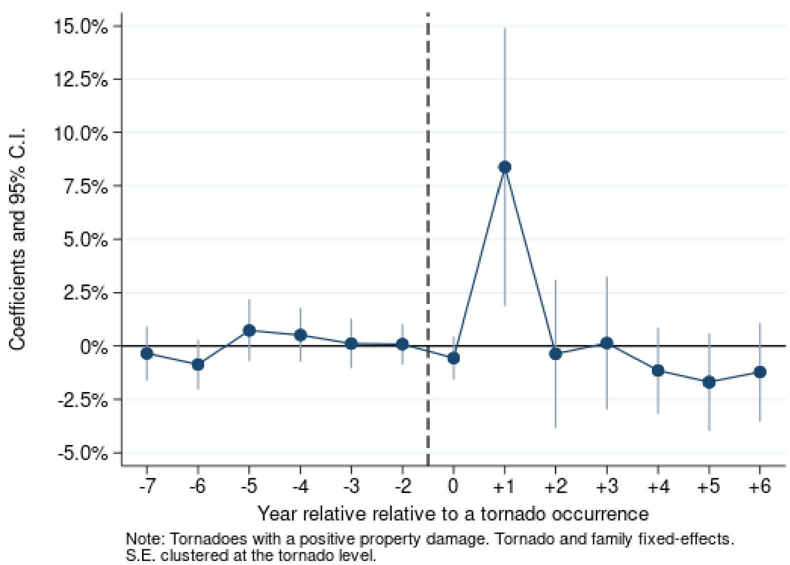
III.V Households and Businesses

The occurrence of a tornado has the potential to cause displacement among affected residents, leading to temporary or even permanent relocation. The destructive nature of tornadoes can render homes uninhabitable, necessitating the evacuation or displacement of individuals and households as they seek alternative housing arrangements. In my analysis, I found that households living within affected areas by a tornado are more likely to move to new zip codes the year after the extreme weather event occurred. That is, one year after a tornado occurred, people living within an area impacted by a tornado have a +8 8% (p-value = 01) higher probability of living in a different zip code than a year before relative to barely touched regions and the year before a tornado happened.

I only found a difference in the change in the probability of moving to a new zip code between treatment and control areas in the first year. This result suggests that affected households move at a higher rate once, and their change in the probability of moving remains constant after the first move between people affected and not barely affected by tornadoes.

Figure II-13: Impact of a Tornado on the Locations of Households

Probability of a Household Living in Different Zip Code relative to year, 1

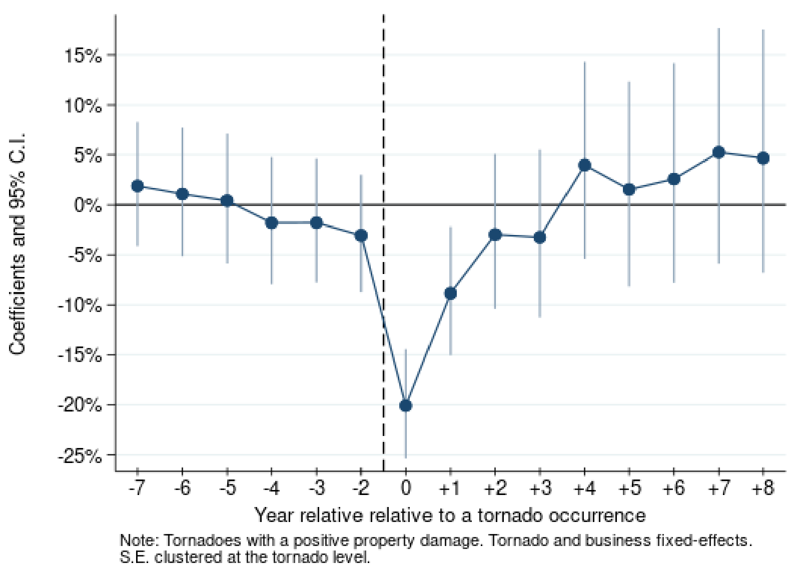


Localized *climate shocks* could also affect the local economy by creating physical damage, disruption of operations, and significant financial losses. The impacts could be larger for small businesses, given their limited resources, lower access to insurance markets, and less robust contingency plans. I find that the number of small businesses (i.e., businesses with 1 or 2 employees) in areas affected by tornadoes decreased following a tornado (Figure II-14) and went back to baseline after three years.

Following a tornado, there was a decline in the number of these businesses, with the affected regions experiencing a reduction of -21% (p-value = 0001) relative to the barely touched regions and the year preceding the tornado. One year after the tornado, these areas still showed a significant decrease of -8.9% (p-value = 01) in small businesses. However, the difference between affected and barely affected areas was no longer statistically significant after two years, suggesting a recovery to trend scenario for the local economy. I didn't find an impact on businesses with more than 2 employees.

Figure II-14: Impact of a Tornado on the Number of Small Businesses

Change in the Number of Small Businesses



IV Conclusion

In this paper, I show how localized extreme weather events, specifically tornadoes in the United States between 1996 and 2019, impact various aspects of the affected regions. By utilizing the exogenous direction and width of tornado paths, I compare the treated areas (those directly affected) with control areas (the regions on the tornado periphery) to analyze their effects on the housing sector, the fiscal capacity of towns, its school system, household migration, and the overall economic capacity.

Following a tornado, properties in the affected regions sell at a discounted price relative to those areas barely touched by it. This effect persists for several years with increasing sales price discounts as time passes, suggesting a post-disaster no-recovery growth path. Second, the decrease in sale prices following a tornado coincides with an increase in the supply of listings, an increase in the number of properties foreclosed, and a decrease in the probability that an investor buys a property within the affected regions. Third, properties affected in the suburbs experience a continuous down-permanent adjustment in their sale prices after a tornado occurs, whereas those within the central business district experience considerable variation in their adjustment. In contrast, those in rural areas do not suffer a decline in prices, suggesting that tornadoes act as a risk adjustment to investors and homebuyers within the city. Fourth, neither allowing affected regions to access recovery funds released by a FEMA presidential declaration nor the intensity of a tornado affects the behavior of property sale prices within the affected regions. Fifth, following a tornado, localities experience an abrupt and permanent tax base decrease, affecting the amount of property taxes collected per property. Sixth, I find evidence that math, reading, and language test scores for grades 3 to 8 decreased for affected school districts. Seventh, I find that tornadoes tend to displace people almost immediately and abruptly reduce the percentage of small businesses in the affected regions. Finally, tornadoes not only affect areas directly impacted by them, but they also generate spillover effects on adjacent areas.

The results of this paper suggest that places impacted by a localized shock have an immediate adverse impact and, generally, follow a no-recovery path, making it difficult for people to live in the affected areas and for businesses to operate, which leads to a decrease in property values and a decline in the demand for real estate in the affected region. A no-recovery path brings severe problems at the local level since communities rely on local taxes to fund public services. Low-income populations rely on public services at a higher rate relative to higher-income populations; shocks of this type may amplify inequality and reduce economic mobility. During the aftermath of localized shock, inequality plays a key role in the capacity of communities to bounce back. Since lower-income and marginalized communities are often located in areas that are more vulnerable to the destructive effects of tornadoes, such as low-lying areas or areas with older and less sturdy properties, these communities are more likely to experience the devastating effects of tornadoes or be more likely to experience lasting consequences of a localized shock. Tornadoes can also exacerbate existing inequalities. For example, lower-income and marginalized communities may have less access to resources and support for recovery, such as financial or housing assistance. This can make it more difficult for them to rebuild and recover from the disaster, leading to further disparities. Tornadoes can also have a long-term impact on inequality. The destruction caused by these storms can disrupt economic activity and lead to a decline in property values, making it difficult for residents to recover financially. This can perpetuate existing economic disparities and make it difficult for lower-income and marginalized communities to access economic mobility and growth opportunities.

Shock stabilizers or public policies designed to reduce the impacts of non-trivial shocks should be encouraged and demanded within local populations. Shock stabilizers, in the form of economic policies and measures designed to help communities recover from the effects of extreme weather events, could reduce inequality during the aftermath of a catastrophe. These policies and actions are intended to help cushion the impact of the disaster on the economy and to provide support to individuals and families who have been affected. Some examples of shock stabilizers that may be used in the aftermath of an extreme weather event include (a) unemployment insurance that provides temporary income support to individuals who have lost their jobs due to the disaster; (b) temporary housing assistance, which gives temporary housing for individuals and families who the disaster has displaced; (c) financial assistance that supports individuals and families who have suffered losses as a result of the disaster; and (d) public works programs that provide temporary employment for individuals who have lost their jobs as a result of the disaster.

Overall, the role of shock stabilizers is to provide support to individuals and families affected by extreme weather events and to help cushion the impact of the disaster on the economy. These policies and measures could help communities recover and rebuild after a disaster. They should be triggered automatically and shouldn't depend on political cycles or emergency declarations to be activated. Shock stabilizers could help impacted areas regain their ability to mobilize resources and move from a situation of lesser wealth to greater wealth. As expressed in the introduction, how a place rebuilds (*or not*) reflects an area's economic outlook and social composition. This process unveils the effectiveness, or lack thereof, of political maneuvering and policy implementations in safeguarding the most vulnerable populations. Introducing shock stabilizers could provide a significant defensive measure to bolster protection for these vulnerable groups.

Chapter III

Eviction Moratoria Expiration and COVID-19 Infection Risk

Joint work with *Mariana Arcaya, ScD, MCP* and *Atheendar Venkataramani, MD, PhD*

I Introduction

On September 4, 2020, the US Centers for Disease Control and Prevention (CDC) enacted a national eviction moratorium because "the evictions of tenants could be detrimental to public health control measures to slow the spread of the virus that causes COVID-19" (Centers for Disease Control and Prevention, 2020). The moratorium came at a time when an estimated 47.0% of individuals in renter-occupied housing were behind on their payments and were likely to leave their homes due to eviction, sequalae of the United States' long-standing housing affordability crisis and the COVID-19 pandemic's impact on employment and income (US Census Bureau, 2020*b*; Benfer et al., 2021).

A growing body of evidence suggests that eviction activity may be associated with increased COVID-19 infection rates. For example, a study using ecologic data on COVID-19 infection rates and the timing of state-level eviction bans found that COVID-19 rates increased after eviction moratoria expired (Leifheit et al., 2020). Other investigations using simulations have since found that households experienced an increased risk of infection not just due to personal experiences but also due to spillover from the transmission processes amplified by community evictions (Nande et al., 2021).

However, public health surveillance data limitations do not allow for exploring differential policy effects based on individual-level health and socioeconomic characteristics. Understanding whether expiring eviction moratoria are particularly dangerous for people and local geographies that have already experienced disproportionate effects of the pandemic, including individuals with preexisting health problems and low-income communities, could help to inform how nonpharmaceutical interventions are deployed with an equity focus. For example, shelter-in-place orders, which protect professional class workers but not essential workers from occupational exposures, likely have different distributional impacts than eviction moratoria, which we expect to disproportionately protect lower-income and rent-burdened populations and places.

We used detailed healthcare claims data from a large database in the United States to conduct what we believe to be the first individual-level analysis of how eviction policy affects the hazard of a COVID-19 diagnosis within health and neighborhood-level socioeconomic strata. We used a difference-in-differences research strategy to compare changes in the risk of being diagnosed with COVID-19 before and after lifting state-level eviction moratoria vs. the same changes in states that maintained these moratoria. We also assessed how associations between eviction moratoria and the risk of COVID-19 diagnosis varied by an individual's Charlson Comorbidity Index (CCI) score as well as by zip code-level poverty and rent burden prevalence to test the hypotheses that (1) individuals with poorer baseline health, as measured by the CCI, will experience a higher risk of infection after moratoria are allowed to expire because baseline health status and eviction risk are both socially patterned and (2) individuals in low-income and rent-burdened communities will be at heightened risk of infection after expiring moratoria due to higher risk of exposure to eviction-related COVID-19 transmission driven by local evictions and subsequent

II Data and Empirical Strategy

Our work combines data from the OptumLabs® Data Warehouse (OLDW) coupled with a novel survival difference-in-differences estimator.

II.I Data

We used deidentified administrative claims data from the OptumLabs® Data Warehouse (OLDW), which includes medical claims and enrollment records for individuals with commercial insurance and Medicare Advantage (MA) but does not include those with Medicare fee-for-service or Medicaid. The database contains health information on nearly 200 million enrollees across the United States, representing a diverse mixture of ages, ethnicities, and geographical regions across the United States (OptumLabs, 2020). The Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects exempted this study from review and the requirement for informed consent because it involved private de-identified information. This study adheres to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline for cohort studies.

Study Population

Our study cohort, organized as an individual-weekly panel (n = 254,847), included all individuals with commercial insurance and MA who (1) lived in a state in which an eviction moratorium was issued and (2) were diagnosed with COVID-19 during the period between the week the state first issued its eviction moratorium and the week the CDC issued the nationwide eviction moratorium (Raifman et al., 2020). Our primary analytic sample (i.e., balanced sample) also included a control group comprising an equal number of randomly selected individuals who were not diagnosed with COVID-19 in the same time period and states. We focused on an analytic sample that contained all individuals with a COVID-19 diagnosis to increase the statistical power to detect differences in the association of the eviction moratorium policy with COVID-19 diagnosis by stratifying variables.

II.II Empirical Strategy

Outcome, Exposure, and Covariates

Our primary outcome measure was a binary variable that varied by week, indicating whether the individual was diagnosed for the first time in that week with the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) code U07.1. Our exposure variable enabled an event-time study, with time centered on the week a state's court, governor, or legislature lifted its eviction moratorium for the first time (Raifman et al., 2020).

We included the following covariates: the number of weeks that had passed since the issuance of a state mask mandate, a stay-at-home or shelter-in-place order, the closure of schools, the state began lifting business restrictions, and the reopening of movie theaters (Raifman et al., 2020); weekly county-level COVID-19 cases lagged by two weeks (Dong, Du and Gardner, 2020); weekly state-level COVID-19 tests lagged by 2 weeks (Dong, Du and Gardner, 2020); zip code-level poverty rate (US Census Bureau, 2020*a*); week and state fixed effects; an individual's sex, age (centered at 65 years), type of insurance (commercial or Medicare Advantage), and latest industry of employment; and whether the individual had a Z code, i.e., a diagnosis of problems related to unemployment (ICD-10 code, Z56), problems related to housing and economic circumstances (ICD-10 code Z59), or problems related to bereavement (ICD-10 code, Z64.4) before 2020. We included an individual's CCI score as a baseline measure of global comorbidity before the pandemic and the study period began (Quan et al., 2011). We used an individual's available claims history from 2017 to 2020 to obtain a continuous positive index that we stratified into four categories (0,1,2, or 3).

Event-time Cox hazards model design

To study the association between lifting the eviction moratorium on the hazard of being diagnosed with COVID-19 in a given week, we used a Cox regression model with time-dependent covariates in an event-time type specification (Venkataramani et al., 2020; Clotfelter et al., 2008). This approach models the weekly probability of being diagnosed with COVID-19 at a given period conditional on having been observed without a positive diagnosis previously, where the treatment is defined as lifting the eviction moratorium and treated individuals are compared with individuals living in states that had not yet lifted their moratoria.

This study used the time from when individuals entered the study until either a COVID-19 diagnosis or the end of the study period, just like in a classic Cox analysis. Unlike a standard Cox model, however, we also used information on time since the treatment occurred (i.e., since the eviction moratorium was lifted) for the treated individuals. This method allows us to understand whether the association between expiring eviction moratoria and a COVID-19 diagnosis changed over time, which is useful when studying events that develop exponentially, such as epidemics, while also relaxing the proportional hazards assumption.

Specifically, we fitted the following model using partial likelihood :

$$\lambda(t \qquad \mathbf{Z}(t)) = \lambda \ (t) \exp \quad \boldsymbol{\beta} \, \mathbf{Z}(t) \tag{III.1}$$

$$\lambda\,(t \qquad \mathbf{Z}(t)) = \lambda \ (t) \exp \qquad \beta \ (D \qquad T \ ) + \Upsilon X \qquad + \Psi M \quad + S \quad + \delta \ + u \tag{III.2}$$

The dependent variable ($\mathbf{Z}($)) denotes the probability that an individual, , living in state, , during week is diagnosed with COVID-19. is a binary variable for the treatment group, i.e., those states that implemented an eviction moratorium but lifted it. is a binary variable that equals 1 for those treated states during the week, , relative to the week when the state lifted its moratorium. The exposure variable is bottom coded before week 15 and top-coded after week 12, implying that dynamics wear off after these points. This decision follows prior literature to avoid difficulties interpreting results due to sample size imbalances created by differences in the timing of lifting the moratoriums. $\mathbf{X}$ is a vector of time-varying covariates (i.e., non-pharmaceutical interventions and COVID-19 cases and tests), while $\mathbf{M}$ is a vector of time-invariant covariates (i.e., sex, age, type of insurance, work-industry, CCI, and z-codes diagnoses). are state-fixed effects that adjust for potential confounding from time-invariant state-level factors or baseline differences in socioeconomic characteristics, while are weekly fixed effects that adjust for nationwide secular trends in the outcome. is a vector of residuals. Standard errors are clustered at the state and week levels. We used the Breslow method for ties when running our partial likelihood estimation.

The coefficients of interest are captured by, , showing the difference in outcomes for leads and lags of lifting the eviction moratoria relative to a reference week (i.e., the week a state lifted their moratorium) and relative to all states that did not lift their eviction moratorium during the reference period.

The causal identifying assumption is that COVID-19 diagnosis risk in exposed states would have continued along the same trajectories without exposure (Venkataramani et al., 2020). We cannot directly test this assumption. Nevertheless, potential violations can be probed by examining outcome trends for events weeks before lifting the eviction moratorium. We formally tested this through a joint significance test simultaneously of all the terms before the eviction moratorium was lifted.

The primary analysis focused on being diagnosed with COVID-19 in the entire sample. We also conducted analyses stratifying by a series of individual- and zip code-level risk factors that could plausibly modify the association of expiring eviction moratoria with COVID-19 risk as time since treatment passed. Specifically, we stratified our sample by an individual's CCI score, by zip code-level poverty rate, measured by whether the percentage of individuals living below the poverty line was greater or less than 10%, a cut point commonly used to designate low-poverty neighborhoods (Ravallion, 2002; Lian, Schootman and Yun, 2008); and by zip code-level rent-burden prevalence, measured by whether more or less than half of the households renting a unit were spending at least 30% of their household income on rent, a cut point that divided our sample roughly in half and allowed us to compare higher and lower rent-burdened places with equal sample size. We tested whether the association in these subgroups increased as time since treatment passed through a joint significance test. For all models, we plotted fully adjusted hazard ratios (HRs) with 95% CIs by week, adjusted for clustering at the state and week level, centered on the week the eviction moratorium expired. Each adjusted HR shows the difference in outcomes for leads and lags of lifting the eviction moratoria relative to a reference week (i.e., the week a state lifted its moratorium) and relative to all states that did not lift their eviction moratorium during the reference period.

II.III Survival Analysis

To investigate temporal trends, however, we also fit survival curves to the data, estimating the hazard of being diagnosed with COVID-19 at time, , for total times, . Using the time-varying outcome variable, , we defined survival at, , as Pr [= 0], which is equal to Pr[T], and risk at, , as Pr [= 1], which is equal to Pr[T]. The hazard at, , is defined as Pr [= 1 = 0]. For = 1 the hazard is equal to the risk because everybody is, by definition, alive = 0. The survival probability at is the product of the conditional probabilities of surviving each interval between 0 and . More generally, the survival at is:

$$\Pr[\quad = 0] = \quad \Pr[\quad = 0 \qquad = 0] \tag{III.3}$$

That is, the survival at equals the product of one minus the hazard at all previous times (or the risk at , which is just one minus the survival). To estimate the hazard at any, , Pr [= 1 = 0], we approximate the hazards through a parametric, logistic event-time study, model restricted to individuals who survived through, , as:

$$\operatorname{logit} \Pr[\quad = 1 \quad = 0 \qquad] =$$

$$ {}_{,t} \sum_{k\,-} {}_{k} \; {}_{mk} \; {}_{m} \quad \sum_{k\,-} {}_{k} \; {}_{mk} \; {}_{m} \quad \sum_{k\,-} {}_{k} \; {}_{mk} \; {}_{m} \quad {}_{imt} \quad {}_{im} \; {}_{m} \; {}_{imt} \tag{III.4} $$

Where, , is a flexible time-varying function, $= \; + \; +$. As in the event-time Cox hazards model design, is a binary variable for the treatment group, i.e., those states that implemented an eviction moratorium but lifted it. is a binary variable that equals 1 for those treated states during the week, , relative to the week when the state lifted their moratorium. The exposure variable is bottom coded before week 15 and top-coded after week 12, implying that dynamics wear off after these points. This decision follows prior literature to avoid difficulties when interpreting results due to sample size imbalances created by differences in the timing of lifting the moratoriums. **X** is a vector of time-varying covariates (i.e., non-pharmaceutical interventions and COVID-19 cases and tests), while **M** is a vector of time-invariant covariates (i.e., sex, age, type of insurance, work-industry, CCI, and z-codes diagnoses). are state-fixed effects that adjust for potential confounding from time-invariant state-level factors or baseline differences in socioeconomic characteristics, while are weekly fixed effects that adjust for nationwide secular trends in the outcome. is a vector of residuals. Standard errors are clustered at the state and week level.

After fitting the logistic event-time study model, we then computed estimates of the survival $\Pr[\quad = 0 \quad = \quad = \quad = \quad = \;]$, under different scenarios, by multiplying the estimates of one minus the estimates of:

$$\Pr[\quad = 1 \quad = 0 \quad = \quad = \quad = \quad = \;]$$

Provided by the logistic model for each individual week. We computed two opposing counterfactual scenarios: (1) every state that implemented an eviction moratorium maintained it throughout the study period, and (2) every state that implemented an eviction moratorium lifted it on week 17. We chose week 17 because it was the first week a state lifted its eviction moratorium (see Table B.1). The difference in the survival probability of both counterfactual scenarios at the end of the study tells us what would have been the increase or reduction in the average survival probability for an individual had all the states never lifted their eviction moratoria relative to a scenario where all the states lifted the moratoria on the first week of the experiment.

To compute a confidence interval around this difference, we bootstrapped (with replacement) this exercise 50 times. We took the standard deviation of the 50 differences to build a 95% confidence interval around our original result.

II.IV Sensitivity Analysis

In sensitivity analyses, we estimated every model with a random sample of all the individuals in the OLDW to ensure that our primary design did not introduce selection bias by choosing individuals by our outcome (Hernan and Robins, 2021). Given the size of our original database and our computational limit, we worked with a 2% random sample. We overlaid these results on those calculated from the same model but with a balanced sample to assess for evidence of bias from our sample selection design. We also conducted the same exercise stratifying for the individual- and zip code-level risk factors previously described. Finally, we assessed whether expiring eviction moratoria were associated with an increase in an individual's probability of eviction by estimating our main models' expiring moratoria on the hazard of a zip code change in our claims data, a crude proxy for mobility.

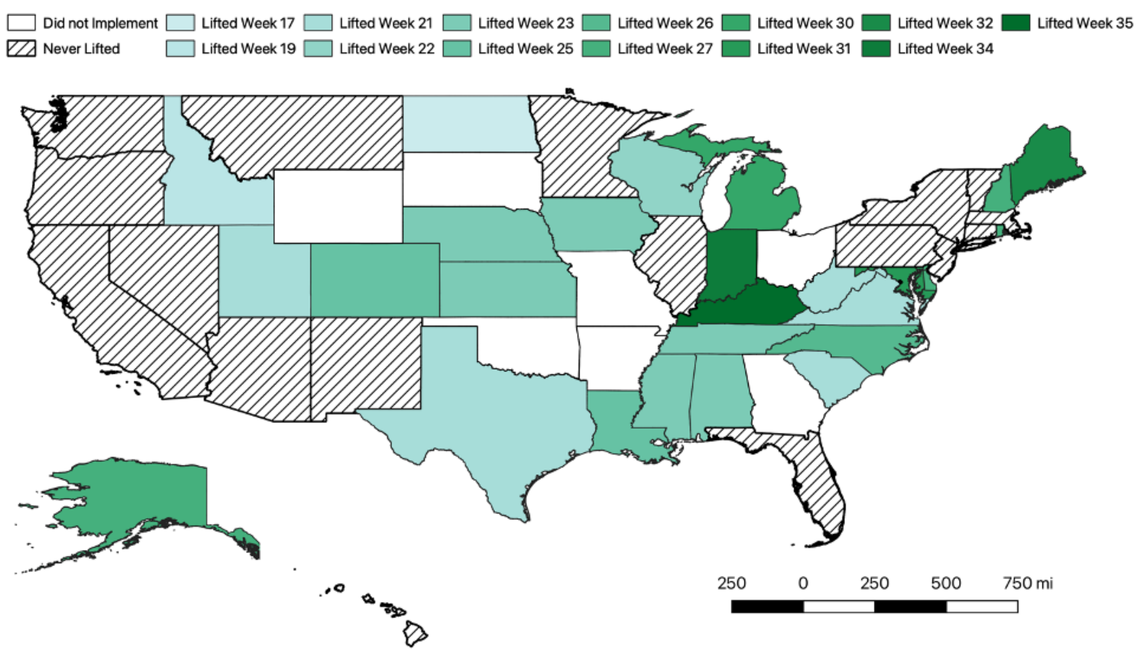
III Results

III.I Descriptive Statistics

Our study sample resided in 43 states and the District of Columbia because 7 states did not implement an eviction moratorium during our study period. (Figure III-1 and Table B.1 for exact dates). These states accounted for 88.8% of the total US population in 2019 and 89.6% of the US COVID-19 cases during the study period (Dong, Du and Gardner, 2020; US Census Bureau, 2020*a*). Overall, 18 states (40.9%) never lifted their eviction moratorium during the study period, so they were included in the control group. The remaining 26 states (59.1%) functioned as the treatment group.

Figure III-1: US States by Eviction Moratorium Lifting Status

Week of the Year



Notes: The map shows the distribution of states in the US that participated in the study, i.e., 44 states (7 states never implemented an eviction moratorium before the end of the study). Eighteen states served as the control group, which never lifted their eviction moratorium during the study period. Twenty-six states functioned as the treatment group that lifted their eviction moratorium throughout the study period. The intensity of green provides the variation in the timing of lifting the eviction moratorium for the treatment group. Alaska and Hawaii are not to scale.

During the study period, our sample included 9,475,897 individual-week observations for 509,694 individuals (254,847 [50.0%] diagnosed with COVID-19; mean [SD] age, 47.0 [23.6] years; 239,056 [53.3%] men). Baseline demographic, health, and socioeconomic characteristics were similar in exposed vs. unexposed states (see Table III.1), although there were higher COVID-19 diagnoses in states that lifted their eviction moratoria.

Table III.1: Baseline Characteristics of Individuals Included in the Estimation Sample, Stratified by Exposure Status

	Individuals by whether the state lifted the eviction moratorium, Number (%)	
Baseline characteristics	No	Yes
COVID-19 diagnosis	141,050(53.15)	113,797(46.57)
Age, mean (SD), y	47.88(22.94)	45.02(22.37)
Sex		
Male	123,961(46.72)	115,095(47.12)
Female	141,359(53.28)	129,123(52.88)
Insurance		
Commercial	190,935(71.95)	183,716(75.19)
Medicare Advantage	74,424(28.05)	60,619(24.81)
Charlson Comorbidity Index score, mean (SD)	0.69(1.10)	0.58(1.02)
Flag		
Unemployment	288(0.10)	247(0.10)
Housing and economic circumstances	326(0.12)	352(0.14)
Bereavement	383(0.14)	426(0.17)
Zip code, mean (SD), %		
Poverty rate[a]	11.31(7.39)	12.51(8.07)
Rent burden prevalence[b]	50.27(10.30)	45.79(9.85)
Number of individuals	265,359	244,335
Number of states	18	26

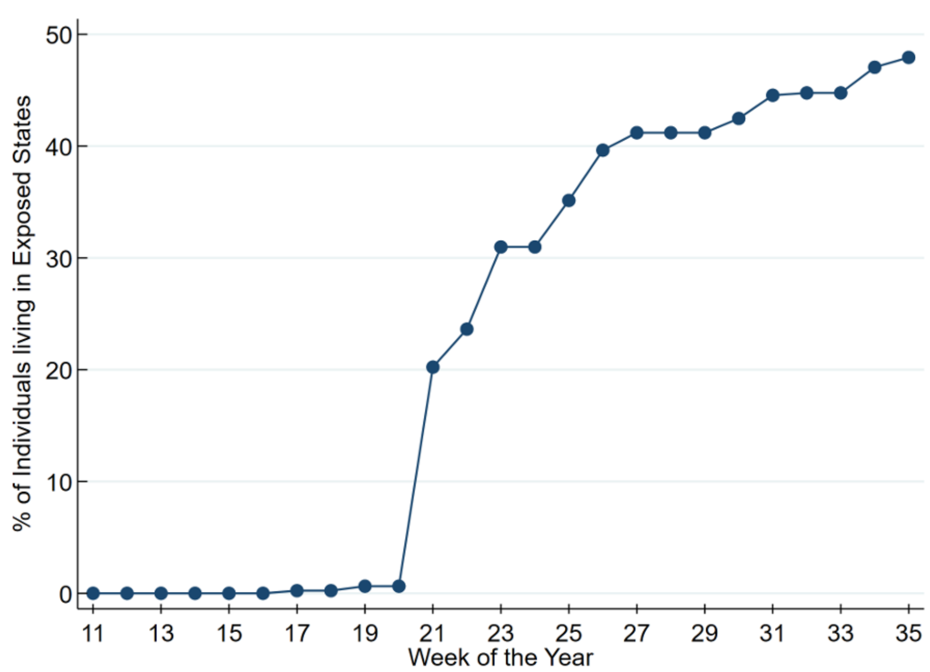
[a]: Percentage of individuals living below the poverty line at the zip code level where the individual lives.

[b]: Percentage of households renting a unit and spending at least 30% of their household income in rent where the individual lives.

As seen in Figure III-2, individuals were exposed to lifting the eviction moratorium from week 17 to 35.

Figure III-2: Trends in Share of Individuals Exposed to Treatment by the week of the year

Percentage



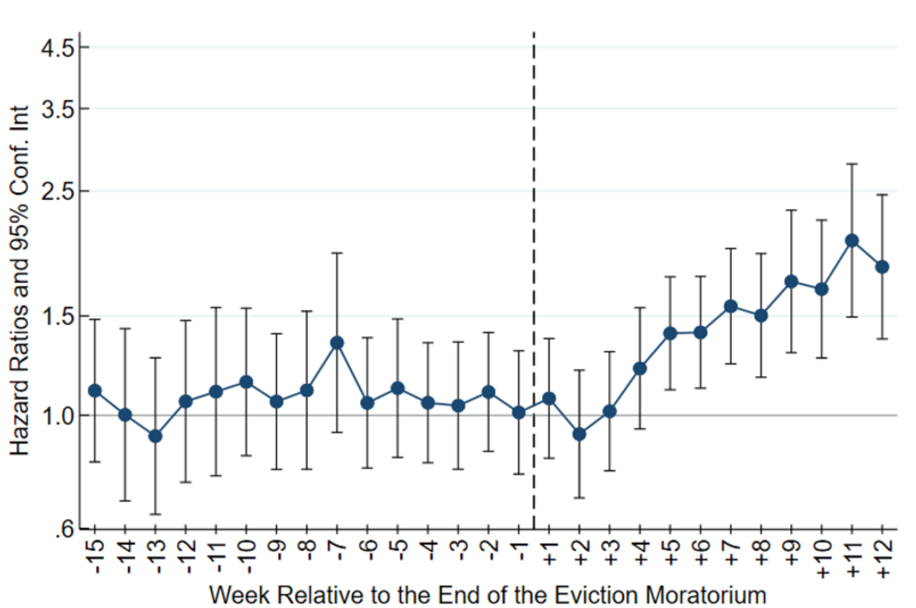
Notes: Trends in the percentage of individuals in the study sample exposed to treatment. Estimates come from the balanced sample. The y-axis shows the cumulative percentage of individuals in the study sample who lived in a state that lifted their eviction moratorium during the study period. At the beginning of the study, in week 11 of 2020, no individuals were exposed to a state lifting their eviction moratorium. By the end of the study period, week 35 of 2020, 244,335 individuals lived in a state that lifted their eviction moratorium.

III.II Eviction Moratoria Expiration and COVID-19 Risk

Figure III-3 plots the fully adjusted HRs of our main model. Before moratoria, there was no difference in trends in COVID-19 diagnosis risk between individuals in states lifting moratoria vs. those keeping them in place, i.e., we cannot reject the jointly null hypothesis in which every coefficient is equal to 1 before the moratoria (= 5 35; P = 98), suggesting that in the absence of exposure, treatment and control groups would have continued along the same trajectory. Individuals living in states that lifted their eviction moratorium, relative to those living in states that never lifted their moratorium, were more likely to be diagnosed with COVID-19 beginning 5 weeks after the eviction moratorium was lifted (HR=1.39; 95%CI,1.11-1.76; P=.004) and reaching an HR of 1.83(95%CI, 1.36-2.46; P 001) at 12 weeks or longer.

Figure III-3: Event Study Estimates of the Association between Lifting the Eviction Moratorium and COVID-19

Hazard Ratios

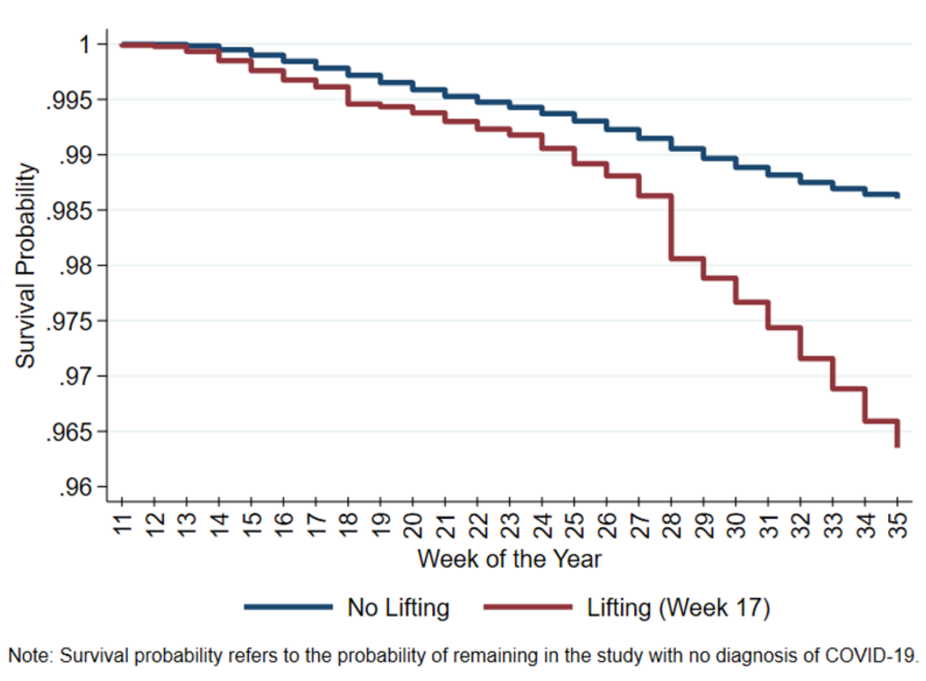


Note: Hazard Ratios and 95% Confidence Intervals. Y-axis is in log scale. The vertical dashed line represents the end of the eviction moratorium. X-axis represents the number of weeks relative to the end of the eviction moratorium.

Looking at the cumulative difference in the hazard of COVID-19 infection during the study period (Figure III-4), we observed an average 2.4-percentage point (95%CI, 0.3-4.3 percentage points) higher probability of remaining in the study with no diagnosis of COVID-19 (P = 01) between the counterfactual scenarios in which every state lifted the eviction moratorium in week 17 of the year vs. never lifting it.

Figure III-4: Survival Curves on the Association between Lifting the Eviction Moratorium and COVID-19

Survival Probability at   conditional on surviving up to   1



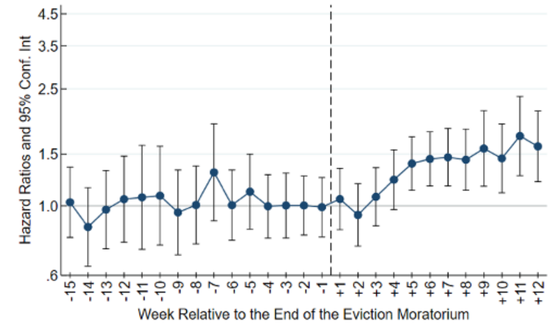
Scenario/Week	11	15	19	23	27	31	35
No Lifting	463,520	463,277	462,210	461,086	459,940	458,351	457,051
Lifting (Week 17)	463,520	462,824	461,005	459,960	457,997	452,705	446,612

Figure III-5 plots the time-varying association between expiring eviction moratoria on individuals by baseline health strata, showing that associations increased with CCI score. The magnitude of the association increased as time since lifting an eviction moratorium passed for individuals with greater CCI scores. Individuals with a CCI of 3 or greater living in a state that lifted its eviction moratorium had an HR of 2.36 (95%CI, 1.67-3.36; P   001) after 12 weeks compared with those living in a state that never lifted its moratorium. The healthiest group (i.e., CCI score 0) was the only subgroup among the health strata where the associations plateaued after week 4 (= 3 54; P = 83).

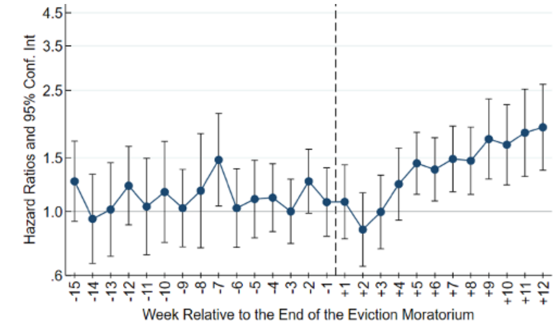
Figure III-5: Event Study Estimates of the Association between Lifting the Eviction Moratorium and COVID-19, stratified by Charlson Comorbidity Index

Hazard Ratios

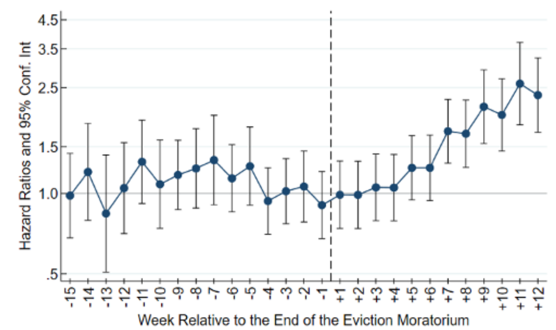
(a) Charlson Comorbidity Index = 0



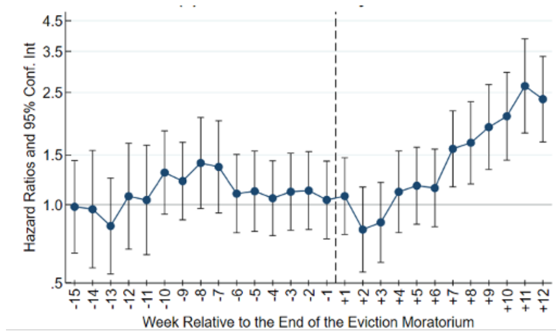
(b) Charlson Comorbidity Index = 1



(c) Charlson Comorbidity Index = 2



(d) Charlson Comorbidity Index 3

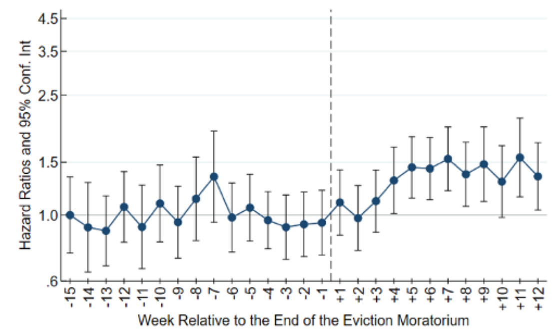


Note: Hazard Ratios and 95% Confidence Intervals, stratified by the Charlson Comorbidity Index of the individual calculated at the beginning of the year 2020. Y-axis is in log scale. The vertical dashed line represents the end of the eviction moratorium. X-axis represents the number of weeks relative to the end of the eviction moratorium.

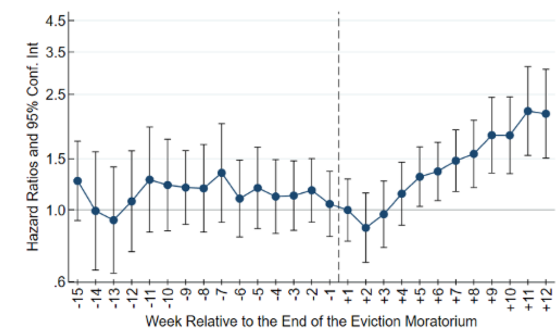
Figure III.6 plots the associations between moratoria and COVID-19 diagnosis risks by area-level poverty rates and rent burden, showing increasing associations for individuals in zip codes with higher levels of each. For areas with high poverty and a high rent burden, we can reject the null hypothesis that HRs after week 4 were equal for both groups (high poverty: = 16 04; P = 02; high rent burden: = 25 82; P 001). Those living in non-affluent areas had an HR of 2.14 (95% CI, 1.51-3.05; P 001), while those living in areas with high rent burden had an HR of 2.31 (95% CI, 1.64-3.26; P 001). However, we found statistically significant higher hazards for individuals living in low-income and rent-burdened rate areas where the eviction moratoria were lifted compared with those living in control states, although they did not increase as time passed since lifting the eviction moratorium. In both the low-poverty and low rent-burdened rate models, we cannot reject the null hypothesis that HRs after week 4 were all equal (low-poverty: = 5 79; P = 0 57; low rent burden: = 4 35; P = 74)

Figure III.6: Event Study Estimates of the Association between Lifting the Eviction Moratorium and COVID-19, stratified by Poverty and Rent-Burdenship Rate
Hazard Ratios

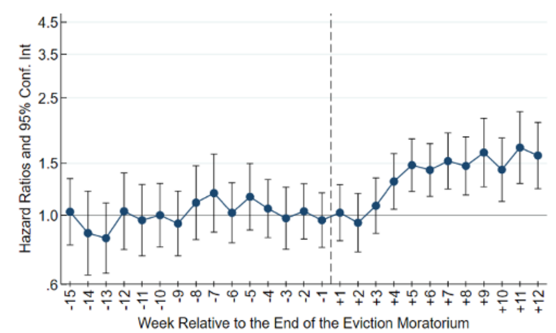
(a) Poverty Rate 10



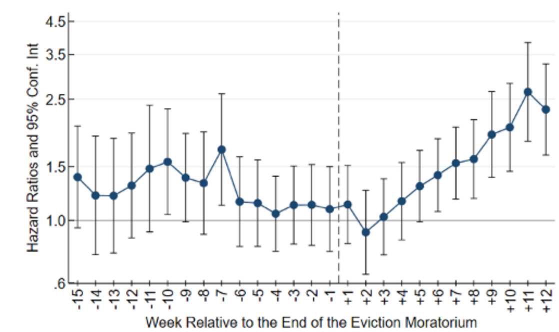
(b) Poverty Rate 10



(c) Rent-Burdened Rate 50



(d) Rent-Burdened Rate 50



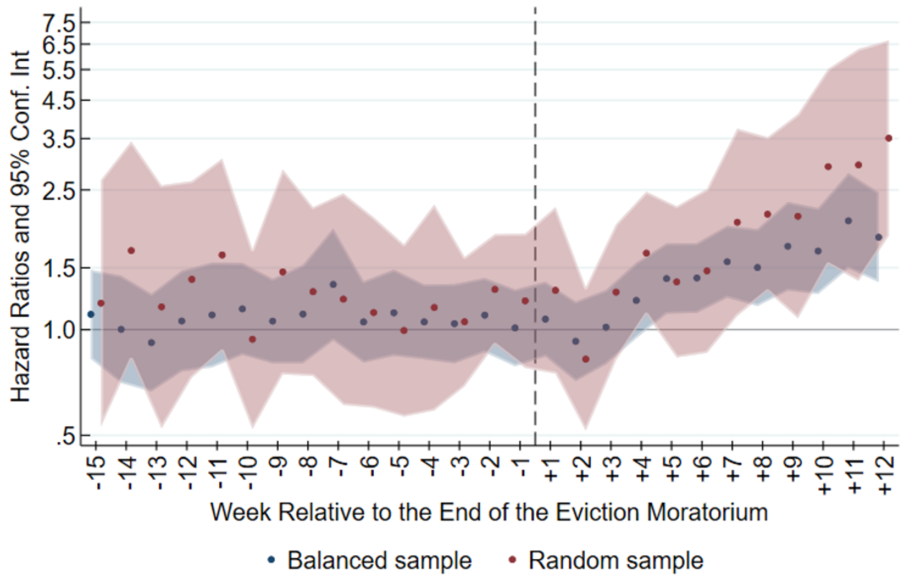
Note: Estimates come from the balanced sample. Y-axis is in log scale. The vertical dashed line represents the end of the eviction moratorium. X-axis represents the number of weeks relative to the end of the eviction moratorium. Hazard Ratios and 95% Confidence Intervals, stratified by the Poverty Rate and Rent-Burdened rate at the zip code level where the individual lived during the study period.

III.III Sensitivity Analyses

Sensitivity analyses showed that the coefficients and confidence intervals of the balanced sample fell within the confidence intervals of the 2% random sample (Figure III.7), providing evidence that our sample selection design did not bias our estimates.

Figure III.7: Event Study Estimates with Different Sample Designs of the Association between Lifting the Eviction Moratorium and COVID-19

Hazard Ratios



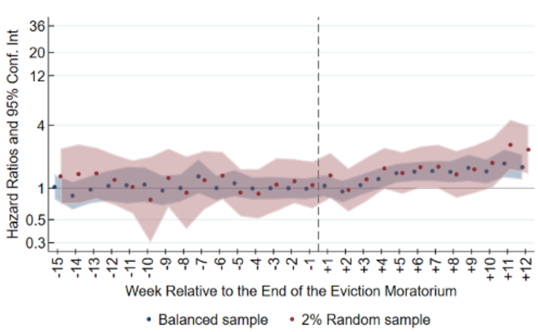
Note: Y-axis is in log scale. The vertical dashed line represents the end of the eviction moratorium. X-axis represents the number of weeks relative to the end of the eviction moratorium. Hazard Ratios and 95% Confidence Intervals. We used, in separate models, both the balanced sample (in blue) and the 2% random sample (in red) to conduct this analysis.

Furthermore, we found the same pattern when conducting the same analysis but using the CCI score and the poverty and rent burden rates subgroups (Figure III.8 and III.9).

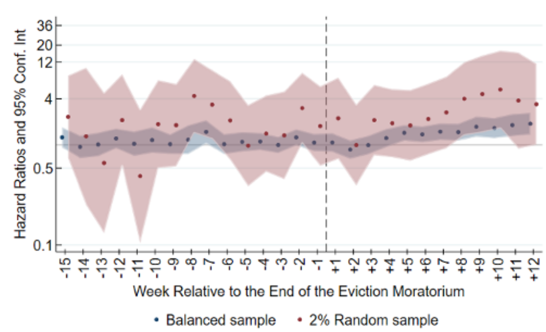
Figure III.8: Event Study Estimates with Different Sample Designs of the Association between Lifting the Eviction Moratorium and COVID-19, stratified by CCI

Hazard Ratios

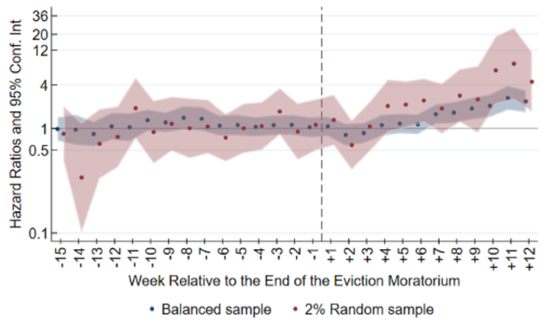
(a) Charlson Comorbidity Index = 0



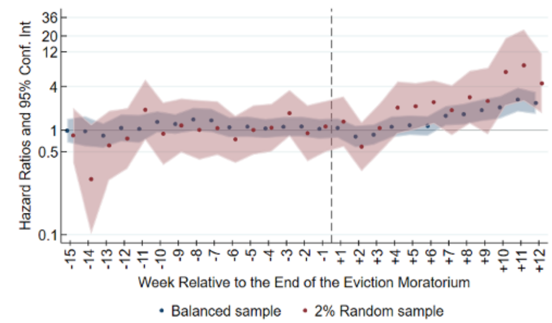
(b) Charlson Comorbidity Index = 1



(c) Charlson Comorbidity Index = 2



(d) Charlson Comorbidity Index 3

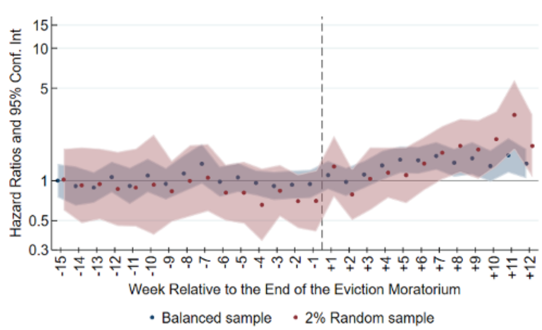


Note: Y-axis is in log scale. The vertical dashed line represents the end of the eviction moratorium. X-axis represents the number of weeks relative to the end of the eviction moratorium. Hazard Ratios and 95% Confidence Intervals. We used, in separate models, both the balanced sample (in blue) and the 2% random sample (in red) to conduct this analysis.

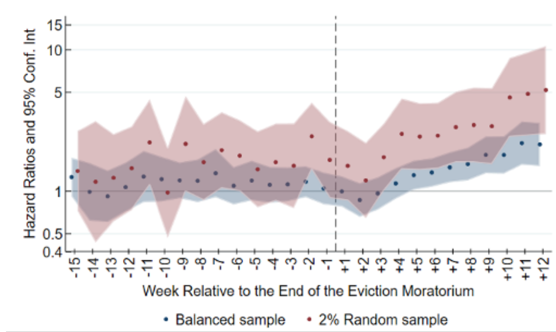
Figure III.9: Event Study Estimates with Different Sample Designs of the Association between Lifting the Eviction Moratorium and COVID-19, stratified by Poverty and Rent-Burdenship Rate

Hazard Ratios

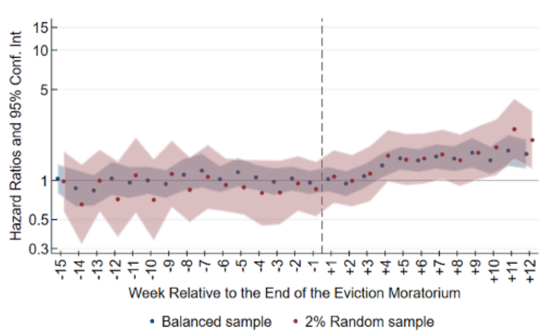
(a) Poverty Rate 10



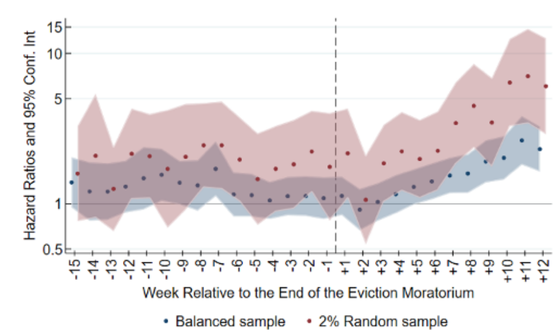
(b) Poverty Rate 10



(c) Rent-Burdened Rate 50



(d) Rent-Burdened Rate 50

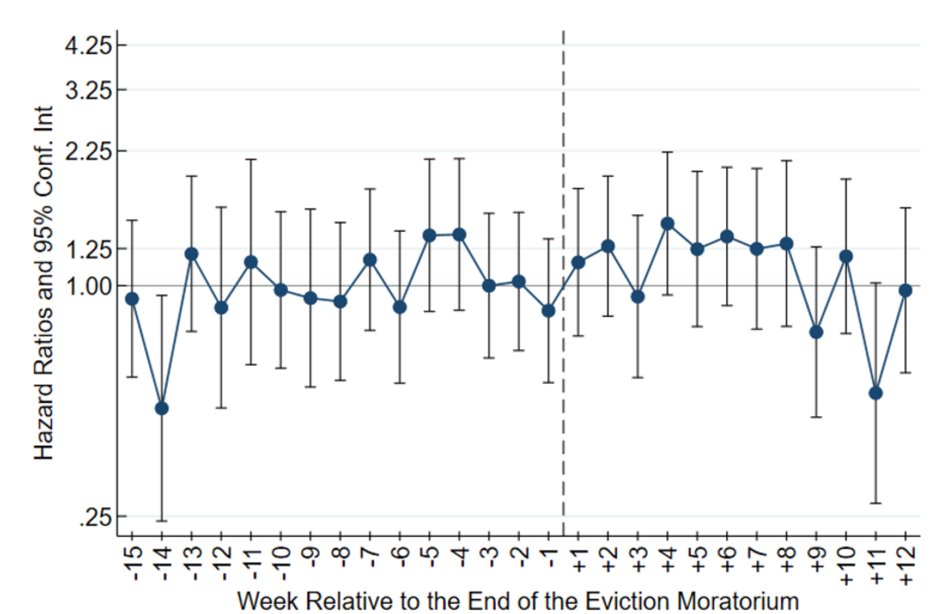


Note: Y-axis is in log scale. The vertical dashed line represents the end of the eviction moratorium. X-axis represents the number of weeks relative to the end of the eviction moratorium. Hazard Ratios and 95% Confidence Intervals. We used, in separate models, both the balanced sample (in blue) and the 2% random sample (in red) to conduct this analysis.

We found no association between expiring eviction moratoria and whether an individual in our data set changed their zip code, suggesting that personal eviction experience was not the main mechanism by which expiring eviction moratoria caused increased COVID-19 hazard (Figure III.10).

Figure III.10: Event Study Estimates of the Association between Lifting the Eviction Moratorium and Changing Zip Code Address

Hazard Ratios

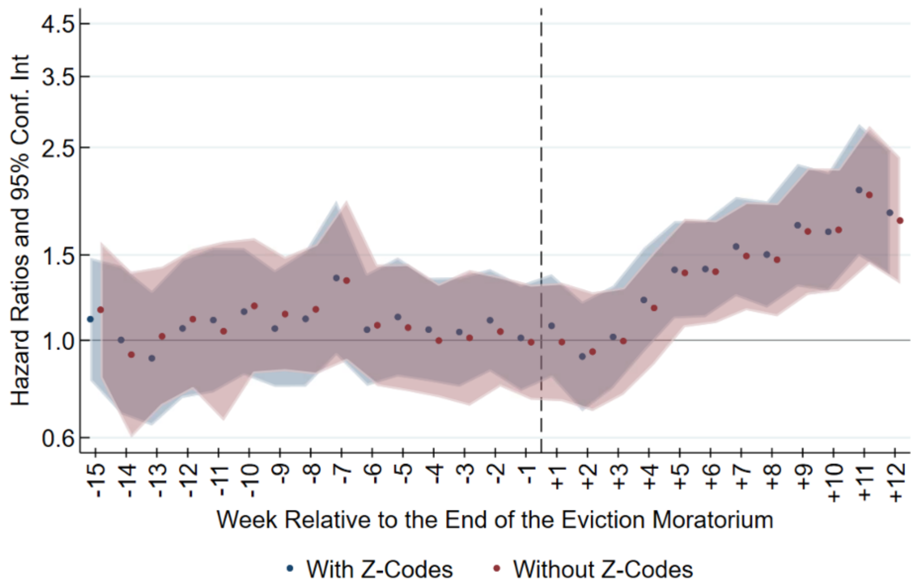


Note: Estimates come from the balanced sample. Y-axis is in log scale. The vertical dashed line represents the end of the eviction moratorium. X-axis represents the number of weeks relative to the end of the eviction moratorium. Hazard Ratios and 95% Confidence Intervals, where the main outcome is a binary variable indicating whether the individual requested a change of residential zip code.

Finally, while ICD-10 Z codes are underused by practitioners (Guo et al., 2020), excluding these covariates did not affect results (Figure III.11).

Figure III.11: Event Study Estimates of the Association between Lifting the Eviction Moratorium and COVID-19, with and without Z-Codes

Hazard Ratios



Note: Y-axis is in log scale. The vertical dashed line represents the end of the eviction moratorium. X-axis represents the number of weeks relative to the end of the eviction moratorium. Hazard Ratios and 95% Confidence Intervals. We used separate models of our main specification, both with z-codes (in blue) and without z-codes (in red), to conduct this analysis.

IV Discussion

Using individual-level healthcare claims data, we found that lifting eviction moratoria was associated with an increase in the hazard of a COVID-19 diagnosis beginning 5 weeks after an eviction moratorium was lifted and persisting for at least 12 weeks after that point. As what we believe to be the first study on eviction policy and COVID-19 diagnoses to use individual-level data, we found that the hazards associated with lifting eviction bans increased with time among individuals with preexisting health problems. Our findings suggest that even individuals with no comorbidities were put at risk by expiring eviction moratoria after controlling for age and social factors, such as insurance type, occupational industry, history of unemployment, problems related to housing and economic circumstances, and area-level covariates. The result of the sensitivity analysis showing no association of expiring eviction moratoria on the hazard of individuals in this data set changing zip codes is consistent with previous findings in the literature, i.e., an individual's hazard of COVID-19 diagnosis was not just affected by personal experiences with eviction but also by spillovers from the transmission process created by evictions within a community (Nande et al., 2021). While previous ecological evidence showed that area-level COVID-19 incidence increases after eviction moratoria are lifted, these county-level analyses have not been able to answer the question of who, specifically, is put at risk by allowing evictions to occur during the COVID-19 pandemic (Leifheit et al., 2020).

Our findings are clear that the hazard of COVID-19 diagnosis increases for all individuals when eviction bans are allowed to expire, but that individuals with preexisting health problems and those living in areas with higher poverty or with a higher prevalence of rent-burdened households have disproportionately higher risk as time since ending the moratoria passes. As such, eviction moratoria should be thought of as a health equity intervention that has helped narrow the gap in risk between affluent and non-affluent neighborhoods and between individuals based on preexisting health conditions, which, especially after age adjustments, are known to be associated with social determinants of health, including individual-level socioeconomic status and exposure to racism (Kawachi and Subramanian, 2018).

Our investigation was designed as an event-time study that exploits the variation of some states implementing, lifting, or maintaining eviction moratoria while also including the timing of other COVID-19-related policy changes, such as mask mandates and school closures, that could have been timed in concert with eviction policy changes and could also affect COVID-19 hazard as well as with a set of individual- and area-level covariates to isolate the associations of expiring eviction moratoria (Angrist and Pischke, 2009; Lechner, 2011). In the weeks before the eviction moratorium was lifted, there was no statistically significant difference in the HR of being diagnosed with COVID-19 between the states that lifted and did not lift their eviction moratoria, suggesting that the probability of being diagnosed with COVID-19 would have evolved similarly in all states absent the treatment. While we created a control balanced panel of individuals who were and were not diagnosed with COVID-19 during the observation period to provide power to the stratified analyses, we also conducted our main model on a 2% random sample of individuals who were not selected concerning the outcome and found similar results, albeit with wider confidence intervals.

IV.I Limitations

This study has limitations. First, we cannot rule out the chance that our associations could be explained by residual confounding, despite our methods and sensitivity analyses. Second, we relied on COVID-19 diagnoses as our outcomes. Thus, we are not including asymptomatic cases or individuals not interacting with the health sector despite having COVID-19. Third, our data set did not include information from individuals with Medicaid or those who are uninsured. However, since many of these individuals are at high risk of eviction and COVID-19, including them would strengthen the associations between expiring moratoria and COVID-19 (Allen et al., 2019). Thus, our results should be considered a lower bound. Additionally, for privacy reasons, we did not have access to beneficiary race and ethnicity. We so cannot describe the implications of allowing eviction moratoria to expire for racial and ethnic disparities in COVID-19 infection.

V Conclusion

In this study with a difference-in-differences analysis, residents in states that lifted an eviction moratorium experienced an increased risk of being diagnosed with COVID-

19 compared with residents of states that maintained moratoria. The magnitude of associations increased over time after the moratoria were lifted among individuals with more comorbidities and for those living in higher poverty and rent-burdened zip codes. Beyond lessons for managing the COVID-19 pandemic as new variants spread, this study suggests that a housing policy that protects individuals with low income and/or more comorbidities can promote health equity and create protection for groups with more advantages.

Chapter IV

House Flipping Within Rapidly Changing Neighborhoods

I Introduction

House flipping, or the purchase of housing by an investor who attempts to profit from buying low and selling high rather than for occupation, has become increasingly popular in recent years (Depken, Hollans and Swidler, 2009; Leung and Tse, 2017). Thanks to the rise of reality TV shows, podcasts, seminars, blogs, or books that glamorize house flipping, people have been attracted to the idea of flipping houses as a means of making a quick profit. House flipping media popularity has been reflected in the housing market since house flipping in 2022 constituted 8.4 percent of all home sales in the United States, representing the largest figure since at least 2005 (ATTOM, 2023).

The rise in flipping activity has coincided with a period where housing supply shortage and housing demand remain at historically high levels (Betancourt, Gardner and Palim, 2022; Joint Center for Housing Studies, 2022). It's thus not surprising that as the number of investors and house flippers increases, so too does the concern about the harms and benefits of this practice at the local level (see Demsas (2023), National Association of Realtors (2022), and Zhang (2022) for examples of opposing points of view).

With historically rising demand for housing and short housing supply, investors and house flippers have found that the real estate market offers potentially attractive monetary returns. In this context, there is also growing concern about the impact of house flipping on rapidly changing neighborhoods, including the potential displacement of long-term residents, the disruption of community stability, and how it could crowd out first-time and minority homebuyers. In this sense, this chapter of my dissertation explores the following questions: (a) What is the impact of house flipping in neighborhoods experiencing rapidly changing housing and demographic processes? (b) What mechanisms do house flippers and investors use to acquire properties in these neighborhoods, and what type of properties do they target? and (c) Could local expertise of residents living in these rapidly changing neighborhoods help us understand why certain areas are more likely to be heavily flipped and what harms and benefits house flipping creates on local communities and residents?

To provide insights into these questions, I employ a mixed-methods approach incorporating various data sources and analytical techniques. Specifically, I use housing transaction records from the state of Massachusetts, digitized scans of HOLC maps, and a Participatory Action Research (PAR) collaboration with the Healthy Neighborhoods Research Consortium (HNS) coupled with descriptive trajectories, hedonic regression models, and a spatial regression discontinuity design. To ensure the robustness of my findings, I also apply inverse probability weights to my spatial regression discontinuity analysis using novel information from the digitized HOLC maps.

I have the following set of results. First, within HNS neighborhoods, almost a quarter of all affordable unit housing sales (i.e., 23%) were flipped between 2008 and 2021, compared to only 7% of expensive unit housing sales. Second, flipped properties generated higher financial gains on average than non-flipped sales, with gains increasing over time and reaching a difference with non-flipped sales of almost 40% higher in 2021. By disentangling the flipping flow, I found that compared to similar properties, houses to be flipped were sold at a lower price at purchase and were sold for a higher amount at the sale than not-flipped houses. Third, homes that were later flipped were more likely to be bought by investors (e.g., LLCs) than homes that were not flipped, and they were also more likely to be bought with all cash. For example, in 2021, homes that were later flipped were 42% more likely to be bought by an investor and 20% more likely to be bought with cash than homes that did not get flipped. Fourth, properties within neighborhoods formerly delineated by HOLC with a grade "D" (i.e., Red) and "C" (i.e., Yellow) were 12.2% more likely to be flipped than those within neighborhoods formerly delineated by HOLC as grade "A" (i.e., Green).

Regarding the PAR collaboration, resident researchers helped disentangle the effects of house flipping on their neighborhoods. While house flipping led to benefits such as new schools, shopping centers, and increased home equity, it also caused harm, such as displacement of tenants, poor-quality housing, and increased evictions. Finally, in group discussions, resident researchers suggested implementing policies targeted to increasing access to information, supporting home repairs, revising zoning laws, creating cooperative and land trusts with limited-equity covenants, and providing financial assistance to first-time homebuyers to better adjust rapidly changing neighborhoods to house flipping.

This chapter builds upon and extends the existing literature in the following ways. First, recent empirical studies have started highlighting the extent of investors in the housing market and their impact on the volatility of housing prices (Bayer et al., 2011; Depken, Hollans and Swidler, 2009; Lee and Choi, 2011; Leung and Tse, 2017). This research suggests that house flipping has a complex and non-linear impact on the volatility and cycle of housing prices. For example, Lee and Choi (2011) found that when more flippers entered the housing market, they created a positive upward movement in home prices. Leung and Tse (2017) added flippers to a housing market search model and found that flipping tends to occur in sluggish and tight markets, resulting in a rapid turnover, a high vacancy rate, and higher housing prices than under a counterfactual scenario. Furthermore, house flipping in a tight and liquid market can create inefficiencies and a surplus loss as the efficiency gain from faster turnover is unlikely to be large enough to offset the loss from more houses being left vacant in the hands of flippers. While existing studies have shed light on the effects of house flipping on average or median house prices, they have yet to explore how the practice affects the entire price distribution or delve into the strategies used by flippers to gain a competitive edge in the market. My research contributes to the house flipping and housing investors' literature by providing estimates of the specific segments of the housing market targeted by flippers, analyzing their impact on housing affordability, exploring whether flippers the condition of the houses they flip, and examining the average returns on investment within rapidly changing neighborhoods. Additionally, this study delves into the strategies and mechanisms utilized by flippers to gain a competitive advantage in the sector.

Second, my research adds to the existing body of knowledge on how historically racialized-driven disinvestment affects the housing market and society. Structural racism plays a major role in perpetuating economic inequities (Bailey et al., 2017; Bassett, Chen and Krieger, 2020). This term refers to the macro-level systems, social forces, institutions, ideologies, and processes that interact with one another to generate and reinforce inequities among racial and ethnic groups in the various systems in society, such as housing, education, employment, earnings, benefits, credit, media, healthcare, and criminal justice (Powell and Colyvas, 2008). These macro-level systems often reinforce discriminatory beliefs, values, and resource allocation, as research has shown (Bailey et al., 2017), and they can hurt economic mobility outcomes, as studies have found (Chetty et al., 2020). This chapter adds to the existing literature by examining the association between historical disinvestment and modern-day house flipping.

Lastly, I contribute to the growing literature that uses big data coupled with PAR methods. For instance, Daepp et al. (2022) and Costanza-Chock (2020) offer valuable insights into how urban planners can effectively engage with communities using big data. These studies emphasize the importance of deep collaboration between academic and resident researchers at every stage of the research process, including developing research questions, data analysis, and interpretation of results. This chapter provides an additional framework and a practical example for iteratively designing housing research and analyzing complex data sets in neighborhoods grappling with rising housing prices, quick neighborhood change, and heightened investor activity.

The remainder of this paper is structured as follows. Section II introduces the HNS Consortium and the research question. Section III describes the data, the empirical strategy, and the collaborative data analysis process. Section IV goes through the empirical results. Section V presents the findings of the participatory action research.

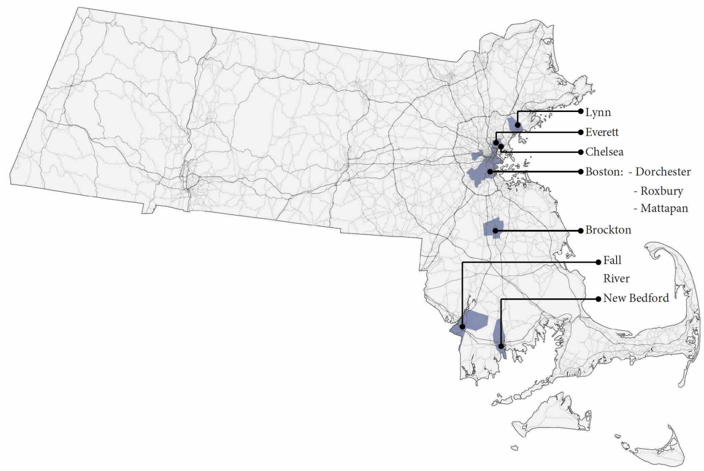
Section VI concludes.

II The Healthy Neighborhood Study (HNS) Consortium

The HNS Consortium is a network of academic researchers, partners from community-based organizations, regional advocacy organizations, government agencies, and residents of nine neighborhoods (i.e., Lynn, Everett, Chelsea, Roxbury, Dorchester, Mattapan, Brockton, New Bedford, and Fall River) in the greater Boston, Massachusetts, area (see Figure IV.1 for spatial locations). The HNS consortium uses PAR as a research practice to integrate diverse perspectives between urban development and community health within low-income communities in the early-to-mid stages of transformational economic growth. The neighborhoods are all dense, mixing single-family and multifamily housing with a high reliance on public transportation (Arcaya et al., 2018).

Through the HNS consortium, a diverse group of resident researchers, either current or recent residents of the study neighborhoods, work with the HNS network (i.e., academic researchers, partners from community-based organizations, regional advocacy organizations, and government agencies) throughout all stages of the research process. The consortium brings a range of perspectives and experiences to ensure a more comprehensive understanding of neighborhood change. For further detail on the HNS PAR model and methodology, see Arcaya et al. (2018) and Binet et al. (2019).

Figure IV.1: Healthy Neighborhood Study Locations



The HNS Consortium PAR process comprises five yearly phases involving resident researchers, academic researchers, and community partner organizations, which are (i) scoping, (ii) study design, (iii) data collection, (iv) data analysis, and (v) action, see Daepp et al. (2022) for further details. In 2020, while conducting the second phase, the Consortium conducted its annual collaborative research design workshops to set a shared agenda and develop research questions. During this phase, resident researchers and community partner organizations focused on making meaning of the experience of neighborhood change, arriving at the following potential research question: *"What are the systems of benefit and harm influencing development in HNS neighborhoods, and how do they work?"*

The HNS Consortium partnered with outside data partners and researchers to answer this question. In this context, this chapter explores the practice of house flipping as a system of benefit and harm influencing development within the HNS neighborhoods.

III Data, Empirical Strategy, and Collaborative Data Analysis

In this section, I discuss the data used in this study and the empirical strategy used to analyze it. I also describe the Collaborative Data Analysis conducted with the HNS consortium.

III.I Data

The data sets used in this paper come from multiple sources: (a) property transactions from The Warren Group, (b) approved property permits from the City of Boston, and (c) digitized HOLC maps and area description files from the University of Richmond and Markley (2023).

Property Transactions

Real-estate information (i.e., property transactions) comes from the Warren Group, consisting of property and mortgage transaction records gathered from the county-level registry of deeds offices in Massachusetts from 2008 to 2021. Each record contains unique property and transaction identifiers, sale date, sale price, names of the buyers and seller, mortgage amount, assessor parcel number (APN), geographic characteristics of the property, and basic characteristics of the property, such as the number of rooms, bathrooms, among others. The data is organized at the property level, and I restrict it to property transactions (*not mortgage transactions*) that occurred within the nine neighborhoods of the HNS study.

Approved Property Permits

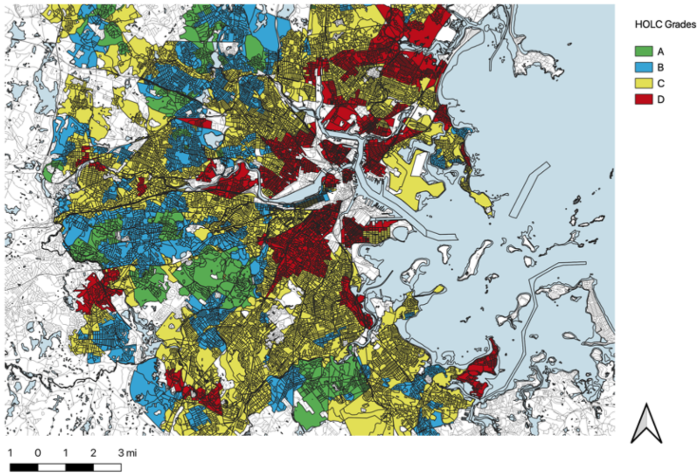
This dataset contains approved property permits issued by the City of Boston and is organized according to individual parcels of land. Each observation contains information regarding a permit granted by the City of Boston since 2009 for a specific property, with any permits that are in the process of being approved or have been denied, deleted, voided, or revoked not included. The types of property permits covered by this dataset include short-form, electrical, plumbing, gas, electrical low voltage, long-form/alteration, electrical fire alarms, certificate of occupancy, excavation, electrical temporary service, an amendment to a long-form, erect/new construction, use of premises, and foundation. The unique identifier of this dataset is the assessor parcel number (APN). As with the property transactions data, I restrict the approved property to those occurring within the HNS neighborhoods of Boston (i.e., Dorchester, Roxbury, and Mattapan).

I merged the approved property permits dataset using the unique APN and Transaction IDs. Since permits with a zero valuation primarily involve administrative paperwork, I excluded them from the dataset. My objective is to identify physical changes made to properties. To create a consolidated dataset, I aggregated the cost and count of permits issued the year before the property sale date. For instance, a property may have received several permits between August 12$^{\text{th}}$, 2014, and August 11$^{\text{th}}$, 2015, before its sale on August 12$^{\text{th}}$, 2015. I combined these permits into one row, allowing me to have a single record per transaction and APN identifiers.

Digitized HOLC Maps

The Mapping Inequality initiative by the University of Richmond has made available digitized scans (i.e., spatial polygons) of HOLC redlining maps from the National Archives, offering geocoded renderings of the original maps for 149 cities within the entire continental US (Nelson et al., 2023), while Markley (2023) provided the characteristics used by real-estate professionals used as inputs to create these maps. Both datasets provide geocoded renderings of the original HOLC maps for the continental US. The renderings provide geographic information on the four-color schema assigned to neighborhoods by the Federal Home Loan Bank Board in 1932. Those colors are A/"Best" = green; B/"Still Desirable" = blue; C/"Definitely Declining" = yellow; and D/"Hazardous" = red. Figure IV.2 visually represents Boston's four-color schema.

Figure IV.2: HOLC Maps in the Greater Boston Area, Massachusetts



The digitized HOLC maps are used in the following ways. First, I spatially merge properties and HOLC polygons using latitude and longitude coordinates from each property transaction to assign each property to a HOLC grade (i.e., A, B, C, or D). Second, to create HOLC boundaries, I create borders (i.e., straight lines) from two adjacent HOLC polygons (e.g., polygon B and C) and assign a unique ID to each one. To each HOLC boundary, I create a 100 meters buffer around it, and I spatially merge properties and HOLC buffers. From the last step, each spatially merged property has a HOLC letter, the distance to the boundary in meters, and the unique ID of the boundary. Given the spatial nature of the HOLC maps, I do not restrict my observations to the HNS neighborhoods but to all of those occurring within Massachusetts HOLC locations.

III.II Empirical Strategy

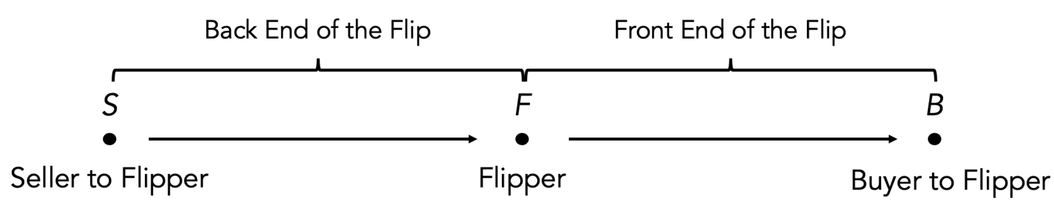
In this section, I discuss how I define house flipping; how I identify investors and entities; what target estimands I seek to estimate, and their respective estimators.

Defining House Flipping

House flipping—i.e., an acquired arm's length property transaction that is resold within weeks or months—by individuals and investors, is a key concept of this research. I defined a property transaction as a house flip when the same property was bought and sold at arm's length[1] manner within less or equal to 12 months, similar to Depken, Hollans and Swidler (2009). To minimize entry mistakes, capture fair market values, and reduce measurement error, I discard property transactions where (a) the sale price is lower than $1,000 (in 2019 USD dollars); (b) they were sold twice in less than 15 days; or (c) where the flipper was a bank, federal agency, holding or relocation company. Every transaction price was measured in real 2019 USD dollars using the CPI index.

Of importance, this research thinks of a house flip as a flow that involves three parties. Figure IV.3 shows a visual representation of this flow (i.e., the back end of the flip and the front end of the flip) and the three involved parties (i.e., the seller to the flipper, the flipper, and the buyer to the flipper).

Figure IV.3: Cycle of a House Flip



Identifying Investors and Entities

Using natural language processing techniques, I employed a supervised text classification model to identify an investor by implementing the steps described below.

First, I extracted the full names of the buyers and sellers of every property transac-

[1] An arm's length transaction occurs via the market, and it excludes transactions between related individuals, which might carry out transactions at non-market prices.

tion. In total, I had 446,943 unique full names. Second, I randomly chose 10% of those names. I classified each name as an investor (using a dummy variable) through a manual process, considering the presence of words or abbreviations associated with investors, such as Associations, Trustees, Companies, Limited Liability Partnerships, Joint Ventures, and Corporate Trusts. Names that did not appear to belong to an individual, bank, or public entity were also identified as investors. Third, I constructed a document matrix of the full names by utilizing the dataset that included the investor classification variable alongside their corresponding full name.[2] Fourth, using the investor variable as a target and the document matrix as the features, I trained a random forest model. Fifth, I utilized the trained model to predict the probability of a given full name belonging to the investor category for the remaining 90% of the unlabeled data. A name was classified as an investor if the probability of being in this category exceeded 0.95. In the sixth step, I appended the resulting predicted dataset with the initial manually labeled dataset to form the final investor-labeled dataset.

To identify an entity, I followed similar steps as with investors. For the training dataset, I classified each name as an entity (using a dummy variable) through a manual process, considering the presence of words or abbreviations associated with entities such as banks, credit unions, or public corporations such as HUD, the Massachusetts Department of Transportation, among others.

Target Estimands

This research focuses on answering the following empirical questions: (a) What is the impact of house flipping in neighborhoods experiencing rapidly changing housing and demographic processes? (b) What mechanisms do house flippers and investors use to acquire properties in these neighborhoods, and what type of properties do they target?

[2] As a pre-processing step, I performed several actions on the full name data, including converting it to lowercase, removing punctuation, stemming words, and eliminating sparse terms.

Given these questions, I seek to identify two target estimands.

Target Estimand (a) According to the potential outcomes framework, a property transaction denoted by, , can have two potential outcomes depending on whether it is flipped or not. For property transaction, , I denote these potential outcomes as ( ), where denotes the property's transaction potential outcome had it being flipped, and the property's transaction potential outcome had it *not* being flipped. Thus, the causal effect of interest for a property transaction is defined as the difference among the potential outcomes for the same property transaction. However, given the fundamental problem of causal inference, I cannot estimate individual causal effects for the *same* property transaction (Angrist and Pischke, 2009). Instead, I estimate the average treatment effect (ATE) defined as:

$$\text{ATE} \quad = \quad\quad\quad = \quad\quad\quad \text{(IV.1)}$$

Where [ ] represents the expected value, and and are the potential outcomes for a property transaction, , while is the mean outcome for the flipped property transactions and is the mean outcome for the non-flipped property transactions.

Target Estimand (b) The second target estimands involve the pairwise comparison of four treatments. For every property transaction, , I denote these potential outcomes as ( ), where denotes the property's transaction potential outcome had it being located within HOLC zone A, B, C, and D, respectively. Given that I am dealing with multiple treatments, we then have 6 potential pairwise target estimands (McCaffrey et al., 2013). For example, the ATE between zone A and B would be defined as:

$$\text{ATE} \quad = \quad\quad = \qquad \text{(IV.2)}$$

Where, is the mean outcome for the property transactions located in zone A and is the mean outcome of property transactions located in zone B. In this study, we thus have 6 pairwise ATE, i.e., = ( ), = ( ), = ( ), (=), = ( ), and = ( ). We will have treatment effect heterogeneity if at least one ATE differs from another.

Estimators

Estimator (a) To provide an estimate of estimand IV.1, I use the following estimator, which seeks to understand the returns of both sides of a house flip and its characteristics, i.e., the back end of the flip and the front end:

$$= \quad + \quad \text{BEF} \quad + \quad \text{FEF} \quad + \quad\quad + \quad + \quad\quad + \quad + \qquad \text{(IV.3)}$$

Where, , is the outcome of property transaction, , during month, , and year, . As , I will use the natural logarithm of the sale price of the transaction, the type of payment (i.e., cash or mortgage), and the type of buyer (i.e., an investor or not an investor). BEF and FEF are dummy variables indicating whether the property transaction was the back end of a flip or the front end of the flip, respectively.[3] are property characteristics, that change over time, such as the age of the house and the deed type. , , and are property-, month-, and year-fixed effects, respectively. Standard errors are clustered at the property level. The coefficients of interest are and , which indicate the average treatment effect of a flip's back end and front end, respectively.

[3]Property transactions that weren't flipped have both dummy variables equal to zero.

As an approximation of the average total revenue per flipping transaction, I add, + , and I calculate the standard error of this linear additive estimate as, = + + 2 Cov . Where, and , are the standard errors of and from equation IV.3, while Cov is the covariance of both estimates calculated from the covariance matrix. I also present interactions between BEF and FEF with to show how the back end and front end of a flip evolve through time. The last results are presented as average marginal effects of BEF and FEF over yearly dummies. I also present the estimated average yearly revenue of a transaction by adding, + , by year.

Estimator (b) To provide estimates of the pairwise estimands IV.2, I use the following baseline estimator that seeks to model the probability of a property being flipped as a two-way fixed effects regression and each property's grade (A, B, C, or D) with the following functional form:

$$= + \quad 1\ \text{Grade} = \quad + \quad + \quad + \quad + \quad + \quad + \qquad \text{(IV.4)}$$

Where, , is the outcome of property transaction, , during month, , and year, . 1 Grade = represents an indicator variable taking the value of the assigned HOLC credit grade, where the property transaction lies—the HOLC grade "A" is the basis of comparison. and are characteristics of the property, , that change and do not change over time, respectively. , , and are zip codes, month and year-fixed effects, respectively.[4] Standard errors are clustered at the ZIP code level. Coefficients of equation IV.4 will help me test the hypothesis stating that relative to properties within areas formerly delineated by HOLC as greenlined (i.e., zone A), properties lying in formerly bluelined (i.e., zone B), yellowlined (i.e., zone C), and redlined (i.e., zone D) areas have a higher likelihood of being flipped. That is, the average treatment effects

[4]Given that HOLC grades do not change over time, one cannot control for property fixed-effects.

of these areas are 0, 0, and 0.

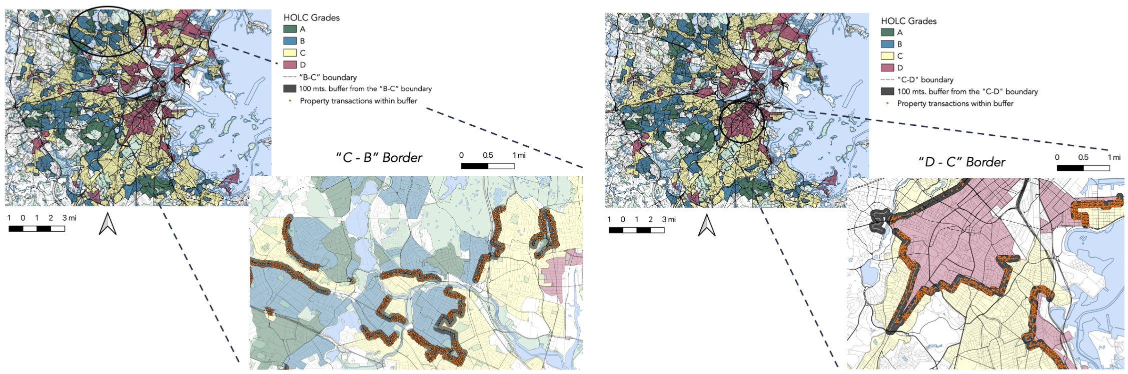
Spatial Regressions Discontinuity Design with Multiple Treatments. A key concern from estimator IV.4 is that the maps created by HOLC may have reflected and codified pre-existing differences in neighborhoods but were not associated with future disinvestment. A second concern with estimator IV.4 is that it doesn't fully control for potential confounders that affect the likelihood of an area receiving a specific HOLC grade and its contemporaneous likelihood of a property being flipped.

To overcome this limitation, I employed a Spatial Regression Discontinuity Design methodology, allowing a more precise comparison between neighboring housing units on either side of a tightly defined distance from a HOLC boundary, i.e., within a 100-meter buffer. The running variable in this approach is the distance from the centroid of each property transaction to the boundary lines of two adjacent HOLC zones. By considering changes over time in the probability of a transaction being a flip for almost adjacent property transactions located on either side of a HOLC boundary within a tightly defined geographic band, typically a few city blocks or even a few hundred meters, I remove potentially important but typically hard-to-measure factors influencing properties on both sides of a border. For instance, as noted by Aaronson, Hartley and Mazumder (2021), individuals residing in properties located within proximity to one another but separated by a boundary are likely to have similar access to local area amenities such as public transportation, retail stores, schools, and job opportunities. As these factors could potentially confound our treatments, the SRDD "controls" them by tightly comparing geographically proximate properties.

This method also allowed me to test for the impact of residing within a yellowlined zone, in addition to redlining. Figures IV.4 (a) and (b) present examples of the Spatial Regression Discontinuity Design Methodology for both "C-B" and "D-C" borders, respectively. Each dot in both figures represents a property transaction, while the black shaded area shows a 100 meters buffer around each HOLC border.

Figure IV.4: Spatial Regression Discontinuity Design Methodology

(a) “C-B” Border (b) “D-C” Border



The Spatial Regression Discontinuity Design Methodology has as a unit of observation the arm's length property transaction at the year level. Property transactions are organized as a panel, where every row of the data represents a property transaction per year. The estimator of this strategy has the following functional form:

$$= \quad + \quad + \quad (\quad \text{Dist} \quad) + \quad + \quad + \quad + \quad + \quad + \qquad \text{(IV.5)}$$

Where, , is the outcome of property transaction, , located on boundary, , during year, . For instance, , will represent a dummy variable indicating whether the property transaction, , located on the boundary, , during the year, , was a flip. Dist represents the distance, in meters, of a property transaction, , centroid to the boundary, . The sign of Dist is determined by the change in how HOLC “perceived” zones between two adjacent HOLC-graded neighborhoods. Properties falling within a lower “HOLC-perceived” neighborhood relative to the adjacent neighborhood take on a negative value. For example, when comparing a “D” graded neighborhood vs. an “A” graded neighborhood, those properties within the “D” graded neighborhood will have a negative Dist sign. is a dummy variable representing whether the property transaction, , lies on one side of the boundary; for example, whether the property lies on the “D” side of the “C-D” boundary. () represents a flexible polynomial that uses both Dist and . and are characteristics of the property, , that change and do not change over time, respectively. , , and are zip code, month, and year fixed effects, respectively. are the residuals. Standard errors are clustered at the boundary level. represents the average treatment effect, and it provides an estimate of the estimand in equation IV.2. The identifying assumption behind this approach is that neighborhood and housing quality varies continuously across HOLC borders, while HOLC's past grades shift discontinuously at the boundaries.

In equation IV.5, takes the form of a dummy variable, representing whether the property transaction, , lies on one side of the boundary. Given that there are four HOLC categories (i.e., A/"Best" = green; B/"Still Desirable" = blue; C/"Definitely Declining" = yellow; and D/"Hazardous" = red), I will focus on four different treatments. Those are boundaries of HOLC areas A (i.e., = 0) and B (i.e., = 1); boundaries of HOLC areas A (i.e., = 0) and C (i.e., = 1); boundaries of HOLC areas B (i.e., = 0) and C (i.e., = 1); and boundaries of HOLC areas C (i.e., = 0) and D (i.e., = 1). To calculate from estimator IV.5, I use Calonico, Cattaneo and Titiunik (2014)'s estimator to implement a local polynomial regression discontinuity estimation with robust bias-corrected confidence intervals and inference procedures. For each outcome, I provide three different procedures: (i) conventional RD estimates with a conventional variance estimator; (ii) bias-corrected RD estimates with a conventional variance estimator, and (iii) bias-corrected RD estimates with a robust variance estimator.

Spatial Regressions Discontinuity Design with Multiple Treatments and Propensity Scores. As Aaronson, Hartley and Mazumder (2021), Bayer, Ferreira and McMillan (2007), and Dhar and Ross (2012) have expressed, a SRDD may still fail to satisfy the assumption of balancing for pre-treatment covariates or continuity across the border. For instance, Aaronson, Hartley and Mazumder (2021) show that before the HOLC maps were introduced, "D" and "C" zones had a higher proportion of African American residents, lower housing values and rents, and lower homeownership rates compared to

"A" and "B" zones. To reduce potential confounding, I complemented my SRDD with inverse probability (IP) weighting to balance pre-treatment covariates and make the continuity assumption more plausible. Assuming conditional exchangeability, weighting the SRDD by the IP of being assigned to a treatment would produce unbiased estimates of estimand IV.2. For this method to be effective, the critical task is to identify a function, ( ), that can accurately predict the treatment variable, , which, in our case, is the HOLC grades, using observable characteristics, , of each zone.

Previous studies have also employed IPW to address confounding, often in conjunction with other natural experiments, such as the difference-in-differences method (e.g., Xu (2022)). However, these studies typically employ a binary treatment, treating redlining as the sole treatment and assuming that residing in yellowlined, bluelined, and greenlined areas had the same impact when comparing it against residing in redlined areas. This approach ignores the fact that, as Aaronson, Hartley and Mazumder (2021), Aaronson et al. (2022), and Wassmer (2023) explain, formerly yellowlined areas experienced similar harms as formerly redlined zones. Failure to account for this distinction would underestimate the impact of HOLC maps. Moreover, other studies that rely upon IPW methods often use pre-HOLC census data and geographic approximations to calibrate, ( ); however, as Rothstein (2017) explain, the realtors and assessors who provided input for HOLC maps frequently relied on subjective and sometimes fabricated approximations of neighborhoods' appearances. Therefore, using census data to calibrate ( ) would lead to biased estimates. Lastly, much of the literature on IPW assumes linear approximations of the function, ( ), which can limit the estimator's ability to generate unbiased estimates.

To address these limitations, I employ three strategies. Firstly, I utilize estimand IV.2, which considers multiple treatments. Secondly, instead of relying on census data as previous studies to fit, ( ), I employ the data used to assign risk grade available for most cities in their "area description" as features, , to predict HOLC grades, which was coded and organized by Markley (2023). Figures C.1 (a) and (b) show two examples of area description files in Boston, MA. From these area description files, I used the Black population percentage, the "foreign-born" population percentage, the median family income, the occupation class (categorized and one-hot encoded from "Low Mid," "Lower," "Mid Mix," "Up Mid," and "Upper"), the average property age, the home repair status (categorized and one-hot encoded from "Fair," "Fair-Good," "Fair-Poor," "Good," and "Poor"), and the mortgage availability (categorized and one-hot encoded from "Fair," "Fair-Good," "Fair-Poor," "Good," and "Poor"). Finally, I utilize a generalized boosted model with 3,000 trees to calibrate a flexible function ( ), using absolute standardized mean difference and Kolmogorov-Smirnov statistic as stopping rules. With this new flexible function and the "area description" variables, I predict a propensity score for each observation and use it as a weight for the SPRDD presented in equation IV.5.

III.III Collaborative Data Analysis

The Collaborative Data Analysis (CDA) occurred at the Conservation Law Foundation on Saturday, April 1st, 2023. The CDA began with a presentation from one of MIT's academic researchers to the resident researchers of the main findings of the research. The purposes of CDA are for the HNS research team to jointly interpret data, drawing on the different forms of expertise different researchers bring to the study. CDA activities are, therefore, internal peer-to-peer researcher discussions about results, not efforts to collect data from resident researchers about the phenomenon under study. To present findings at CDA, I divided the main findings into five sections, i.e., (i) What is house flipping? (ii) What housing is being flipped? (iii) Why does house flipping occur? (iv) Who is flipping? and How? and (v) Where is house flipping occurring?

After presenting the results to CDA participants, we engaged in an activity around the question: "Where is house flipping occurring?" The main objective of this activity was to understand block-level conditions in HNS sites and discuss how communities are changing locally in specific places where properties are being flipped at a higher rate. For each HNS neighborhood, we designed and printed HNS maps with 250 meters by 250 meters hexagon grid where each hexagon represented the percentage of the total flipped transactions over the study period within that hexagon (see Figure C.11 for individual maps used during the activity). We color-coded each grid with the following labels and colors, i.e., 0% flipped transactions (blue color), 0 - 25% flipped transactions (orange color), and 25 - 100% (blue color). We printed each map and placed them on the wall.

We then looked at maps of HNS neighborhoods that depicted the location of flipping hotspots and discussed why they are hotspots and what the consequences of flipping have been. We used pink stickies to point out and write down the harms and green stickies to point out and write down any benefits associated with house flipping hotspots. We tried to be as specific as possible about who is being harmed and how, and who is benefiting and how? After the mapping exercise, we engaged in a group activity where researchers from different sites discussed common threads across communities related to house flipping and discussed house flipping as a system of benefit and harm. The group discussion covered questions such as: How does the house flipping system work? What results does it produce? and where is the best place to intervene if needed? We documented the discussion and used insights from it to inform the discussion section of this paper.

IV Results

Understanding the behavior of housing investors and their impact on the housing market is a complex question that requires empirical research. I begin analyzing the spatial and temporal patterns of house flipping in HNS neighborhoods, which involves buying and reselling a property for a profit. Then, I explore the association of historical disinvestment with the likelihood of a house being flipped nowadays. Lastly, I present the results of the collaborative data analysis to better understand the local living experiences of house flipping in rapidly changing neighborhoods.

IV.I House Flipping Trends in the Greater Boston Area

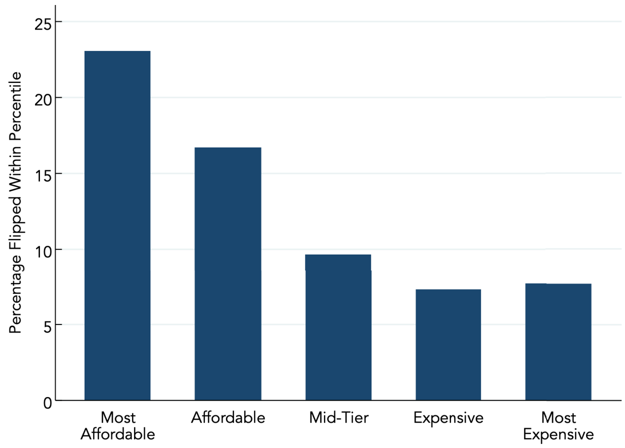
Previous studies on housing investors and house flipping, such as Depken, Hollans and Swidler (2009) and Li, Yavas and Zhu (2023), have used historical data to examine the trends and prevalence of this practice over time. They have documented the abnormal returns made by flippers and how their investments vary over the housing market cycle. However, while these studies highlight the positive returns associated with house flipping, they provide limited insight into how such activity impacts affordability or how flipping patterns occur across the price distribution.

As seen in Figure IV.5, house flipping occurs more often at the lower price spectrum of the price distribution, i.e., within the most affordable units. Between 2008 and 2021, 23.05% of all the affordable units in HNS neighborhoods were flipped. As the price percentile increased, the percentage of flips decreased rapidly until it remained almost constant after the 4^{th} and 5^{th} quintiles, i.e., expensive and most-expensive housing units. For example, 16.7% of all the affordable housing transactions, i.e., 2^{nd} price quintile, were flipped, whereas approximately 7.2% of housing transactions within the expensive price percentile were flipped.

In other words, between 2008 and 2021, within HNS neighborhoods, the most affordable and affordable flipped units accounted for 36.14% (3,888 housing units) and 26.00% (2,797 housing units) of all flipped units, respectively. When mid-tier house flips are included (14.93%), almost 80% of all house flips were from the most affordable, affordable, and mid-tier housing units.

Figure IV.5: Flipping Rate of Residential Properties, by Price Affordability

Percentage of Total Transactions within each Quintile



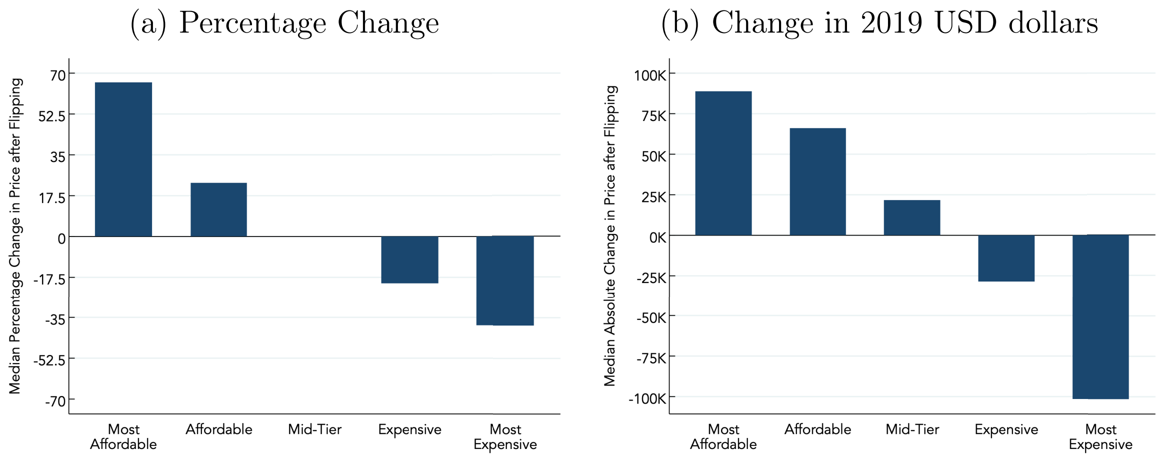
Note: Price quintiles are calculated within the city. Most affordable refers to the bottom quintile of the sale prices within a neighborhood, affordable refers to the next quintile of the sale prices within a neighborhood, and so on, until most expensive refers to the top quintile of the sale prices within an HNS neighborhood.

As house flipping remains higher among the cheapest housing in HNS neighborhoods, it increases the concern regarding its impact on affordability and access to homeownership. In other words, does house flipping crowd-out first-time and affordable homebuyers by reducing the supply of affordable housing?

According to Figures IV.6 (a) and (b), house flipping in HNS neighborhoods was a profitable venture between 2009 and 2021, with positive returns in both percentage and absolute terms for most flipped units in the most affordable, affordable, and mid-tier price ranges. However, the median gross profit and return on investment were not distributed evenly across the price spectrum, and both declined as the original price of the flipped property increased. In fact, the expensive and most expensive housing units resulted in negative median gross profits and return on investment.

The median gross profits for the most affordable, affordable, mid-tier, expensive, and most expensive house flips were \$88,660, \$65,807, \$21,476, -\$28,687, and -\$101,514, respectively. Similarly, the median return on investment was 65.9%, 22.8%, 0.0%, -20.2%, and -38.6%, in the same order.

Figure IV.6: Median Change in the Price of Flipped Houses, by Price Affordability



Note: Price quintiles are calculated within the city. Most affordable refers to the bottom quintile of the sale prices within a neighborhood, affordable refers to the next quintile of the sale prices within a neighborhood, and so on, until most expensive refers to the top quintile of the sale prices within an HNS neighborhood. Prices in 2019 dollars.

The legal nature of the house flipper (i.e., investor or individual) matters for the functioning of the real estate market. The legal structure of investors (e.g., limited liability company (LLC)) protects from personal liability if something goes wrong during the flip. Individuals and investors could also face different tax implications when selling the flip. For example, investors can be taxed at different rates[5], deduct higher expenses from the flip sale[6], or take advantage of section 1031 exchanges after the property sale.[7]

Within HNS neighborhoods, 30.53% of total transactions that were eventually flipped (i.e., the back-end of the flip) were done by investors, whereas 10.3% of flipped transactions (i.e., the front-end of the flip) were done by investors. Irrespective of the initial sale price, flipped housing units are bought at a similar rate by investors (as seen in

[5]The profits from a house flip are generally taxed as ordinary income for individuals, which means they are subject to the same tax rates as wages and salaries. The tax rate on regular income can range from 10% to 37%, depending on the income earned. However, suppose the flip is a capital asset held for over a year. In that case, the profits may be taxed at the lower capital gains tax rate, generally 0%, 15%, or 20%, depending on the taxpayer's income level.

[6]An investor may have more flexibility in deducting expenses because they may be able to structure the flip as a business and deduct additional costs, such as office supplies, travel expenses, and professional fees.

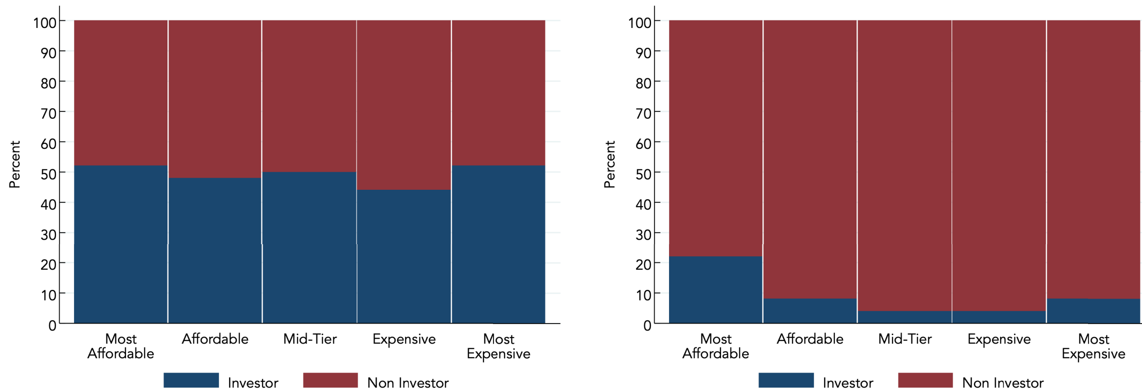
[7]A Section 1031 exchange, also known as a like-kind exchange, allows investors to defer paying taxes on the profits from the sale of an investment property by reinvesting the proceeds in another investment property.

Figures IV.7 (a) and (b)).

Figure IV.7: Type of Flipper, by Price Quintile

Percentage of Total Transactions within each Quintile

(a) Seller to Flipper (b) Flipper to Buyer



Note: An investor represents a buyer whose full name contains either Association, Trustee, Company, Limited Partnership, Joint-Ventures, or Corporate Trust. Price quintiles are calculated within the year and city. Most affordable refers to the bottom quintile of the sale prices within a neighborhood, affordable refers to the next quintile of the sale prices within a neighborhood, and so on, until most expensive refers to the top quintile of the sale prices within an HNS neighborhood. Prices in 2019 dollars.

Another pathway in which flipping can have a negative impact on the housing supply for first-time or low-income homebuyers is through the payment type used to purchase the property. Cash payments by investors, compared to traditional mortgages used by typical homebuyers, can create more favorable scenarios for them to acquire properties. Since securing a mortgage can be time-consuming and difficult, house owners may prefer to sell to buyers with ample liquidity to avoid potential issues with the buyer's ability to obtain a mortgage.

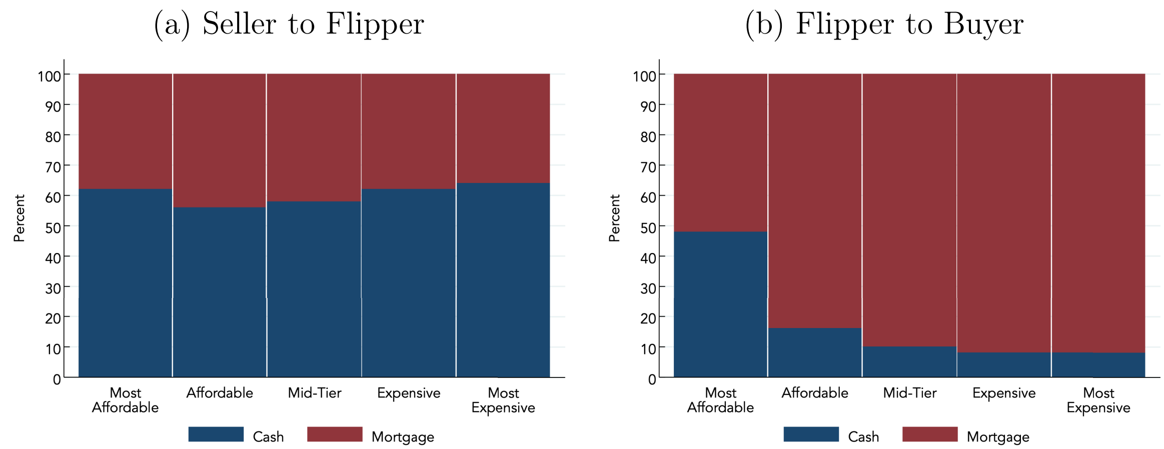
Cash payments can also have a broader impact on affordable housing. Investors purchasing affordable units with cash can worsen the existing affordable housing shortage. Increased demand for affordable units by investors may make it difficult for low- and moderate-income households—that rely on private and public mortgages to secure reasonably priced housing—to outbid cash investors, especially during a thick housing market period. Figure IV.8 (a) shows how flippers pay for a property with either cash or mortgage, while figure IV.8 (b) shows how homebuyers pay flippers to acquire a property. Both figures are ordered by the sale price quintile of the flip, indicating the percentage of total transactions within each quintile paid for with cash or mortgage. When analyzing cash purchases by price affordability, heterogeneity in the payment type emerges.

Figure IV.8 (a) reveals that cash payments are the dominant form of payment for house flippers, accounting for 62% of all back-end flips. Even for the most affordable flipped units, cash payments were used in nearly 62.5% of transactions. Remarkably, cash payments remained the preferred mode of payment for flipped properties, representing almost 64% of all transactions for the most expensive properties. This trend could negatively impact the housing market's middle-income segment by driving up prices and reducing affordability. As competition for high-value properties increases, finding suitable properties may become challenging for middle-income and high-income homebuyers, leading them to remain in their current homes for longer. This could, in turn, slow down the market's filtering process.

In contrast, buyers of flips display a different payment preference compared to flippers, as shown in Figure IV.8 (b). Buyers of the most affordable flips paid approximately 48% of all flipped transactions with cash. However, for mid-tier, expensive, and most expensive flips, cash payments were made only 10% of the time.

Figure IV.8: Type of Payment, by Price Quintile

Percentage of Flips within each Quintile



Note: Price quintiles are calculated within the city. Most affordable refers to the bottom quintile of the sale prices within a neighborhood, affordable refers to the next quintile of the sale prices within a neighborhood, and so on, until most expensive refers to the top quintile of the sale prices within an HNS neighborhood. A cash offer is one where the buyer doesn't use a mortgage to purchase the house.

To better understand the value added by house flippers in the housing market, it is crucial to examine whether they conducted repairs or additions to the flipped properties. For instance, flippers may revitalize neighborhoods by repairing blighted properties and making them habitable again. In such cases, the externalities of their work could extend beyond individual homes and contribute to broader community development. However, flippers may also engage in practices that give them an advantage in the housing market without enhancing the housing stock. For example, they may use all-cash transactions to outcompete other buyers but not invest in additional repairs or improvements to the property. In other words, flippers speculate on the direction of the housing market without investing in revitalizing the house.

Among flipped transactions, 33.6% had at least one approved property permit between the time a flipped property was bought and eventually re-sold, i.e., when the flipper owned the house. In other words, 66.4% of the total flipped transactions never underwent any repair or addition that required an approved property permit within HNS

Boston neighborhoods. Furthermore, among the flipped transactions paid initially in an all-cash manner, only 2.89% of them had at least one approved property permit when the flipper owned the house. Lastly, over half (53.9%) of the flipped transactions did not have any approved property permit during the flipper's ownership were also paid for with all-cash. The last suggests that a substantial portion of the flipped properties did not undergo any significant repairs or additions that required approval and were likely bought by flippers to resell them for a profit quickly rather than make improvements to the properties.

The Effects of House Flipping on the Real Estate Market

Table IV.1 presents estimates of coefficients and using estimator IV.3. Both estimates provide the average percentage difference in sale prices for property transactions, with the first one capturing the difference for property transactions on the back end of a flip (i.e., BEF) and the second one for properties on the front end of a flip (i.e., FEF), relative to non-flips. Two statistically significant price results are worth highlighting.

On the one hand, properties flipped (i.e., BEF) get purchased at an average *discount* of = 31 2% (Table IV.1, row 1 of column (5)); saying it differently, the average percentage difference in purchase price between a flipped house and an almost identical property that has not been flipped is about -27.9%. On the other hand, properties flipped (i.e., FEF) sell at an average percentage price difference of = 8 8% (Table IV.1, row 3 of column (5)); in other words, the total average percentage difference in sell price, i.e., + , between a flipped house and an almost identical property that has not been flipped is about +19.1%.

The payment type also differed between the back end of a flip and the front end (Table C.1). On average, properties flipped (i.e., BEF) have a 21.9% higher probability of being paid in cash (Table C.1, row 1 of column (5)). In a similar way, flipped properties (i.e., FEF) have a 2.9% more probability of being paid in cash (Table C.1, row 3 of column (5)). Regarding investor purchases, back-flipped housing transactions were 18.2% more likely to be bought by an investor than non-flipped ones, whereas front-flipped sales were -7.1% less likely to be bought by an investor than non-flipped transactions (Table C.2, rows 1 and 3 of column (5)).

Table IV.1: Hedonic Regression and House Flipping, Elasticity of the Sale Price

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Back End of the Flip	-.345	-.263	-.211	-.223	-.279
	(.006)	(.006)	(.006)	(.007)	(.010)
Front End of the Flip	-.296	-.214	-.161	-.141	-.088
	(.006)	(.006)	(.005)	(.005)	(.009)
Month & Year Fixed Effects	No	Yes	Yes	Yes	Yes
Zip Codes Fixed Effects	No	No	Yes	Yes	No
House Characteristics	No	No	No	Yes	Yes
Property Fixed Effects	No	No	No	No	Yes
# of Observations	93,932	93,932	93,919	90,047	56,037

Note: Coefficients are in the form of β from equation IV.3. Standard errors are clustered at the property level. * p 0.05, ** p 0.01, and *** p 0.001. House characteristics include the property's age, type of deed, living area of the property in square feet, size of the property in square feet, number of bedrooms in the property, number of bathrooms in the property, and number of parking spaces.

Effects of House Flipping Over Time

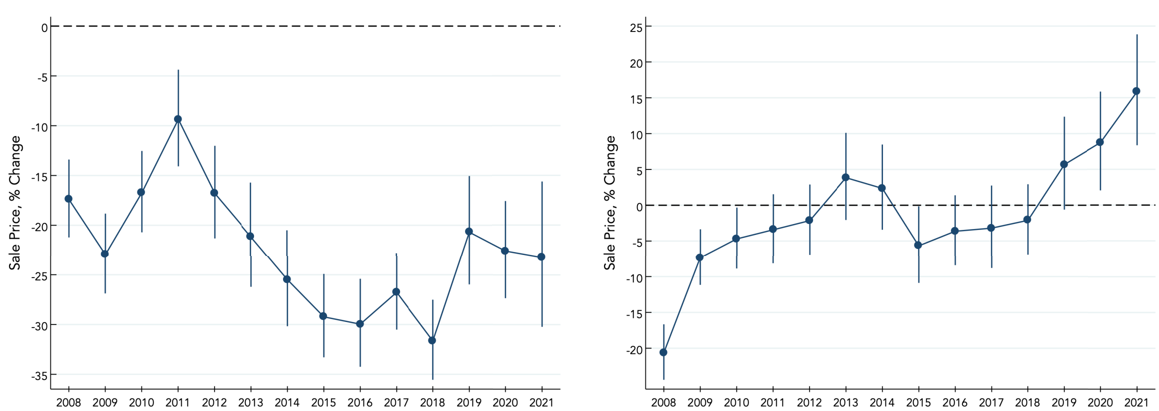
Before moving to the next section, let's discuss how house flipping impacted the real estate market over time. Figure IV.9 presents estimates of how house flipping discounts ( ) and premiums ( ) evolved between the years 2009 and 2021. Between 2009 and 2021, the house flipping discount was always negative even though it decreased as time passed, going from -36.8% in 2009 to -19.2% in 2019. The decrease in the discount could potentially be attributed to the ability of house flippers to locate distressed properties in the wake of the 2009 Financial Crisis, which offered higher returns on investment. As the availability of distressed properties decreased, the average discount at which house flippers could purchase properties also decreased.

A similar trend was observed for the house flipping premium. As time passed, house flipping premiums increased, ranging from a negative premium of -7.03% in 2009 to a positive one of 18.72% in 2019. One possible explanation for this pattern is similar to the explanation for the discount. The rise in house flipping premiums coincided with the Financial Crisis. After the Financial Crisis, a flood of properties was on the market, which led to negative returns for those who flipped houses, possibly due to oversupply. However, as the real estate market recovered from the crisis, the supply of properties decreased, resulting in an increase in the house flipping premium.

Figure IV.9: Hedonic Regression and House Flipping,
Percentage Change in Sale Price by Year

Average Marginal Effects by Year

(a) Back End of the Flip (b) Front End of the Flip



Note: Vertical lines crossing the estimates are confidence intervals, where the cap represents the confidence interval at the 95% level.

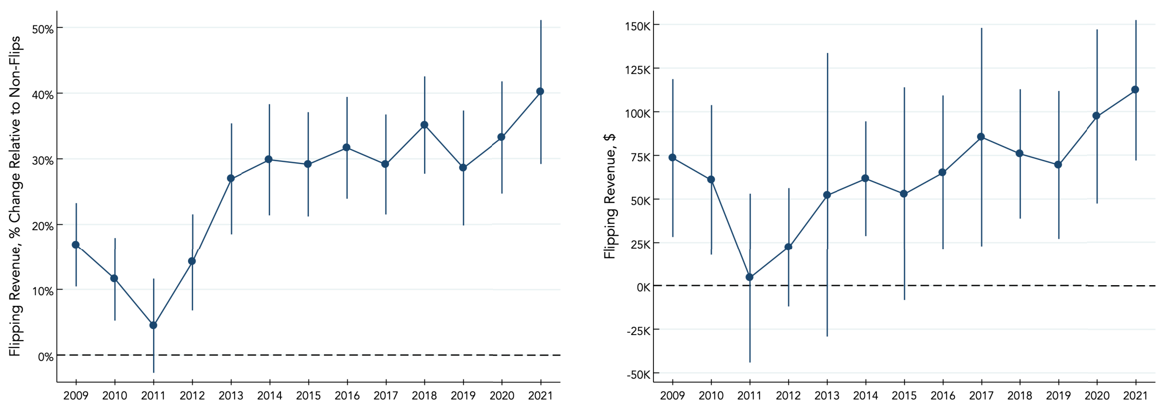
Figures IV.10 (a) and (b) display the average marginal revenue perceived by a flipped transaction relative to a non-flip in both percentage and absolute terms. This includes the back-end and front-end coefficients from equation IV.3. Both the total revenue in percentage and absolute terms have increased since the beginning of the study. In

2009, the entire flipping transaction was sold for 16.8% more than a similar non-flip transaction, reaching 40.1% more in 2021, as shown in Figure IV.10 (a). In dollar terms, the entire flipping transaction was sold for $75,000 more than a similar non-flip transaction, reaching $120,000 more in 2021, as seen in Figure IV.10 (b).

Figure IV.10: Hedonic Regression and House Flipping,
Change in Total Revenue by Year

Average Marginal Effects by Year

(a) Percentage Change (b) Total USD Dollars



Note: Vertical lines crossing the estimates are confidence intervals, where the cap represents the confidence interval at the 95% level.

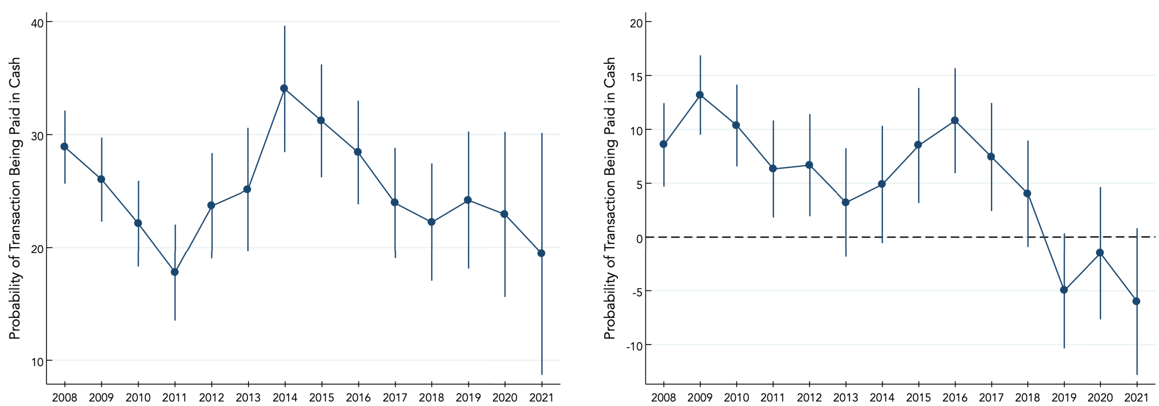
As discussed in the previous section, it is important to determine whether flippers add value to the housing market, as they can play a role in revitalizing urban blight and investing in deteriorating housing stock. As shown in Figure IV.10 (a), within HNS neighborhoods, house flippers earned a statistically significantly higher revenue, on average, compared to non-flippers. However, this finding is altered when we distinguish between flippers who filed at least one property permit and those who did not. Figure C.2 (a) and (b) display the revenue outcomes for flippers in Suffolk County who did not file a permit and those who did, respectively. In both graphs, one notices that those flippers that didn't file a permit and didn't physically alter something of the property (Figure C.2 (a)), had statistically significantly higher revenue, on average, compared to non-flippers. This result isn't statistically significantly different from zero for those flippers that did file a property permit (Figure C.2 (b)).

Mechanisms Over Time. The probability of paying with cash for flipped properties also changed as time passed but always remained positive for the back end of the flip (Figure IV.11). The likelihood of a back end of a flipped paid with cash went from 28.8% in 2009 to 19.4% in 2021, remaining positive throughout every year of the study period (Figure IV.11(a)). In contrast, the probability of paying the front end of a flip with cash turned negative as years passed, reaching -5.9% in 2021 (Figure IV.11(b)).

Figure IV.11: Hedonic Regression and House Flipping, Probability of Paying with Cash by Year

Average Marginal Effects by Year

(a) Back End of the Flip (b) Front End of the Flip



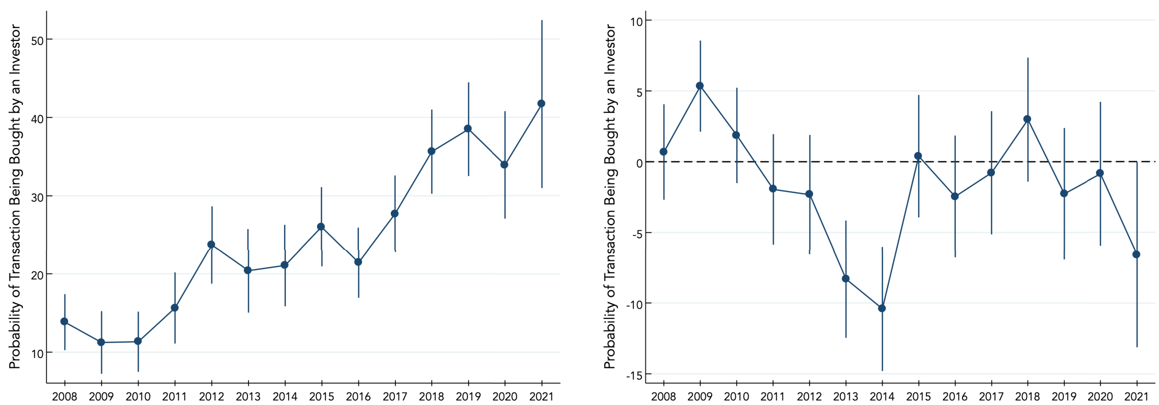
Note: Vertical lines crossing the estimates are confidence intervals, where the cap represents the confidence interval at the 95% level.

The average marginal probability of a house eventually being flipped (i.e., the back end of the flip) and bought by an investor increased at a steep rate between 2008 and 2021 (Figure IV.12(a)). In 2008, a house being flipped was 13.8% more likely to be bought by an investor relative to non-flipped houses, whereas in 2021, this difference reached 41.6%. In contrast, the average marginal probability of a house flipped (i.e., the front end of the flip) and bought by an investor wasn't statistically significantly different from non-flipped houses across time (Figure IV.12(b)).

Figure IV.12: Hedonic Regression and House Flipping,
Probability of Buyer Being an Investor by Year

Average Marginal Effects by Year

(a) Back End of the Flip (b) Front End of the Flip



Note: Vertical lines crossing the estimates are confidence intervals, where the cap represents the confidence interval at the 95% level.

IV.II Tracing the Roots of Modern-Day House Flipping back to Early 20$^{\text{th}}$ Century Disinvestment

Neighborhoods that have historically faced economic and financial disinvestment[8] currently experience lower health and economic opportunity (Chetty, 2021; Krieger and Higgins, 2002). However, disinvested neighborhoods with significant intrinsic residential amenities and abundant residential floor space may offer profit opportunities for investors and house flippers through the following path. Historical conditions of disinvestment may have created a slow historical decline in home values, thereby elevating the risk that mortgaged property owners may find themselves in a negative equity situation, owing more than their property's market worth (Glaeser and Gyourko, 2005). Properties with market values that fall below replacement costs are less likely to un-

[8]Defined here as a lack of adequate access to capital and a stagnation in economic growth

dergo maintenance and improvement (Gyourko and Saiz, 2004; Haughwout, Sutherland and Tracy, 2013; Melzer, 2017). These properties, due to their suppressed acquisition costs and potential for higher resale value, particularly in neighborhoods rich with inherent amenities, then become attractive for investors and house flippers. Despite the potential relationship between historical urban disinvestment and house flipping, there are no papers to my knowledge linking them both. This chapter section explores the association between 20$^{\text{th}}$ century patterns of disinvestment and 21$^{\text{st}}$ century home flipping.

Housing dynamics during the 20$^{\text{th}}$ century were characterized by periods of investment and disinvestment. For example, the Stock Market Crash of 1929 and the subsequent Great Depression served as catalysts of neighborhood and housing disinvestment (Fishback et al., 2020). At the peak of the Great Depression, almost 40% of home borrowers fell behind their mortgages, effectively freezing mortgage and money markets and thereby potentially amplifying the financial catastrophe without any further action (Rothstein, 2017). In this context, in 1933, under Franklin D. Roosevelt's presidency, the Federal Government established the Home Owners Loan Corporation (HOLC) to stabilize the mortgage market by buying and refinancing home loans at risk of default. By 1936, the HOLC had effectively incorporated nearly a million distressed assets into its portfolio. In 1935, long after HOLC started stabilizing the mortgage market, the Federal Home Loan Bank Board established a program called the City Survey program (Fishback et al., 2020). In this program, HOLC staff, alongside local real estate professionals, including local bank loan officers, city officials, and realtors, conducted surveys to understand and assess lending practices, real estate risk levels, and assets held by HOLC in more than 230 cities (Hillier, 2003).

This survey led to the creation of HOLC descriptions of neighborhoods, known as area description files (an example of which is illustrated in Figure C.1), which were subsequently utilized to come up with real-estate residential risk grades and create maps for each neighborhood, making both available only to government officials involved in housing policy (Hillier, 2003). The area description files and maps classified neighborhoods into four zones (i.e., zone D, "Hazardous;" zone C, "Definitely Declining;" zone B, "Still Diserable;" zone A, "Best") based on detailed risk-based characteristics, including housing age, quality, occupancy, mortgage access, and home values. Alongside housing characteristics, non-housing attributes such as race, ethnicity, and immigration status of the neighborhood were also collected (Aaronson, Hartley and Mazumder, 2021).[9] In essence, the HOLC, along with other Federal Agencies' surveys and research initiatives, illustrated the disinvestment trends of the 1930s, which were steeper in neighborhoods with a higher population of African Americans and immigrants, as demonstrated in Table IV.1. In other words, the HOLC surveys and their subsequent maps described the existing conditions across different neighborhoods, and they did a good job of identifying the border between already-racialized areas (Hillier, 2003).

This naturally prompts us to consider how to spatially quantify disinvestment (and, more importantly, racially-driven disinvestment) in the 20$^{\text{th}}$ century to show its association with 21$^{\text{st}}$ century house flipping. Ideally, I would leverage the FHA maps and their ratings which categorized mortgages based on their risk level, thereby encouraging private lenders seeking insurance for their mortgages to adopt these classifications (Fishback et al., 2020). However, since the FHA maps were potentially destroyed around the 1970s (Sagalyn, 1980), I will use recently digitized maps by the University of Richmond with the caveat that these maps did not directly influence the HOLC's primary goal

[9] Around the same time that HOLC was established, the Federal Government also founded the Federal Housing Administration (FHA) in 1934 as an independent agency with the primary purpose of insuring loans for home maintenance, rehabilitation, and new construction. As HOLC, the FHA embarked on their studies of local markets and created their own set of neighborhood color-coded risk maps to "avoid insuring risky properties to keep foreclosure rates down" (Fishback et al., 2020). This increased focus on location represented an additional barrier to capital in these already distressed neighborhoods and had substantial negative impacts on Black homeowners. Historians generally agree that the FHA avoided insuring mortgages for potential borrowers who were Black or lived in Black neighborhoods (Michney and Winling, 2020; Fishback et al., 2021). As Hillier (2003) explains, the FHA was a leader in establishing and promoting standards and procedures for neighborhood appraisals. FHA did more to institutionalize racialized disinvestment than any other agency by categorizing mortgages according to their risk level and encouraging private lenders who wanted mortgage insurance to do the same.

of stabilizing the mortgage market in the wake of the Great Depression nor causally created the practice of "redlining" (Hillier, 2003; Fishback et al., 2020). However, they correlate with and offer a snapshot of patterns of racialized disinvestment and real estate valuation that existed before their creation by the HOLC. In other words, the HOLC maps used in this chapter show the striking extent of disinvestment experienced within African American and immigrant communities in Boston during the early twentieth century.

Historical Disinvestment and Modern Day House Flipping

Associations Between Area Description Files and HOLC Grades. I begin by exploring the associations between the area description files, which were used to create HOLC's maps, and the grades allocated to each neighborhood. Each column of table IV.1 presents the results of regressing a dummy variable—assigned the value 1 if a neighborhood was given a grade of "B," "C," or "D" by the HOLC, and 0 if it received an "A"—against the variables contained within the area description files for each graded neighborhood (see Figures C.1 for examples of two such files).

On average, the presence of Black residents in a neighborhood was found to increase the likelihood of a neighborhood being graded with a letter D (i.e., redlining) compared to being assigned a letter A (i.e., greenlining), but this effect was not statistically significant for neighborhoods with grades C (i.e., yellowlined) or B (i.e., bluelined). In contrast, the presence of foreign-born residents in a neighborhood increased the likelihood of a neighborhood being assigned a letter D, C, or B relative to being assigned a letter A. As well, the average age of properties in a neighborhood was statistically significant in determining whether a neighborhood was assigned with a letter D, C, or B relative to being assigned a letter A. Neighborhoods with lower income were more likely to be assigned a letter D or C relative to being assigned a letter A, suggesting that neighborhoods assigned with letters B and A had similar economic characteristics.

A “higher” occupation class in a neighborhood was also correlated with having a higher likelihood of being assigned a letter A rather than D, C, or B. A similar pattern was observed regarding the mortgage availability and home repair status of housing. Better mortgage availability and lower home repair status were associated with a higher likelihood of being assigned a letter A rather than D, C, or B.

In essence, neighborhoods with Black and foreign-born populations, older properties, lower family income, lower perceived occupation class, and poorer mortgage availability and home repair status were more likely to be assigned a letter D or C rather than a letter A. These results support Hillier (2003)'s and Fishback et al. (2020)'s work explaining that HOLC's maps did not automatically redlined neighborhoods. Instead, they provided a comprehensive snapshot of pre-existing disinvestment before the commencement of HOLC's City Survey program. Furthermore, these results support my argument for utilizing HOLC polygons as proxies to represent disinvestment during the early 20$^{\text{th}}$ century.

Table IV.1: Probability of a HOLC Zone Being Graded B, C, or D

	Zone B	Zone C	Zone D
	(1)	(2)	(3)
		Relative to Zone A	
Zone Has Black Population	0.074	-0.009	0.022***
	(0.041)	(0.010)	(0.005)
Zone Has Foreign Born Population	0.104***	0.049***	0.033***
	(0.020)	(0.008)	(0.006)
Median Family Income ($)	-0.000	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
Median property Age (Years)	0.010***	0.004***	0.001***
	(0.001)	(0.000)	(0.000)
Base = "Lower" Occupation Class			
"Low Mid" Occupation Class	0.118*	0.001	0.011
	(0.049)	(0.008)	(0.006)
"Mid" Occupation Class	0.177***	0.027**	0.009
	(0.040)	(0.009)	(0.010)
"Up Mid" Occupation Class	0.134***	-0.030*	-0.231***
	(0.040)	(0.015)	(0.040)
"Upper" Occupation Class	-0.087*	-0.264***	-0.240***
	(0.038)	(0.019)	(0.034)
Base = "Poor" Home Repair Status			
"Fair" Home Repair Status	0.058	0.063***	0.036***
	(0.061)	(0.015)	(0.008)
"Fair-Good" Home Repair Status	0.104	0.110***	-0.016
	(0.060)	(0.017)	(0.025)
"Fair-Poor" Home Repair Status	-0.009	0.017	0.019***
	(0.065)	(0.015)	(0.005)
"Good" Home Repair Status	-0.075	-0.233***	-0.441***
	(0.062)	(0.024)	(0.052)
Base = "Poor" Mortgage Availability			
"Fair" Mortgage Availability	0.005	0.010	0.011*
	(0.049)	(0.009)	(0.005)
"Fair-Good" Mortgage Availability	-0.158**	-0.021	-0.116**
	(0.061)	(0.023)	(0.042)
"Fair-Poor" Mortgage Availability	0.029	-0.009	0.006
	(0.101)	(0.016)	(0.007)
"Good" Mortgage Availability	-0.121*	-0.120***	-0.236***
	(0.049)	(0.014)	(0.036)
# Observations	2,415	3,420	2,314

Note: Standard errors (in parentheses) are clustered at the state level. * p 0.05, ** p 0.01, and *** p 0.001. Each column represents a different regression.

Associations Between HOLC Grades and House Flipping. Having shown the association between HOLC grades and historical disinvestment, I now test the association between historical disinvestment (*proxied* here as neighborhoods with a D or C grade) and modern-day house flipping practices. Table IV.2 shows estimates of estimator IV.4, supporting the hypothesis that properties in formerly graded D and C neighborhoods are more likely to be flipped than those in former A-graded neighborhoods. Specifically, in Table IV.2, column (4), I find that properties in former redlined (zone D, "Hazardous") and yellowlined (zone C, "Definitely Declining") areas are more likely to be flipped than those in former greenlined areas (zone A, "Best").

Controlling for the year, zip code, and house characteristics, properties in former redlined and yellowlined areas are, on average, 1.7% and 1.0% more likely to be flipped than those in greenlined areas (Table IV.2 (1)-(4)). However, I found no statistically significant difference in flipping rates between greenlined and bluelined (i.e., "Still Desirable") areas. These results are also robust to including fixed effects for the year, zip code, and house characteristics.

Furthermore, properties located within redlined and yellowlined areas not only have a higher likelihood of being flipped but also sell at lower prices and are more prone to foreclosure compared to those in greenlined areas (Tables C.3 and C.5 in the appendix show the results). Specifically, properties in former redlined and yellowlined areas sell for 32.4% and 30.1% less, respectively, than those in former greenlined areas on average (Table C.3, column (4)). Additionally, properties in these same areas are, respectively, 0.23% and 0.41% more likely to be foreclosed than those in former greenlined areas (Table C.5, column (4)). There is no statistically significant association between properties located in redlined or yellowlined zones and those in greenlined areas with the probability of cash purchases, or the buyer is an investor (Tables C.4 and C.6, respectively, column (4)).

Table IV.2: House Flipping Probability, by HOLC Zone

	Model (1)	Model (2)	Model (3)	Model (4)
Relative to Zone A				
Zone B	.003	.003	.004	.003
	(.003)	(.002)	(.003)	(.003)
Zone C	.016	.016	.010	.010
	(.002)	(.002)	(.003)	(.003)
Zone D	.010	.010	.016	.017
	(.002)	(.002)	(.003)	(.003)
Year Fixed Effects	No	Yes	Yes	Yes
Zip Codes Fixed Effects	No	No	Yes	Yes
House Characteristics	No	No	No	Yes
# of Observations	154,235	154,235	154,232	142,262

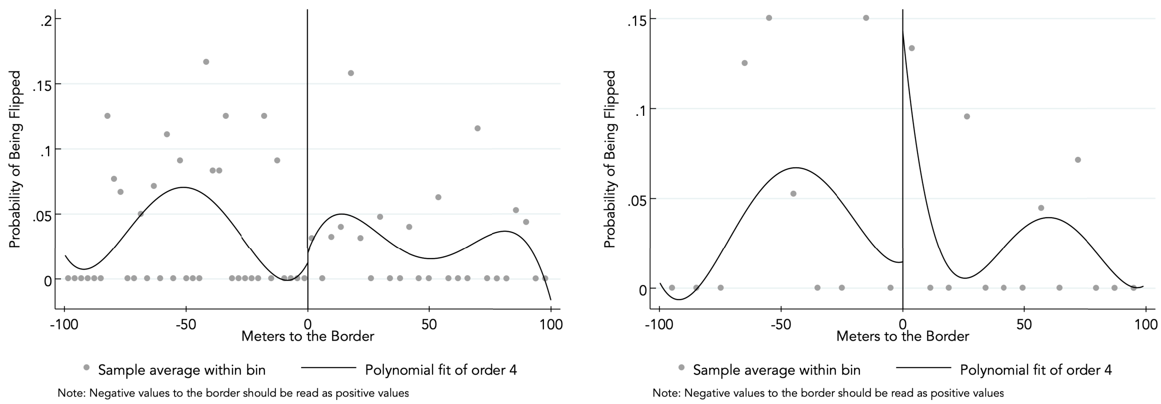
Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. House characteristics include the sale price of the transaction, age of the property, type of deed, living area of the property in square feet, size of the property in square feet, number of bedrooms in the property, number of bathrooms in the property, and number of parking spaces.

Spatial Regression Discontinuity Design with Inverse Probability Weights. As discussed in subsection III.II, estimator IV.4 has two potential limitations. Firstly, the concern is that the maps created by the HOLC may have reflected and codified pre-existing differences in neighborhoods but were not associated with future disinvestment. Secondly, estimator IV.4 may not fully control for potential confounders that affect the likelihood of an area receiving a specific HOLC grade and its contemporaneous likelihood of a property being flipped. Moreover, the estimates presented in Table IV.2 are limited to testing the effect of residing in a redlined zone and do not allow for examining the impact of living within a yellowlined area without altering the baseline HOLC zone. To address this limitation, I utilized a Spatial Regression Discontinuity Design (SRDD), which involves analyzing changes in property-level outcomes in areas that are geographically close but on opposite sides of a HOLC boundary. To increase the precision of the analysis, I also incorporated Inverse Probability Weights, which adjust the kernel function of the SRDD based on the likelihood of a particular area receiving a specific HOLC grade, as further described in subsection III.II.

Table C.7 displays various estimates of , which are obtained from estimator IV.5 for the probability of property transactions being flipped. Each row in the table corresponds to a different estimation method, while each column corresponds to a different average treatment border effect. Figures IV.13(a) and IV.13(b) display Regression Discontinuity Plots for the treatment effect at the borders "A-B" and "A-C," respectively. The vertical difference or jump at the discontinuity (i.e., 0 meters from the border) represents the estimate from estimator IV.5. Both plots indicate an increase in the probability of a property transaction being flipped at the discontinuity point where a lower-rated zone meets a higher-rated zone. The probability of a transaction being a flip showed a small increase of +2.9% between zones A (i.e., "Best") and B (i.e., "Still Desirable"), but this difference was not statistically significant. However, there was a larger statistically significant increase of +12.2% in the same probability for transactions within zone C (i.e., "Definitely Declining") relative to zone A, supporting the "yellowlining" hypothesis suggested by Aaronson, Hartley and Mazumder (2021).

Figure IV.13: Spatial Regression Discontinuity Design with IPW for "A - B" and "A - C" Borders, Probability of a Home being Flipped

(a) "A - B" Border (b) "A - C" Border

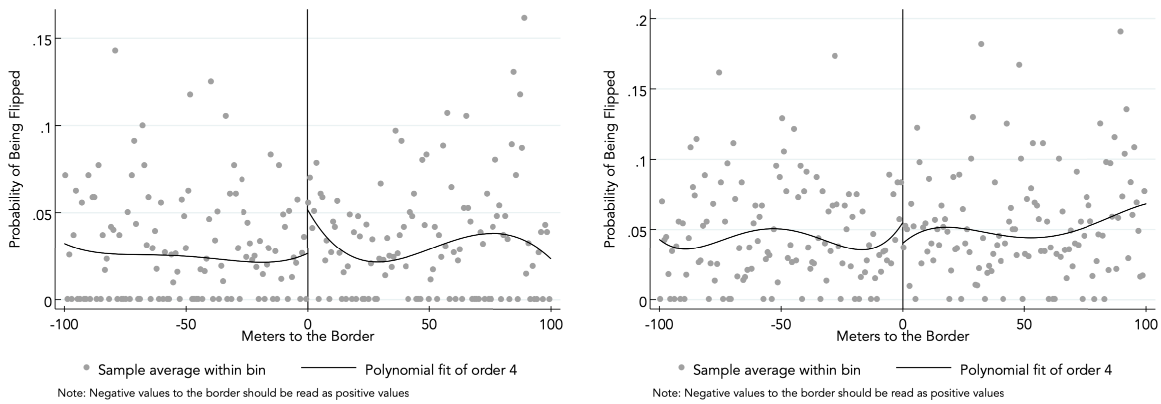


Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and to the right represents the fitted polynomial.

Figure IV.14 (a) and Figure IV.14 (b) display Regression Discontinuity Plots of the treatment effect at the borders "B-C" and "C-D," respectively. Once again, we observe a sharp increase or jump in the outcome variable at the discontinuity point where a lower-rated zone meets a higher-rated zone. The probability of a transaction being a flip is +3.5% higher in zones C than in zones B, and this difference is statistically significant at the 5% level, providing further support for the "yellowlining" hypothesis. However, we do not observe a statistically significant effect in the probability of a transaction being a flip in zone D (i.e., "Hazardous") compared to zone C.

Figure IV.14: Spatial Regression Discontinuity Design with IPW, for "B - C" and "C - D" Borders, Probability of a Home being Flipped

(a) "B - C" Border (b) "C - D" Border



Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and the right represents the fitted polynomial.

In addition to the association between historical disinvestment and house flipping, I found associations between the sale price of a transaction, the probability of a transaction being a foreclosure, and a neighborhood being graded by HOLC as D or C. On the one hand, there is a positive difference in the prices of properties close to Zone C borders compared to those close to Zone A borders. Specifically, I only found that properties close to a Zone C border were sold at a price that was 193% higher than those close to a Zone A border. This result was statistically significant at the .01% level, as evidenced by Figures C.3 and C.4, and summarized in Table C.8. On the other hand, I only found a significant difference in the likelihood of a property being sold as a foreclosure within a close distance of Zone B compared to a close distance of a Zone A. Specifically, properties located within a close distance of Zone B border had a 7.6% higher likelihood of being sold as a foreclosure than those located within a close distance of a Zone A border. This difference was statistically significant at the 1% level and is supported by Figures C.7 and C.8, as well as by the data in Table C.10.

Finally, neither the probability of paying for a transaction with cash (as shown in

Figures C.5 and C.6, and summarized in Table C.9), nor the probability of the buyer in a property transaction being an investor (as shown in Figures C.9 and C.10, and summarized in Table C.11), exhibit statistically significant effects.

IV.III Discussing the Social Importance of House Flipping

The Collaborative Data Analysis (CDA) allowed us to jointly interpret and discuss statistical model results with academic and resident researchers who live and interact with HNS neighborhoods. In total, 13 resident researchers, two public agency researchers, and two academic researchers attended the house flipping CDA sessions.

Discussions that took place during CDA helped put into context the social importance of house flipping. For example, our cross-sector, cross-site research team discussion, thanks to resident researcher contributions, focused on possible implications of house flipping for the displacement of tenants, toxic particles and asbestos rising from demolitions, no renovation and poor quality housing, absentee landlords, long-term residents being kicked out, rent increases, gentrification, homelessness and opioid addiction, increase in evictions and displacement, and a decrease in safety. We also discussed, thanks to the lived experience of resident researcher members, the fact that house flipping was occurring near potentially positive new transportation-related, housing, and commercial developments.

How does house flipping work? Our subsequent group discussion focused on house flipping as a mode of gentrification operating in a "systematic and non-random way," and it was argued that this system relies on those with wealth and access to financial resources to make a profit. Resident researcher insights helped raise questions about how investors target areas with low-income people, vacant lots, and affordable housing. Public investments, such as transit, were also identified as potential triggers of house flipping cycles that deserve study.

A neutralized framing, such as portraying flipping as a business or side job, may contribute to flipping's role in widening inequality, according to our discussion. Cultural representation through media, such as HGTV glorifying flipping or countless books, also contributes to the house flipping system. In this sense, people and investors prioritize "winning the flipping game" over community relationships or other considerations.

However, even though house flippers profit from the lack of housing supply, individuals face few options to build generational wealth for their families. House flipping provides that opportunity, which does not make it inherently right or wrong. However, the system's victims remain the same, with this system keeping certain groups stagnant and removing choice from housing for those not benefiting from flipping.

House Flipping Impacts. The group discussions and the process of collaboratively analyzing the results led to new research directions. For example, we raised questions about health implications while discussing the impact of house flipping on affordability and the environmental impacts of renovations. We discussed how rising housing and rent prices often force families to live in overcrowded conditions, leading to increased homelessness and family separation. These, in turn, can have intergenerational physical and mental health impacts beyond the immediate financial strains. For example, we talked about the feeling of congestion and churn due to a reduction of housing leads to a feeling of "always something happening, where there is no place just to take a breath."

The impact of house flipping on an individual's sense of place, community well-being, and access to green spaces were other significant outcomes that surfaced thanks to resident researchers' contributions to discussions. Priorities for future study include how the proliferation of house flipping reduces open green areas, as property owners fenced off previously accessible public spaces and limited mobility within the neighborhood. Other priority questions include how house flipping disrupts economic vitality by displacing locally owned businesses and replacing them with high-end supermarkets that often increased the prices of food products for the whole community and harmed the livelihoods of those who depend on such businesses.

Policy Solutions and Interventions. Finally, our discussions —thanks to resident researcher contributions —helped identify potential policy interventions and actions related to house flipping that could benefit local communities. Potentially helpful actions include policies targeted to increase access to information, support home repairs, revise zoning laws, creation of cooperative and land trusts with limited-equity covenants, provide financial assistance to first-time homebuyers, and augment the capital tax levied on house flipping was discussed. To combat real estate speculators and maintain affordable housing, community land ownership through cooperative and land trusts with limited-equity covenants may be promising. These policies aim to enable community members to outmaneuver the real estate flippers and secure accessible housing. Additionally, as community members become more aware of the predatory practices employed by flippers in their neighborhoods, potential policies may be aimed at improving access to information. By providing communities with knowledge about the potential profit of flipping, homeowners could avoid selling their properties or benefit from increased sale prices.

Another potential policy designed to compete against all-cash flip buyers could involve providing financial assistance for first-time homebuyers. However, doubts remained about the feasibility of this approach since it would require substantial financial support to level the playing field. The policy proposal focused on supporting home repairs was targeted to prevent homeowners from being compelled to sell their homes in the face of unexpected events and to reduce the potential for flippers to profit from such situations. In our conversations, we also discussed supply-side policies. Specifically, we explored the effects of inflexible zoning laws as a mechanism that restricts housing supply and reduces opportunities for homeowners to earn additional income (e.g., by renting out a potential accessory dwelling unit), thereby decreasing the likelihood of resorting to selling their property for emergency funds. Finally, the idea of an increase in the number of community-based realtors to help ensure that property transactions are carried out in ways that are beneficial to both the financial and community aspects of the transaction was explored.

V Conclusion

The rise of house flipping as a popular investment strategy has been accompanied by debates on its impact on the housing market, particularly on housing supply and affordability. This chapter has focused on investigating the effects of house flipping on rapidly changing neighborhoods in the Greater Boston area, with a particular emphasis on understanding the mechanisms by which flippers acquire properties and the types of properties they target. Through a mixed-methods approach and the collaboration with resident researchers of the HNS consortium, this study has highlighted the complex effects of house flipping on neighborhoods, including the potential displacement of long-term residents and the disruption of community stability. The findings show that flipping disproportionately affects affordable housing sales and homes within historically disinvested areas.

This study contributes to the literature on the impact of investors in the housing market by shedding light on the extent of house flipping and its complex and nonlinear impacts. While previous studies have focused on the average or median prices of homes, this chapter provides a more nuanced analysis of the mechanisms and types of properties targeted by flippers. The insights gained from this study could inform policy decisions to mitigate the negative impacts of house flipping on neighborhoods. Some suggestions include increasing access to information, supporting home repairs, revising zoning laws, creating cooperative and land trusts with limited-equity covenants, and providing financial assistance to first-time homebuyers.

Bibliography

Aaronson, Daniel, Bhashkar Mazumder, Daniel A Hartley, and Martha Harrison Stinson. 2022. "The long-run effects of the 1930s redlining maps on children."

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder. 2021. "The effects of the 1930s HOLC" redlining" maps." *American Economic Journal: Economic Policy*, 13(4): 355–92.

Akao, Ken-Ichi, and Hiroaki Sakamoto. 2018. "A theory of disasters and long-run growth." *Journal of Economic Dynamics and Control*, 95: 89–109.

Allen, Heidi L., Erica Eliason, Naomi Zewde, and Tal Gross. 2019. "Can Medicaid Expansion Prevent Housing Evictions?" *Health Affairs*, 38(9): 1451–1457. Publisher: Health Affairs.

Angrist, Joshua, and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* . 1 ed., Princeton University Press.

Arcaya, Mariana C, Alina Schnake-Mahl, Andrew Binet, Shannon Simpson, Maggie Super Church, Vedette Gavin, Bill Coleman, Shoshanna Levine, Annika Nielsen, Leigh Carroll, et al. 2018. "Community change and resident needs: Designing a Participatory Action Research study in Metropolitan Boston." *Health & place*, 52: 221–230.

Arcaya, Mariana, Ethan J. Raker, and Mary C. Waters. 2020. "The Social Consequences of Disasters: Individual and Community Change." *Annual Review of Sociology*, 46(1): null. _eprint: https://doi.org/10.1146/annurev-soc-121919-054827.

ATTOM. 2023. "Home Flipping Remains Up In 2022 Across U.S. But Gross Profits Fall To Another Low."

Bailey, Zinzi D, Nancy Krieger, Madina Agénor, Jasmine Graves, Natalia Linos, and Mary T Bassett. 2017. "Structural racism and health inequities in the USA: evidence and interventions." *The lancet*, 389(10077): 1453–1463.

Bassett, Mary T, Jarvis T Chen, and Nancy Krieger. 2020. "Variation in racial/ethnic disparities in COVID-19 mortality by age in the United States: A cross-sectional study." *PLoS medicine*, 17(10): e1003402.

Bayer, Patrick, Christopher Geissler, James W Roberts, et al. 2011. *Speculators and middlemen: The role of flippers in the housing market.* Vol. 16784, Citeseer.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A unified framework for measuring preferences for schools and neighborhoods." *Journal of political economy*, 115(4): 588–638.

Benfer, Emily A., David Vlahov, Marissa Y. Long, Evan Walker-Wells, J. L. Pottenger, Gregg Gonsalves, and Danya E. Keene. 2021. "Eviction, Health Inequity, and the Spread of COVID-19: Housing Policy as a Primary Pandemic Mitigation Strategy." *Journal of Urban Health*, 98(1): 1–12.

Betancourt, Kim, Stephen Gardner, and Mark Palim. 2022. "The U.S. Housing Shortage from a Local Perspective."

Binet, Andrew, Vedette Gavin, Leigh Carroll, and Mariana Arcaya. 2019. "Designing and facilitating collaborative research design and data analysis workshops: lessons learned in the healthy neighborhoods study." *International journal of environmental research and public health*, 16(3): 324.

Boddupalli, Aravind, and Kim Rueben. 2021. "State and Local Government Revenues and Racial Disparities." *Washington, DC: Urban-Brookings Tax Policy Center*.

Botzen, W. J. Wouter, Olivier Deschenes, and Mark Sanders. 2019. "The Economic Impacts of Natural Disasters: A Review of Models and Empirical Studies." *Review of Environmental Economics and Policy*, 13(2): 167–188. Publisher: The University of Chicago Press.

Boustan, Leah, Matthew Kahn, Paul Rhode, and Maria Lucia Yanguas. 2020. "The effect of natural disasters on economic activity in US counties: A century of data." *Journal of Urban Economics*, 118(C). Publisher: Elsevier.

Brooks, Harold E, Charles A Doswell III, and Michael P Kay. 2003. "Climatological estimates of local daily tornado probability for the United States." *Weather and forecasting*, 18(4): 626–640.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2014. "Robust data-driven inference in the regression-discontinuity design." *The Stata Journal*, 14(4): 909–946.

Centers for Disease Control and Prevention. 2020. "Temporary Halt in Residential Evictions to Prevent the Further Spread of COVID-19."

Chetty, Raj. 2021. "Affordable Housing as a Pathway to Economic Opportunity."

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2020. "Race and economic opportunity in the United States: An intergenerational perspective." *The Quarterly Journal of Economics*, 135(2): 711–783.

Clotfelter, Charles, Elizabeth Glennie, Helen Ladd, and Jacob Vigdor. 2008. "Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina." *Journal of Public Economics*, 92(5): 1352–1370.

Costanza-Chock, Sasha. 2020. *Design justice: Community-led practices to build the worlds we need.* The MIT Press.

Crespo Cuaresma, Jesús, Jaroslava Hlouskova, and Michael Obersteiner. 2008. "Natural Disasters as Creative Destruction? Evidence from Developing Countries." *Economic Inquiry*, 46(2): 214–226. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1465-7295.2007.00063.x.

Daepp, Madeleine IG, Andrew Binet, Vedette Gavin, Mariana C Arcaya, and Healthy Neighborhoods Research Consortium. 2022. "The moving mapper: Participatory action research with big data." *Journal of the American Planning Association*, 88(2): 179–191.

Demsas, Jerusalem. 2023. "Meet the Latest Housing-Crisis Scapegoat." Section: Ideas.

Depken, Craig A, Harris Hollans, and Steve Swidler. 2009. "An empirical analysis of residential property flipping." *The Journal of Real Estate Finance and Economics*, 39: 248–263.

Dhar, Paramita, and Stephen L Ross. 2012. "School district quality and property values: Examining differences along school district boundaries." *Journal of Urban Economics*, 71(1): 18–25.

Dong, Ensheng, Hongru Du, and Lauren Gardner. 2020. "An interactive web-based dashboard to track COVID-19 in real time." *The Lancet Infectious Diseases*, 20(5): 533–534. Publisher: Elsevier.

Field, Christopher B., Vicente Barros, Thomas F. Stocker, and Qin Dahe, ed. 2012. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change.* Cambridge:Cambridge University Press.

Fishback, Price V, Jessica LaVoice, Allison Shertzer, and Randall Walsh. 2020. "The HOLC maps: How race And poverty influenced real estate professionals' evaluation of lending risk in the 1930s." National Bureau of Economic Research.

Fishback, PV, J Rose, KA Snowden, and T Storrs. 2021. "New evidence on redlining by federal housing programs in the 1930s [Working Paper 29244]."

Fussell, Elizabeth, Narayan Sastry, and Mark VanLandingham. 2010. "Race, socioeconomic status, and return migration to New Orleans after Hurricane Katrina." *Population and Environment*, 31(1/3): 20–42. Publisher: Springer.

Glaeser, Edward L, and Joseph Gyourko. 2005. "Urban decline and durable housing." *Journal of political economy*, 113(2): 345–375.

Guo, Yi, Zhaoyi Chen, Ke Xu, Thomas J. George, Yonghui Wu, William Hogan, Elizabeth A. Shenkman, and Jiang Bian. 2020. "International Classification of Diseases, Tenth Revision, Clinical Modification social determinants of health codes are poorly used in electronic health records." *Medicine*, 99(52): e23818.

Gyourko, Joseph, and Albert Saiz. 2004. "Reinvestment in the housing stock: the role of construction costs and the supply side." *Journal of Urban Economics*, 55(2): 238–256.

Hallegatte, Stéphane, and Patrice Dumas. 2009. "Can natural disasters have positive consequences? Investigating the role of embodied technical change." *Ecological Economics*, 68(3): 777–786.

Haughwout, Andrew, Sarah Sutherland, and Joseph S Tracy. 2013. "Negative equity and housing investment." *FRB of New York Staff Report*, , (636).

Heckman, James J., and Jeffrey A. Smith. 1999. "The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies." *The Economic Journal*, 109(457): 313–348. Publisher: [Royal Economic Society, Wiley].

Hernan, MA, and JM Robins. 2021. *Causal Inference: What If.*

Hillier, Amy E. 2003. "Redlining and the home owners' loan corporation." *Journal of Urban History*, 29(4): 394–420.

Howell, Junia, and James R Elliott. 2019. "Damages Done: The Longitudinal Impacts of Natural Hazards on Wealth Inequality in the United States." *Social Problems*, 66(3): 448–467.

Hsiang, Solomon M., and Amir S. Jina. 2014. "The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth: Evidence From 6,700 Cyclones." National Bureau of Economic Research Working Paper 20352. Series: Working Paper Series.

Joint Center for Housing Studies. 2022. "The State of the Nation's Housing 2022." Harvard University, Cambridge, MA.

Kawachi, Ichiro, and S. V. Subramanian. 2018. "Social epidemiology for the 21st century." *Social Science & Medicine (1982)*, 196: 240–245.

Krieger, James, and Donna L Higgins. 2002. "Housing and health: time again for public health action." *American journal of public health*, 92(5): 758–768.

Lechner, Michael. 2011. "The Estimation of Causal Effects by Difference-in-Difference Methods." *Foundations and Trends(R) in Econometrics*, 4(3): 165–224.

Lee, Jin Man, and Jin Wook Choi. 2011. "The role of house flippers in a boom and bust real estate market." *The Journal of Economic Asymmetries*, 8(2): 91–109.

Leifheit, Kathryn M., Sabriya L. Linton, Julia Raifman, Gabriel Schwartz, Emily A. Benfer, Frederick J. Zimmerman, and Craig Pollack. 2020. "Expiring Eviction Moratoriums and COVID-19 Incidence and Mortality." Social Science Research Network SSRN Scholarly Paper ID 3739576, Rochester, NY.

Leung, Charles Ka Yui, and Chung-Yi Tse. 2017. "Flipping in the housing market." *Journal of Economic Dynamics and Control*, 76: 232–263.

Lian, Min, Mario Schootman, and Shumei Yun. 2008. "Geographic variation and effect of area-level poverty rate on colorectal cancer screening." *BMC Public Health*, 8(1): 358.

Li, Lingxiao, Abdullah Yavas, and Bing Zhu. 2023. "Externalities of residential property flipping." *Real Estate Economics*, 51(1): 233–271.

Logan, John R., Sukriti Issar, and Zengwang Xu. 2016. "Trapped in Place? Segmented Resilience to Hurricanes in the Gulf Coast, 1970–2005." *Demography*, 53(5): 1511–1534.

Marani, Marco, Gabriel G. Katul, William K. Pan, and Anthony J. Parolari. 2021. "Intensity and frequency of extreme novel epidemics." *Proceedings of the National Academy of Sciences*, 118(35). Publisher: National Academy of Sciences Section: Biological Sciences.

Markley, Scott. 2023. "Tabulating Home Owners' Loan Corporation area description sheet data." *Environment and Planning B: Urban Analytics and City Science*, 50(1): 268–280.

McCaffrey, Daniel F, Beth Ann Griffin, Daniel Almirall, Mary Ellen Slaughter, Rajeev Ramchand, and Lane F Burgette. 2013. "A tutorial on propensity score estimation for multiple treatments using generalized boosted models." *Statistics in medicine*, 32(19): 3388–3414.

Melzer, Brian T. 2017. "Mortgage debt overhang: Reduced investment by homeowners at risk of default." *The Journal of Finance*, 72(2): 575–612.

Michney, Todd M, and LaDale Winling. 2020. "New perspectives on new deal housing policy: explicating and mapping HOLC loans to African Americans." *Journal of Urban History*, 46(1): 150–180.

Nande, Anjalika, Justin Sheen, Emma L. Walters, Brennan Klein, Matteo Chinazzi, Andrei H. Gheorghe, Ben Adlam, Julianna Shinnick, Maria Florencia Tejeda, Samuel V. Scarpino, Alessandro Vespignani, Andrew J. Greenlee, Daniel Schneider, Michael Z. Levy, and Alison L. Hill. 2021. "The effect of eviction moratoria on the transmission of SARS-CoV-2." *Nature Communications*, 12(1): 2274. Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Applied mathematics;Policy and public health in microbiology;Viral infection Subject_term_id: applied-mathematics;policy-and-public-health-in-microbiology;viral-infection.

National Association of Realtors. 2022. "Impact of Institutional Investor Buyers on Home Sales and Single-Family Rentals."

Nelson, Robert, LaDale Winling, Richard Marciano, and Nathan Connoly. 2023. "Mapping Inequality."

OptumLabs. 2020. "OptumLabs and OptumLabs Data Warehouse (OLDW) Descriptions and Citation." Eden Prairie, MN.

Powell, Walter W, and Jeannette A Colyvas. 2008. "Microfoundations of institutional theory." *The Sage handbook of organizational institutionalism*, 276: 298.

Quan, Hude, Bing Li, Chantal M. Couris, Kiyohide Fushimi, Patrick Graham, Phil Hider, Jean-Marie Januel, and Vijaya Sundararajan. 2011. "Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries." *American Journal of Epidemiology*, 173(6): 676–682.

Raifman, J, K Nocka, D Jones, J Bor, S Lipson, J Jay, and P Chan. 2020. "COVID-19 US state policy database."

Raker, Ethan J. 2020. "Natural Hazards, Disasters, and Demographic Change: The Case of Severe Tornadoes in the United States, 1980–2010." *Demography*, 57(2): 653–674.

Ravallion, Martin. 2002. "On the urbanization of poverty." *Journal of Development Economics*, 68(2): 435–442.

Rothstein, Richard. 2017. *The color of law: A forgotten history of how our government segregated America.* Liveright Publishing.

Sagalyn, Lynne Beyer. 1980. "Housing on the installment plan: an economic and institutional analysis of contract buying in Chicago." PhD diss. Massachusetts Institute of Technology.

Skidmore, Mark, and Hideki Toya. 2002. "Do Natural Disasters Promote Long-Run Growth?" *Economic Inquiry*, 40(4): 664–687. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1093/ei/40.4.664.

Stott, Peter. 2016. "How climate change affects extreme weather events." *Science*, 352(6293): 1517–1518.

Strobl, Eric. 2011. "The Economic Growth Impact of Hurricanes: Evidence from U.S. Coastal Counties." *The Review of Economics and Statistics*, 93(2): 575–589.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

US Census Bureau. 2020*a*. "2015-2019 American Community Survey 5-Year Data."

US Census Bureau. 2020*b*. "Week 13 Household Pulse Survey: August 19 – August 31 [Housing Table 1b]."

USGCRP. 2018. "Fourth National Climate Assessment." U.S. Global Change Research Program, Washington, DC.

Venkataramani, Atheendar S., Elizabeth F. Bair, Rourke L. O'Brien, and Alexander C. Tsai. 2020. "Association Between Automotive Assembly Plant Closures and Opioid Overdose Mortality in the United States: A Difference-in-Differences Analysis." *JAMA Internal Medicine*, 180(2): 254.

Wassmer, Robert W. 2023. "Past structural racism and present home prices." *Journal of Urban Affairs*, 1–24.

Xu, Wenfei. 2022. "Legacies of institutionalized redlining: a comparison between speculative and implemented mortgage risk maps in Chicago, Illinois." *Housing Policy Debate*, 32(2): 249–274.

Zhang, Sharon. 2022. "Ocasio-Cortez Calls Out Real Estate Firms for "Gobbling Up" Homes and Constricting Housing Market."

Appendix A

Appendix for The Effects of Localized Climate Shocks on Places and People

Table A.1: Balance Tests (1). Pre-Tornado Difference-In-Differences

	(1) Sale Price	(2) Institution	(3) Sq. Ft.	(4) Year Built	(5) Bathrooms
Treatment	331.665	0.000	6.272	-0.029	-0.000
	(1977.325)	(0.001)	(10.392)	(0.253)	(0.008)
Constant	62900.973***	0.058***	1743.463***	1966.645***	2.008***
	(10035.076)	(0.006)	(31.432)	(0.562)	(0.035)
Obs.	2,024,386	2,024,387	1,898,839	1,882,082	1,727,015

Note: *** p 0.001, ** p 0.01, * p 0.05. Tornado and year fixed-effects. Standard Errors Clustered at the Tornado Level. Tornadoes with positive property damage. For Institution, it refers to the probability an institution bought a property.

Table A.2: Balance Tests (2). Pre-Tornado Difference-In-Differences

	(1) Walk Score	(2) Transit Score	(3) Bike Score	(4) Lot Sq. Ft.
Treatment	-0.442	-0.235	-0.173	33048.772
	(0.254)	(0.204)	(0.158)	(44747.188)
Constant	31.925***	30.908***	39.557***	-612839.435
	(0.579)	(0.277)	(0.889)	(724808.240)
Obs.	2,022,617	854,581	2,001,658	1,739,898

Note: *** p 0.001, ** p 0.01, * p 0.05. Tornado and fixed-effects. Standard Errors Clustered at the Tornado Level. Tornadoes with positive property damage.

Figure A.1: Institutional Investors Across Time

Probability that a property was bought by an institutional investor

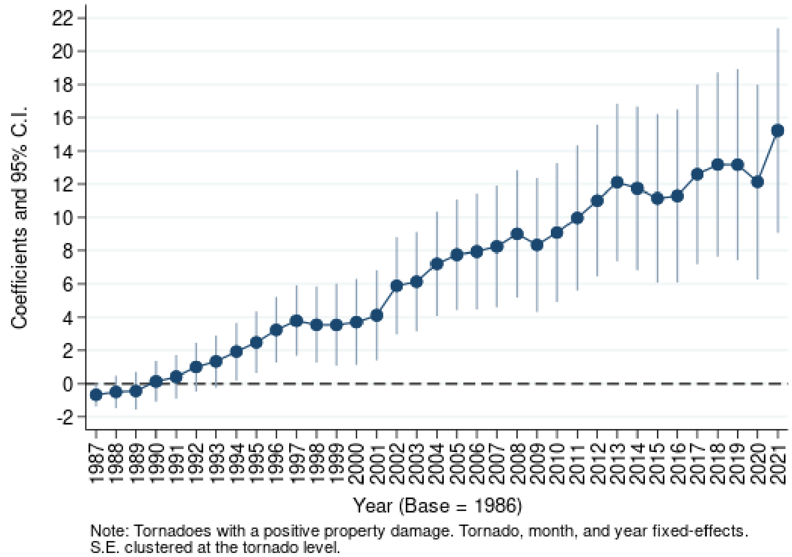


Figure A.2: Impact of a Tornado on Sale Prices for FEMA Disaster Declaration Areas, US Dollars ($1,000)

Marginal Effects relative to Areas not designated FEMA Disaster Declaration

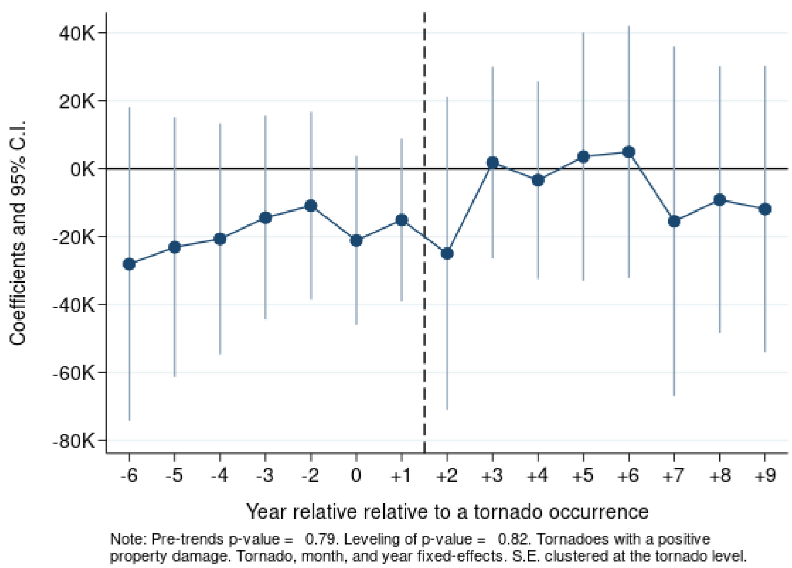
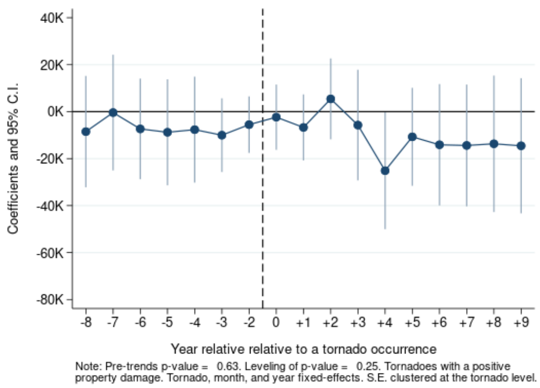


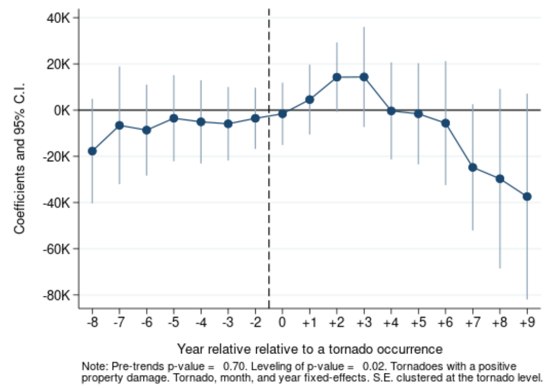
Figure A.3: Impact of a Tornado on Sale Prices by Tornado Magnitude, US Dollars ($1,000)

Marginal Effects relative to Areas Impacted by EF0 Tornadoes

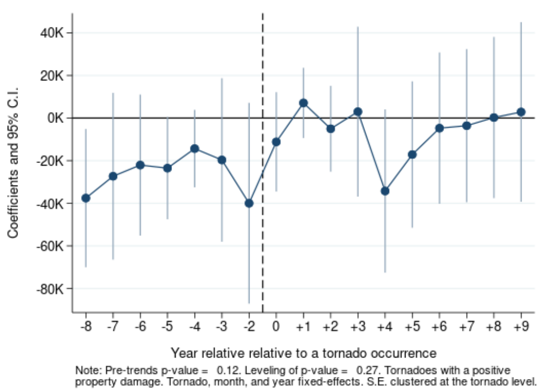
(a) Magnitude EF1



(b) Magnitude EF2



(c) Magnitude EF3



(d) Magnitude EF4

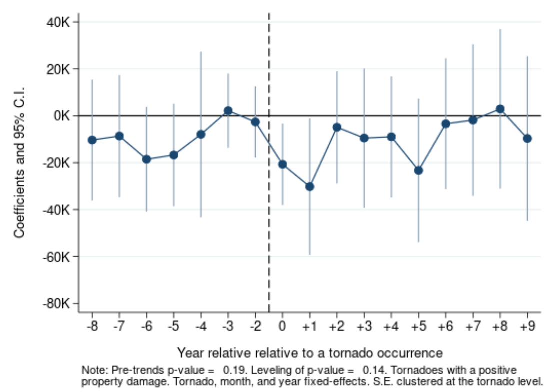


Figure A.4: Impact of a Tornado on the Probability a Property was Bought by an Institutional Investor, by Treatment Design

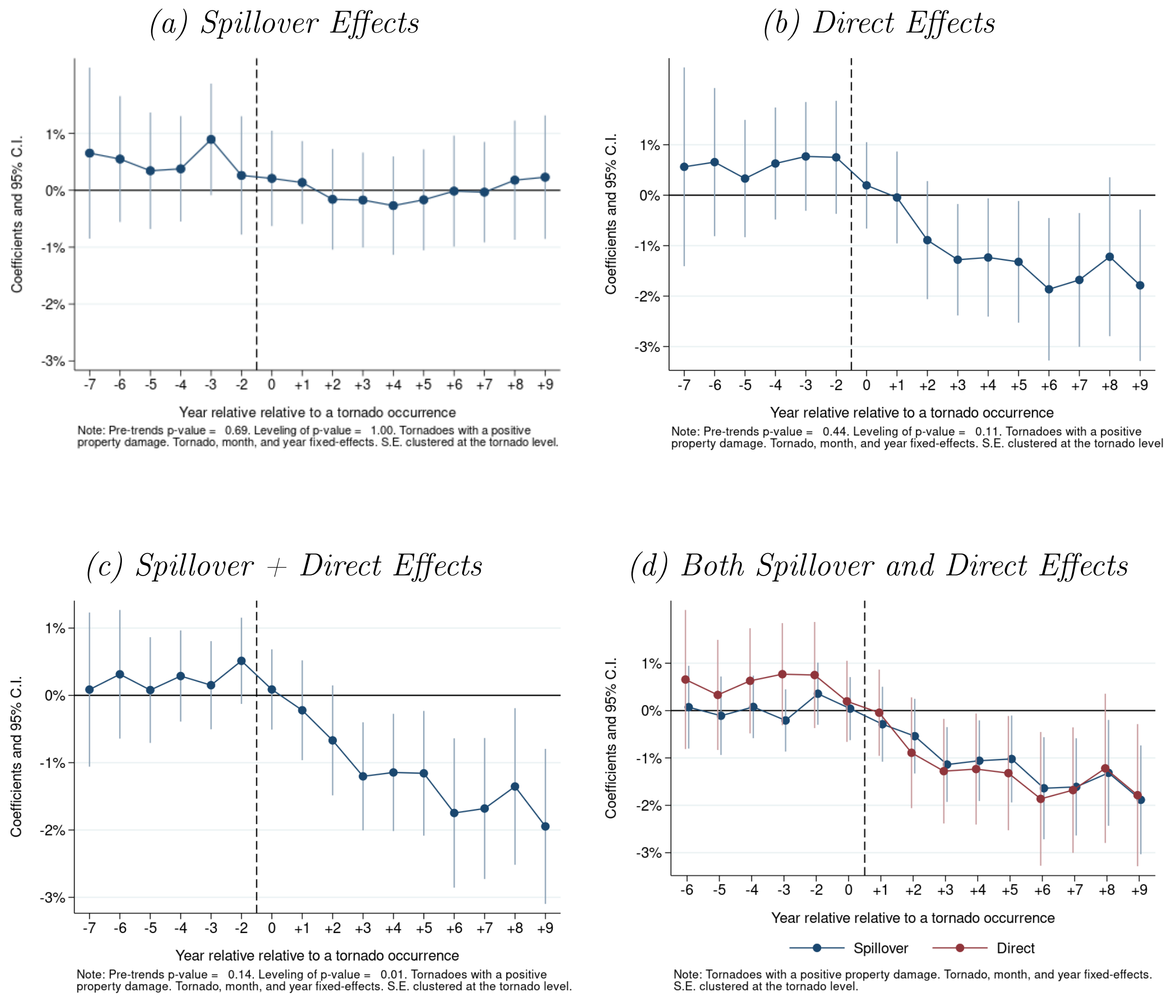
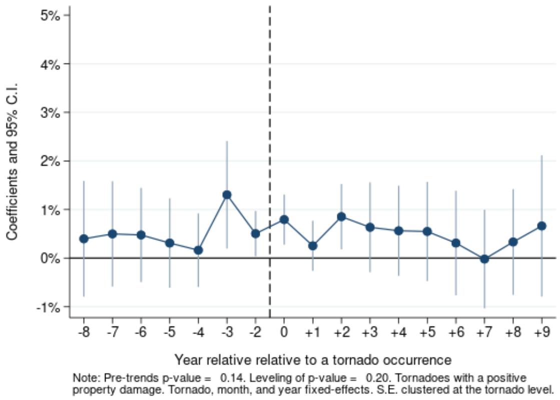
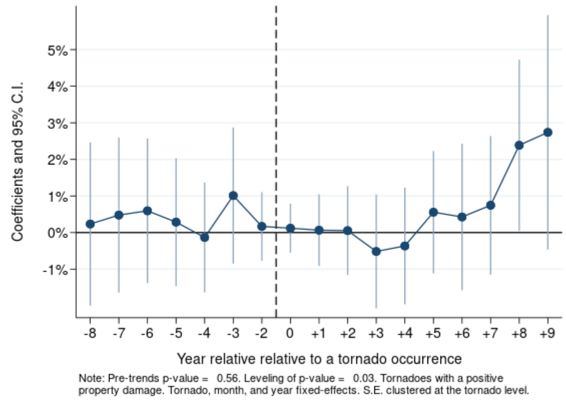


Figure A.5: Impact of a Tornado on the Probability a Property was Sold by a Bank, by Treatment Design

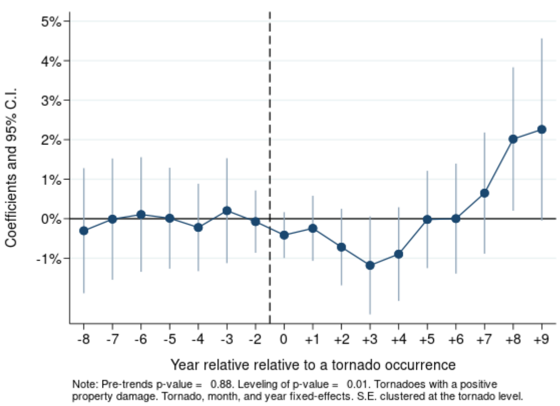
(a) Spillover Effects



(b) Direct Effects



(c) Spillover + Direct Effects



(d) Both Spillover and Direct Effects

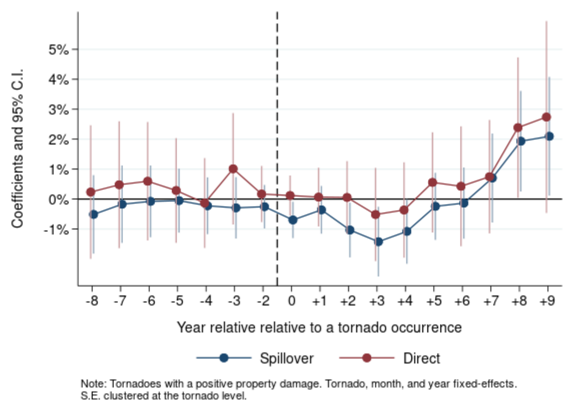
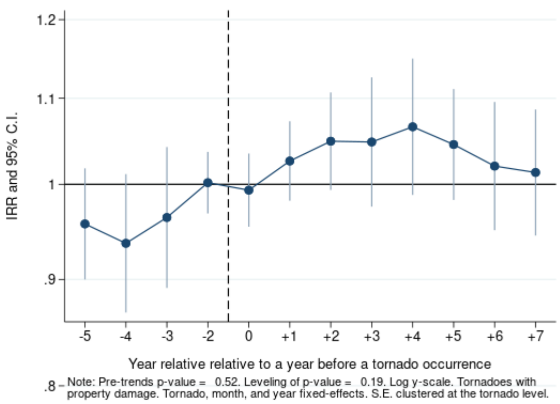
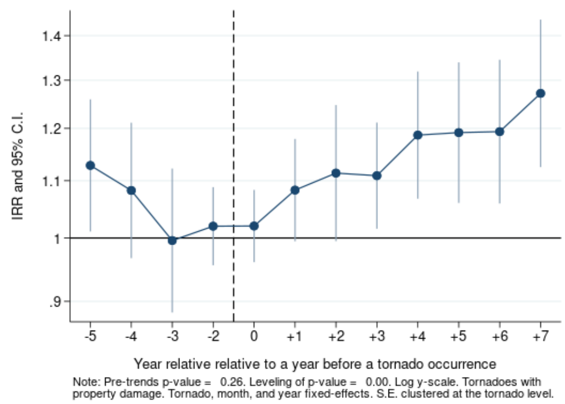


Figure A.6: Impact of a Tornado on the Number of Listings, by Treatment Design

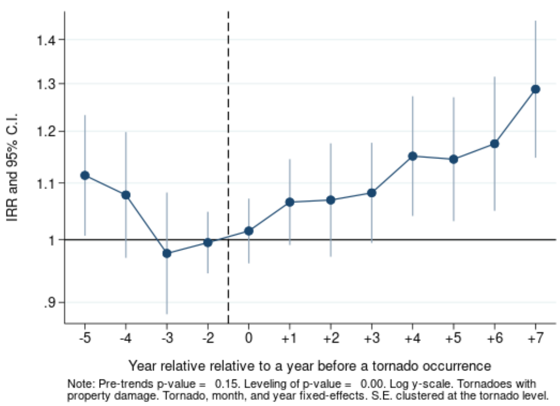
(a) Spillover Effects



(b) Direct Effects



(c) Spillover + Direct Effects



(d) Both Spillover and Direct Effects

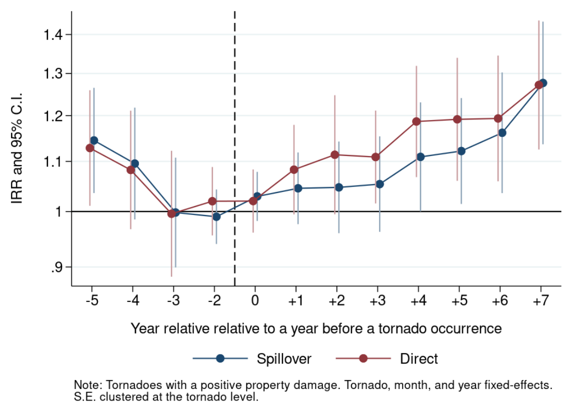
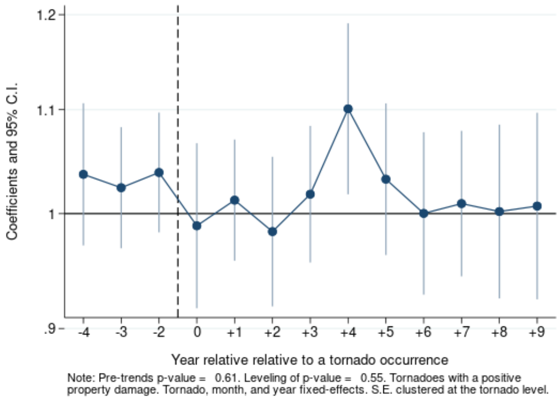
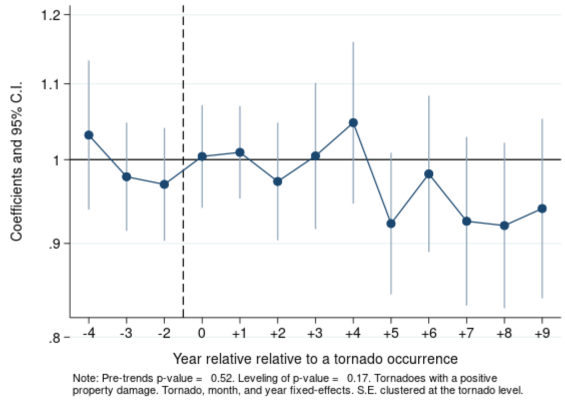


Figure A.7: Impact of a Tornado on the Days on the Market, by Treatment Design

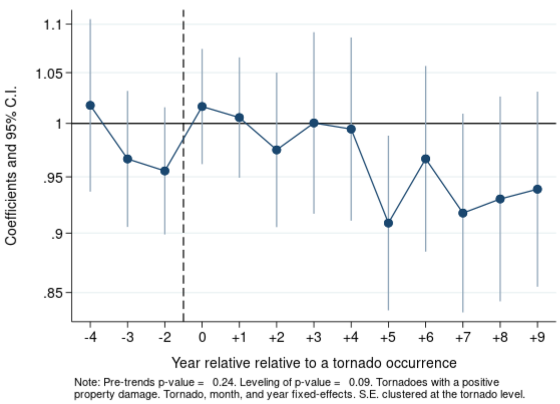
(a) Spillover Effects



(b) Direct Effects



(c) Spillover + Direct Effects



(d) Both Spillover and Direct Effects

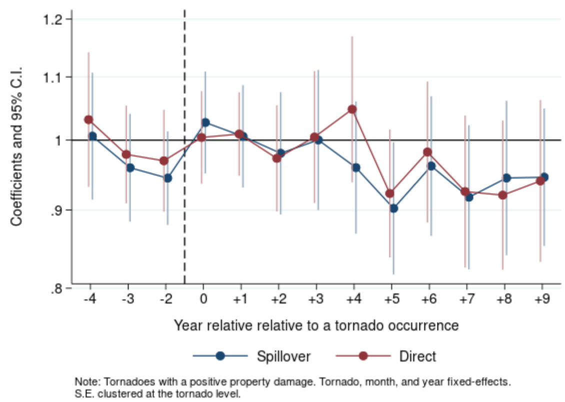
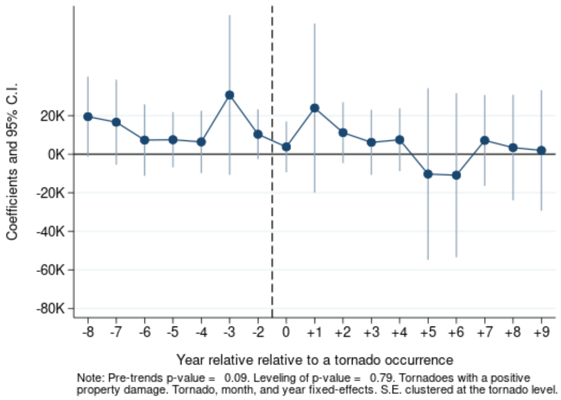
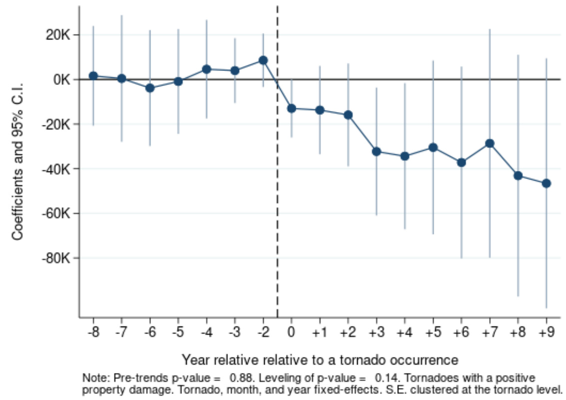


Figure A.8: Impact of a Tornado on the Sale Prices,
by the Distance to the Nearest CBD (4 km) and by Treatment Design

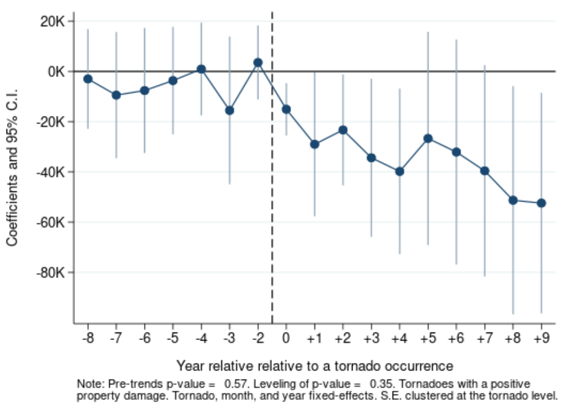
(a) Spillover Effects



(b) Direct Effects



(c) Spillover + Direct Effects



(d) Both Spillover and Direct Effects

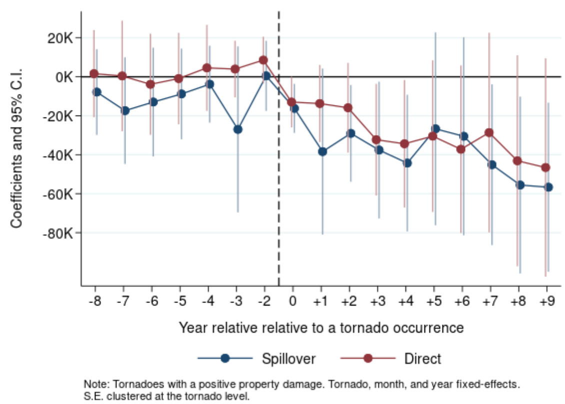


Figure A.9: Impact of a Tornado on the Sale Prices,
by the Distance to the Nearest CBD (4 km and   18 km) and by Treatment Design

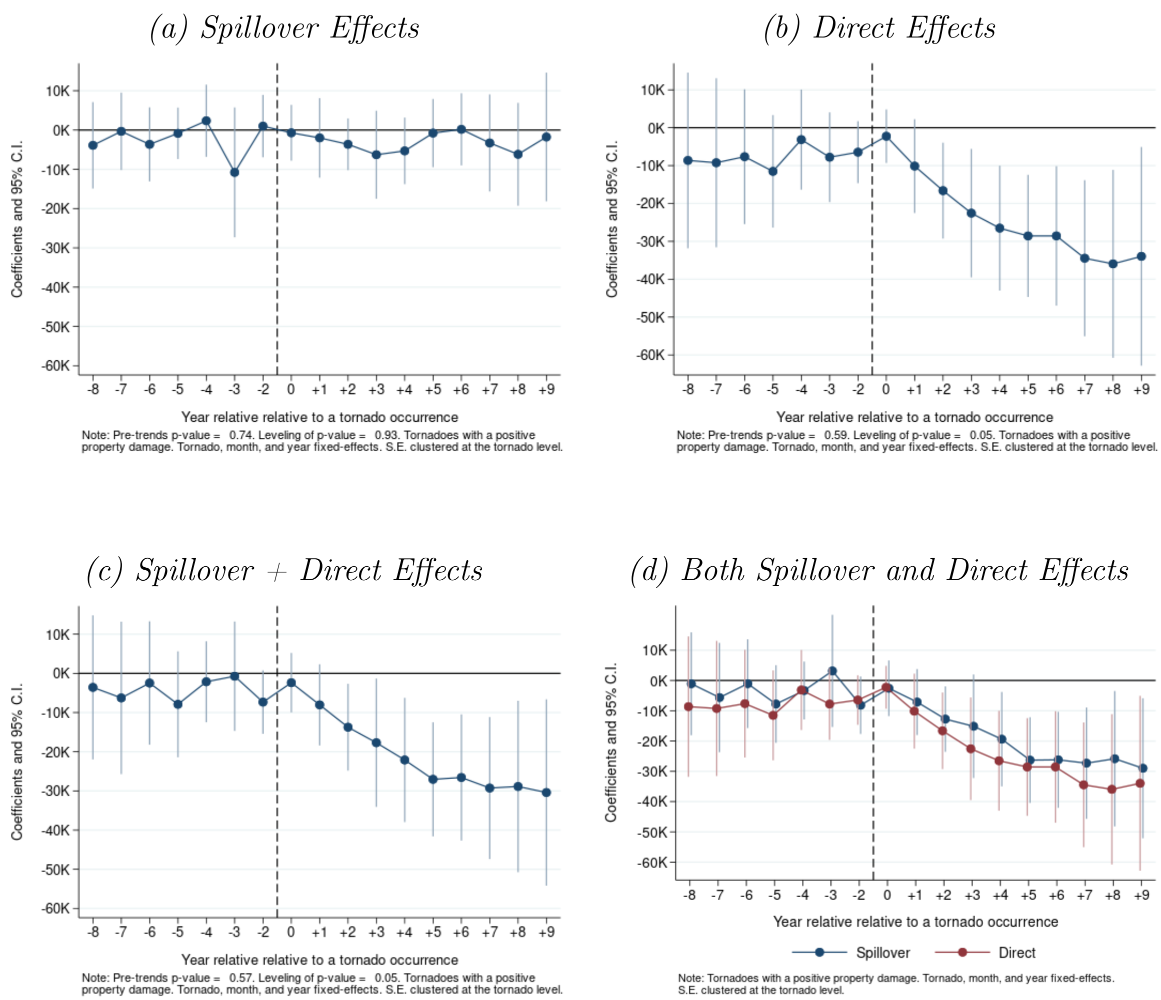
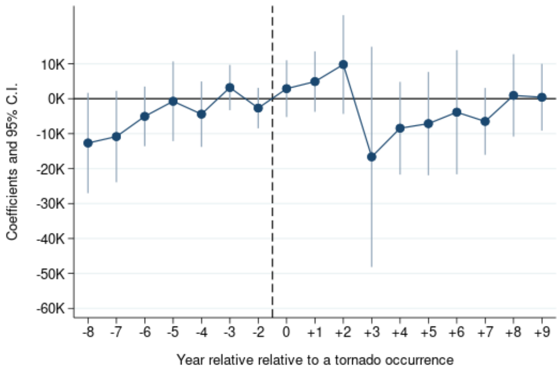


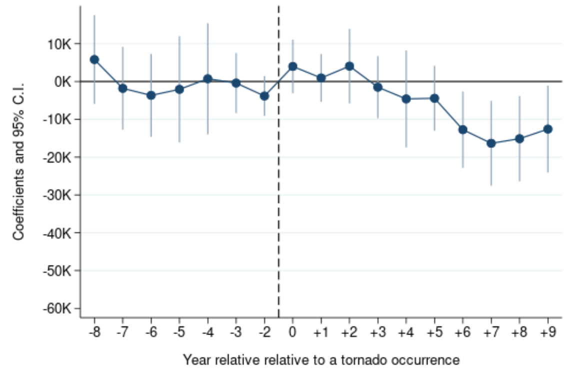
Figure A.10: Impact of a Tornado on the Sale Prices,
by the Distance to the Nearest CBD (18 km) and by Treatment Design

(a) Spillover Effects



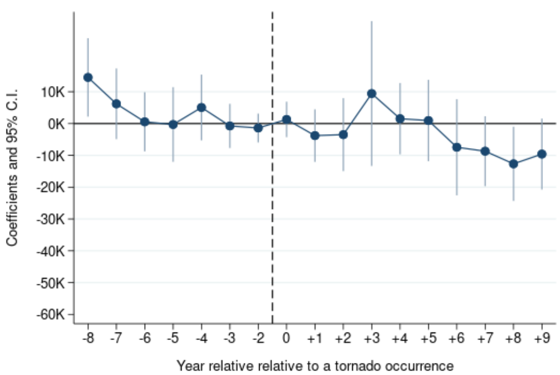
Note: Pre-trends p-value = 0.40. Leveling of p-value = 0.05. Tornadoes with a positive property damage. Tornado, month, and year fixed-effects. S.E. clustered at the tornado level.

(b) Direct Effects



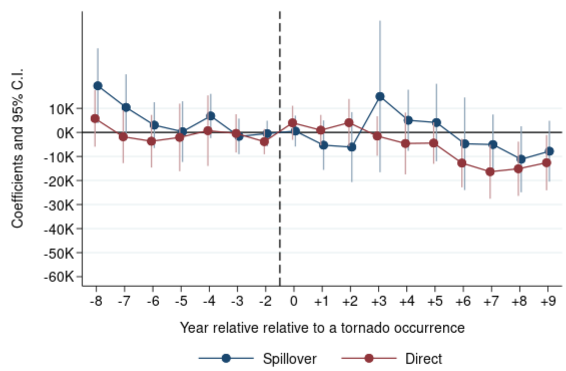
Note: Pre-trends p-value = 0.41. Leveling of p-value = 0.20. Tornadoes with a positive property damage. Tornado, month, and year fixed-effects. S.E. clustered at the tornado level.

(c) Spillover + Direct Effects



Note: Pre-trends p-value = 0.17. Leveling of p-value = 0.14. Tornadoes with a positive property damage. Tornado, month, and year fixed-effects. S.E. clustered at the tornado level.

(d) Both Spillover and Direct Effects

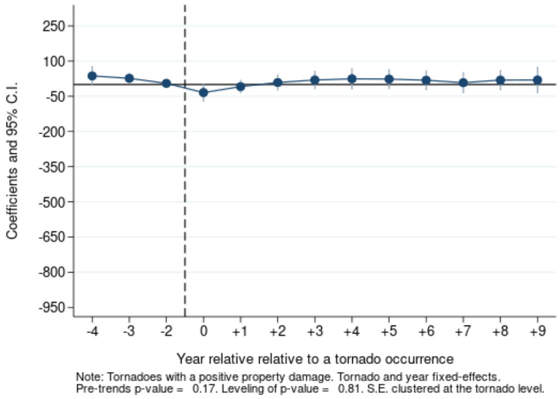


Note: Tornadoes with a positive property damage. Tornado, month, and year fixed-effects. S.E. clustered at the tornado level.

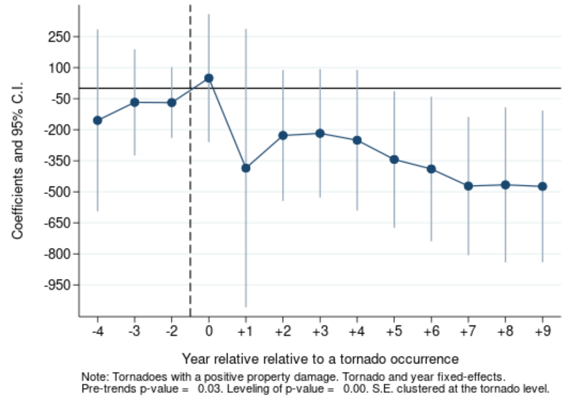
Figure A.11: Impact of a Tornado on Taxes Due per Property, by Treatment Design

US Dollars

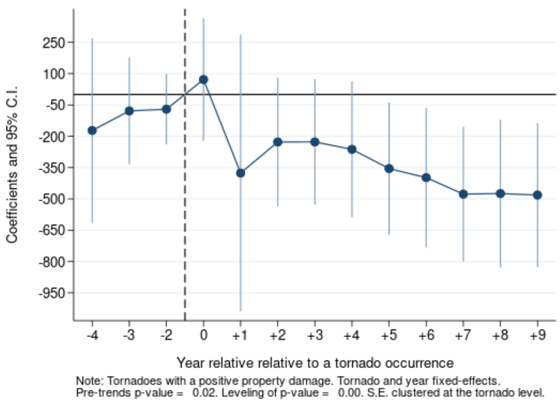
(a) Spillover Effects



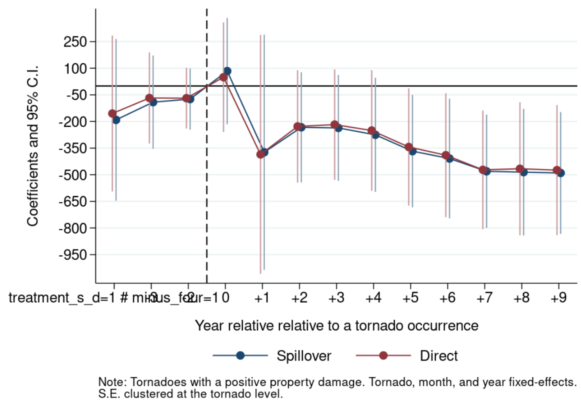
(b) Direct Effects



(c) Spillover + Direct Effects



(d) Both Spillover and Direct Effects



Appendix B

Appendix for Eviction Moratoria Expiration and COVID-19 Infection Risk

Table B.1: US States by Eviction Moratorium Implementation and Lifting Status

State	Implemented an Eviction Moratorium	Lifted their Eviction Moratorium	Week of the Year the the Moratorium was First Lifted
Alabama	Yes	Yes	23
Alaska	Yes	Yes	27
Arizona	Yes	No	NA
Arkansas	No	NA	NA
California	Yes	No	NA
Colorado	Yes	Yes	25
Connecticut	Yes	No	NA
Delaware	Yes	Yes	27
District of Columbia	Yes	No	NA
Florida	Yes	No	NA
Georgia	No	NA	NA
Hawaii	Yes	No	NA
Idaho	Yes	Yes	19
Illinois	Yes	No	NA
Indiana	Yes	Yes	34
Iowa	Yes	Yes	23
Kansas	Yes	Yes	23
Kentucky	Yes	Yes	35
Louisiana	Yes	Yes	25
Maine	Yes	Yes	32
Maryland	Yes	Yes	31
Massachusetts	Yes	No	NA
Michigan	Yes	Yes	30
Minnesota	Yes	No	NA
Mississippi	Yes	Yes	23
Missouri	No	NA	NA
Montana	Yes	No	NA
Nebraska	Yes	Yes	23
Nevada	Yes	No	NA

State	Implemented an Eviction Moratorium	Lifted their Eviction Moratorium	Week of the Year the the Moratorium was First Lifted
New Hampshire	Yes	Yes	27
New Jersey	Yes	No	NA
New Mexico	Yes	No	NA
New York	Yes	No	NA
North Carolina	Yes	Yes	26
North Dakota	Yes	Yes	17
Ohio	No	NA	NA
Oklahoma	No	NA	NA
Oregon	Yes	No	NA
Pennsylvania	Yes	No	NA
Rhode Island	Yes	Yes	27
South Carolina	Yes	Yes	21
South Dakota	No	NA	NA
Tennessee	Yes	Yes	23
Texas	Yes	Yes	21
Utah	Yes	Yes	21
Vermont	Yes	No	NA
Virginia	Yes	Yes	21
Washington	Yes	No	NA
West Virginia	Yes	Yes	21
Wisconsin	Yes	Yes	22
Wyoming	No	NA	NA

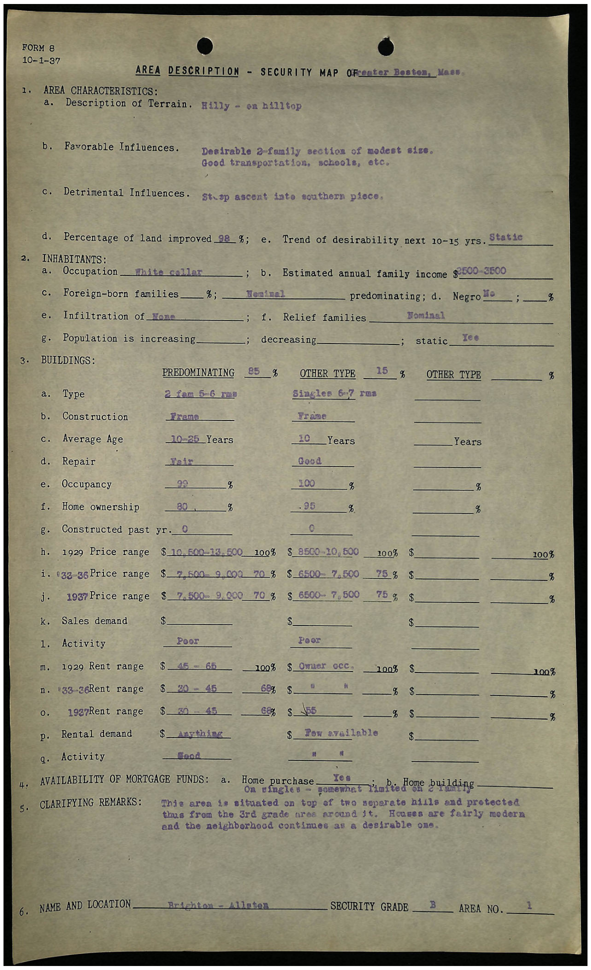
Note: NA = Not Applicable

Appendix C

Appendix for House Flipping Within Rapidly Changing Neighborhoods

Figure C.1: Examples of Area Description Files in Boston, MA

(a) HOLC Zone "B"



FORM 8
10-1-37

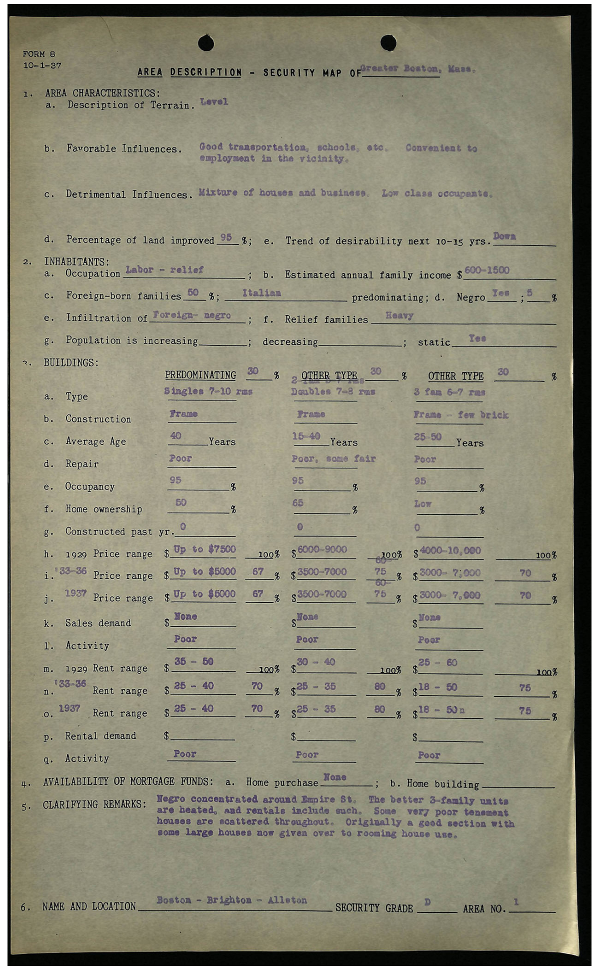
AREA DESCRIPTION - SECURITY MAP OF Greater Boston, Mass.

1. AREA CHARACTERISTICS:
   a. Description of Terrain. Hilly - on hilltop
   b. Favorable Influences. Desirable 2-family section of modest size. Good transportation, schools, etc.
   c. Detrimental Influences. Steep ascent into southern piece.
   d. Percentage of land improved 98 %; e. Trend of desirability next 10-15 yrs. Static
2. INHABITANTS:
   a. Occupation White collar ; b. Estimated annual family income $2500-3500
   c. Foreign-born families ___ %; Nominal predominating; d. Negro No ; ___ %
   e. Infiltration of None ; f. Relief families Nominal
   g. Population is increasing ___ ; decreasing ___ ; static Yes
3. BUILDINGS:

	PREDOMINATING 85 %		OTHER TYPE 15 %		OTHER TYPE ___ %	
a. Type	2 fam 5-6 rms		Singles 6-7 rms			
b. Construction	Frame		Frame			
c. Average Age	10-25 Years		10 Years		Years	
d. Repair	Fair		Good			
e. Occupancy	99 %		100 %		%	
f. Home ownership	80 %		95 %		%	
g. Constructed past yr.	0		0			
h. 1929 Price range	$ 10,500-13,500	100%	$ 8500-10,500	100%	$	100%
i. '33-36 Price range	$ 7,500- 9,000	70 %	$ 6500- 7,500	75 %	$	%
j. 1937 Price range	$ 7,500- 9,000	70 %	$ 6500- 7,500	75 %	$	%
k. Sales demand	$		$		$	
l. Activity	Poor		Poor			
m. 1929 Rent range	$ 45 - 65	100%	$ Owner occ.	100%	$	100%
n. '33-36 Rent range	$ 30 - 45	68%	$	%	$	%
o. 1937 Rent range	$ 30 - 45	68%	$ 55	%	$	%
p. Rental demand	$ Anything		$ Few available		$	
q. Activity	Good		.. ..			

4. AVAILABILITY OF MORTGAGE FUNDS: a. Home purchase Yes ; b. Home building ___
   On singles - somewhat limited on 2 family
5. CLARIFYING REMARKS: This area is situated on top of two separate hills and protected thus from the 3rd grade area around it. Houses are fairly modern and the neighborhood continues as a desirable one.
6. NAME AND LOCATION Brighton - Allston SECURITY GRADE B AREA NO. 1

(b) HOLC Zone "D"



FORM 8
10-1-37

AREA DESCRIPTION - SECURITY MAP OF Greater Boston, Mass.

1. AREA CHARACTERISTICS:
   a. Description of Terrain. Level
   b. Favorable Influences. Good transportation, schools, etc. Convenient to employment in the vicinity.
   c. Detrimental Influences. Mixture of houses and business. Low class occupants.
   d. Percentage of land improved 95 %; e. Trend of desirability next 10-15 yrs. Down
2. INHABITANTS:
   a. Occupation Labor - relief ; b. Estimated annual family income $600-1500
   c. Foreign-born families 50 %; Italian predominating; d. Negro Yes ; 5 %
   e. Infiltration of Foreign- negro ; f. Relief families Heavy
   g. Population is increasing ___ ; decreasing ___ ; static Yes
3. BUILDINGS:

	PREDOMINATING 30 %		OTHER TYPE 30 %		OTHER TYPE 30 %	
a. Type	Singles 7-10 rms		Doubles 7-8 rms		3 fam 6-7 rms	
b. Construction	Frame		Frame		Frame - few brick	
c. Average Age	40 Years		15-40 Years		25-50 Years	
d. Repair	Poor		Poor, some fair		Poor	
e. Occupancy	95 %		95 %		95 %	
f. Home ownership	50 %		65 %		Low %	
g. Constructed past yr.	0		0		0	
h. 1929 Price range	$ Up to $7500	100%	$ 6000-9000	100%	$ 4000-10,000	100%
i. '33-36 Price range	$ Up to $5000	67 %	$ 3500-7000	75 %	$ 3000- 7,000	70 %
j. 1937 Price range	$ Up to $5000	67 %	$ 3500-7000	75 %	$ 3000- 7,000	70 %
k. Sales demand	$ None		$ None		$ None	
l. Activity	Poor		Poor		Poor	
m. 1929 Rent range	$ 35 - 50	100%	$ 30 - 40	100%	$ 25 - 60	100%
n. '33-36 Rent range	$ 25 - 40	70 %	$ 25 - 35	80 %	$ 18 - 50	75 %
o. 1937 Rent range	$ 25 - 40	70 %	$ 25 - 35	80 %	$ 18 - 50	75 %
p. Rental demand	$		$		$	
q. Activity	Poor		Poor		Poor	

4. AVAILABILITY OF MORTGAGE FUNDS: a. Home purchase None ; b. Home building ___
5. CLARIFYING REMARKS: Negro concentrated around Empire St. The better 3-family units are heated, and rentals include such. Some very poor tenement houses are scattered throughout. Originally a good section with some large houses now given over to rooming house use.
6. NAME AND LOCATION Boston - Brighton - Allston SECURITY GRADE D AREA NO. 1

Table C.1: Hedonic Regression and House Flipping
Probability of Paying with Cash

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Back End of the Flip	.569	.516	.517	.188	.219
	(.004)	(.004)	(.004)	(.005)	(.007)
Front End of the Flip	.049	.002	.002	.027	.029
	(.004)	(.004)	(.002)	(.004)	(.007)
Month & Year Fixed Effects	No	Yes	Yes	Yes	Yes
Zip Codes Fixed Effects	No	No	Yes	Yes	No
House Characteristics	No	No	No	Yes	Yes
Property Fixed Effects	No	No	No	No	Yes
# of Observations	93,932	93,932	93,905	90,047	56,037

Note: Standard errors (in parentheses) are clustered at the property level. * p 0.05, ** p 0.01, and *** p 0.001. House characteristics include the sale price of the property, age of the property, type of deed, the sale type, living area of the property in square feet, size of the property in square feet, number of bedrooms in the property, number of bathrooms in the property, and number of parking spaces.

Table C.2: Hedonic Regression and House Flipping,
Probability of the Buyer Being an Investor

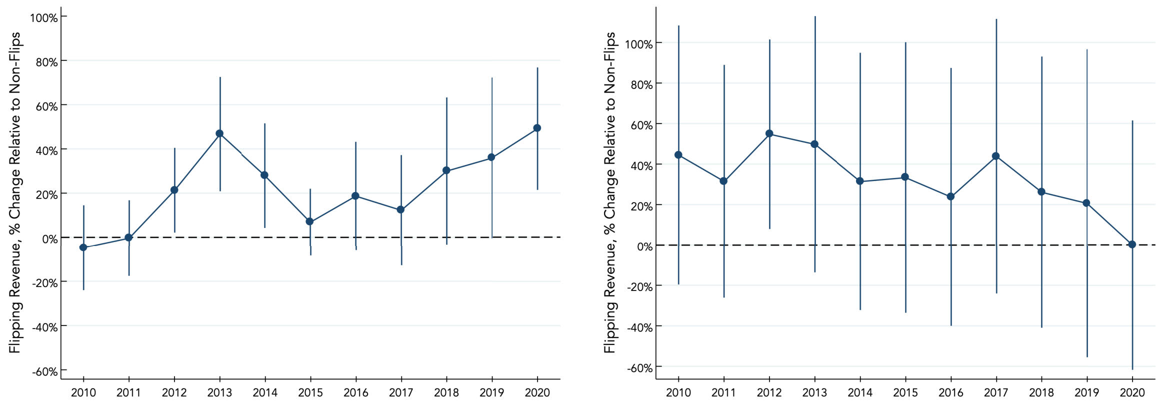
	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Back End of the Flip	.185	.188	.196	.180	.182
	(.004)	(.004)	(.004)	(.006)	(.008)
Front End of the Flip	-.016	-.017	-.009	-.039	-.071
	(.003)	(.003)	(.003)	(.003)	(.007)
Month & Year Fixed Effects	No	Yes	Yes	Yes	Yes
Zip Codes Fixed Effects	No	No	Yes	Yes	No
House Characteristics	No	No	No	Yes	Yes
Property Fixed Effects	No	No	No	No	Yes
# of Observations	93,932	93,932	93,919	90,056	56,037

Note: Standard errors (in parentheses) are clustered at the property level. * p 0.05, ** p 0.01, and *** p 0.001. House characteristics include the sale price of the property, age of the property, type of deed, living area of the property in square feet, size of the property in square feet, number of bedrooms in the property, number of bathrooms in the property, and number of parking spaces.

Figure C.2: Hedonic Regression and House Flipping,
Change in Total Revenue by Year and Approved Property Permit

Percentage Change

(a) **Without** Approved Property Permit (b) **With** Approved Property Permit



Note: Vertical lines crossing the estimates are confidence intervals, where the cap represents the confidence interval at the 95% level.

Table C.3: Change in the Sale Price, by HOLC Zone

	Model (1)	Model (2)	Model (3)	Model (4)
Relative to Zone A				
Zone B	-.283	-.283	-.248	-.184
	(.010)	(.010)	(.008)	(.009)
Zone C	-.534	-.536	-.362	-.301
	(.009)	(.009)	(.008)	(.009)
Zone D	-.452	-.456	-.381	-.324
	(.010)	(.010)	(.009)	(.010)
Year Fixed Effects	No	Yes	Yes	Yes
Zip Codes Fixed Effects	No	No	Yes	Yes
House Characteristics	No	No	No	Yes
# of Observations	154,235	154,235	154,232	142,262

Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. House characteristics include the age of the property, type of deed, living area of the property in square feet, size of the property in square feet, number of bedrooms in the property, number of bathrooms in the property, and number of parking spaces.

Table C.4: Cash Payment Probability, by HOLC Zone

	Model (1)	Model (2)	Model (3)	Model (4)
Relative to Zone A				
Zone B	.026	.027	-.003	.00
	(.006)	(.006)	(.006)	(.006)
Zone C	.043	.044	.003	-.004
	(.005)	(.005)	(.006)	(.006)
Zone D	.043	.045	-.007	-.024
	(.005)	(.005)	(.007)	(.007)
Year Fixed Effects	No	Yes	Yes	Yes
Zip Codes Fixed Effects	No	No	Yes	Yes
House Characteristics	No	No	No	Yes
# of Observations	154,235	154,235	154,232	142,262

Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. House characteristics include the sale price of the transaction, age of the property, type of deed, living area of the property in square feet, size of the property in square feet, number of bedrooms in the property, number of bathrooms in the property, and number of parking spaces.

Table C.5: Foreclosure Probability, by HOLC Zone

	Model (1)	Model (2)	Model (3)	Model (4)
Relative to Zone A				
Zone B	.0013	.0014	.0002	-.0006
	(.0008)	(.0008)	(.0009)	(.0009)
Zone C	.0113	.0115	.0024	.0023
	(.0008)	(.0008)	(.0009)	(.0010)
Zone D	.0096	.0099	.004	.0041
	(.0009)	(.0009)	(.0013)	(.0014)
Year Fixed Effects	No	Yes	Yes	Yes
Zip Codes Fixed Effects	No	No	Yes	Yes
House Characteristics	No	No	No	Yes
# of Observations	154,235	154,235	154,232	142,262

Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. House characteristics include the sale price of the transaction, age of the property, type of deed, living area of the property in square feet, size of the property in square feet, number of bedrooms in the property, number of bathrooms in the property, and number of parking spaces.

Table C.6: Investor Buyer Probability, by HOLC Zone

	Model (1)	Model (2)	Model (3)	Model (4)
Relative to Zone A				
Zone B	.0009	.0009	.0003	.0000
	(.0011)	(.0011)	(.0012)	(.0012)
Zone C	.0012	.0012	-.0011	-.0012
	(.001)	(.001)	(.0012)	(.0012)
Zone D	.001	.001	.0001	.0002
	(.001)	(.001)	(.0013)	(.0014)
Year Fixed Effects	No	Yes	Yes	Yes
Zip Codes Fixed Effects	No	No	Yes	Yes
House Characteristics	No	No	No	Yes
# of Observations	154,235	154,235	154,232	142,262

Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. House characteristics include the sale price of the transaction, age of the property, type of deed, living area of the property in square feet, size of the property in square feet, number of bedrooms in the property, number of bathrooms in the property, and number of parking spaces.

Table C.7: Spatial Regression Discontinuity Design with IPW and Different Estimators, Probability of a Home being Flipped

	A-B Border	A-C Border	B-C Border	C-D Border
Conventional	0.040	0.109	0.029	-0.025
	(0.033)	(0.059)	(0.015)	(0.022)
Bias-corrected	0.029	0.122*	0.035*	-0.033
	(0.033)	(0.059)	(0.015)	(0.022)
Robust	0.029	0.122	0.035	-0.033
	(0.039)	(0.072)	(0.018)	(0.026)
Obs.	902	392	9,980	10,376

Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. Local polynomial analysis. The SRDD estimator uses as covariates the sale price, and zipcode and year fixed-effects.

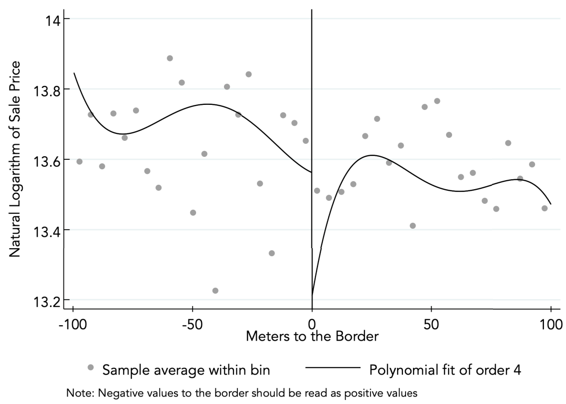
Table C.8: Spatial Regression Discontinuity Design with IPW and Different Estimators, Impact on the Sale Price of a Property Transaction

	A-B Border	A-C Border	B-C Border	C-D Border
Conventional	-0.039	1.575***	-0.083	0.002
	(0.004)	(0.410)	(0.005)	(0.000)
Bias-corrected	-0.066	1.935***	-0.095	-0.000
	(0.006)	(0.503)	(0.006)	(0.000)
Robust	-0.066	1.935***	-0.095	-0.000
	(0.008)	(0.540)	(0.007)	(0.000)
Obs.	931	407	10,272	10,871

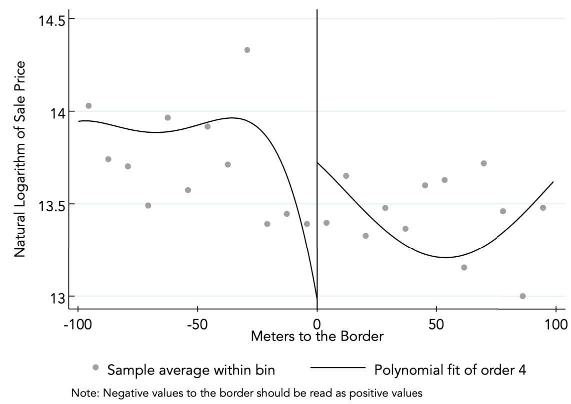
Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. Local polynomial analysis. The SRDD estimator uses as covariates zip code and year fixed-effects.

Figure C.3: Spatial Regression Discontinuity Design with IPW, for “A - B” and “A - C” Borders, Change in the Sale Price

(a) “A - B” Border



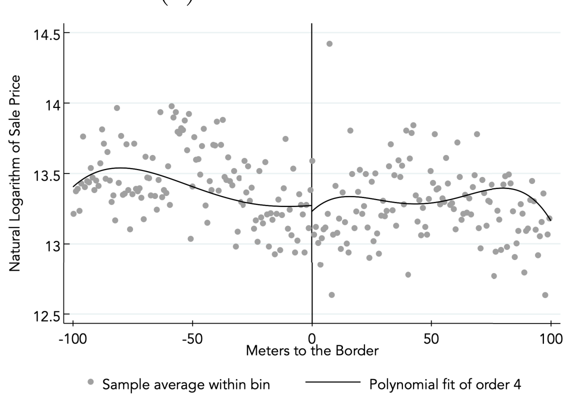
(b) “A - C” Border



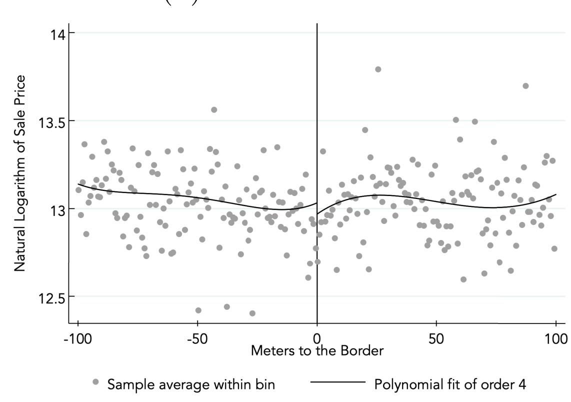
Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and to the right represents the fitted polynomial.

Figure C.4: Spatial Regression Discontinuity Design with IPW, for “B - C” and “C - D” Borders, Change in the Sale Price

(a) “B - C” Border



(b) “C - D” Border



Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and to the right represents the fitted polynomial.

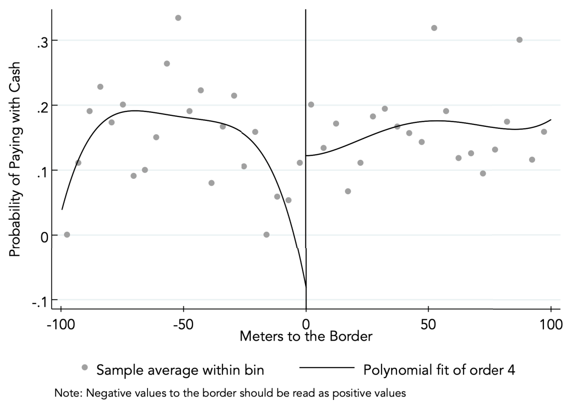
Table C.9: Spatial Regression Discontinuity Design with IPW
and Different Estimators, Probability of Paying with Cash a Property Transaction

	A-B Border	A-C Border	B-C Border	C-D Border
Conventional	0.055	0.116	0.015	0.037
	(0.074)	(0.110)	(0.035)	(0.032)
Bias-corrected	0.014	0.103	0.014	0.037
	(0.074)	(0.110)	(0.035)	(0.032)
Robust	0.014	0.103	0.014	0.037
	(0.089)	(0.134)	(0.041)	(0.038)
Obs.	931	407	10,272	10,871

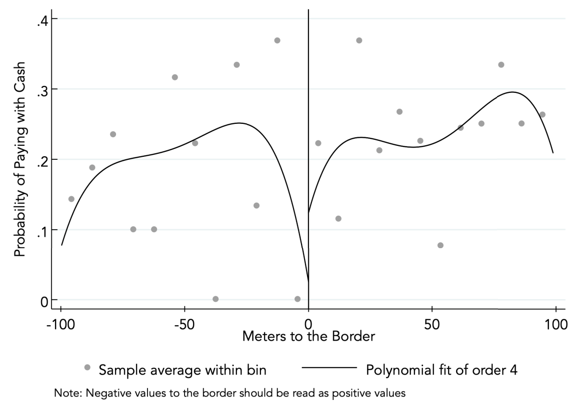
Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. Local polynomial analysis. The SRDD estimator uses as covariates the sale price, zip code, and year fixed-effects.

Figure C.5: Spatial Regression Discontinuity Design with IPW, for “A - C” and “A - C” Borders, Probability of Paying with Cash a Transaction

(a) “A - B” Border



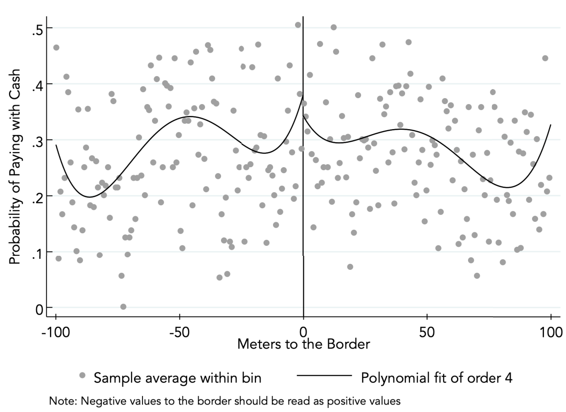
(b) “A - C” Border



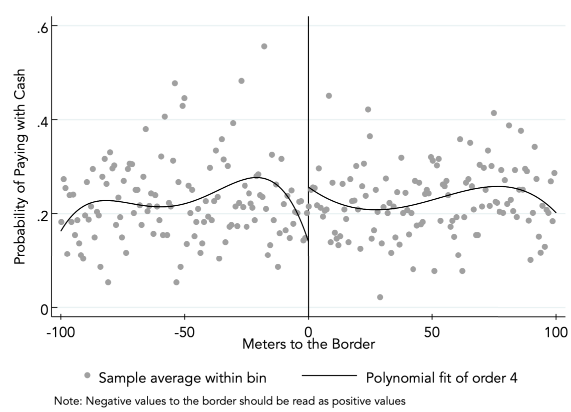
Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and to the right represents the fitted polynomial.

Figure C.6: Spatial Regression Discontinuity Design with IPW, for “B - C” and “C - D” Borders, Probability of Paying with Cash a Transaction

(a) “B - C” Border



(b) “C - D” Border



Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and to the right represents the fitted polynomial.

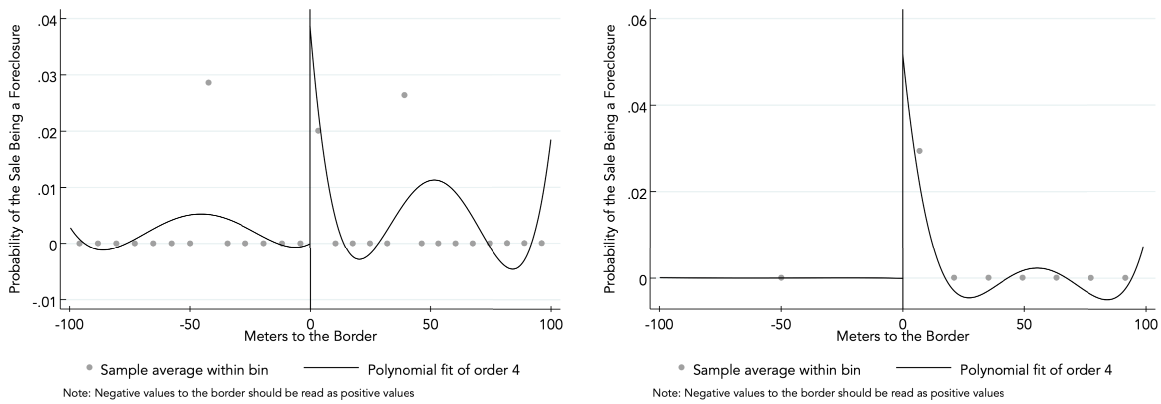
Table C.10: Spatial Regression Discontinuity Design with IPW
and Different Estimators, Probability of a Property Transaction Being a Foreclosure

	A-B Border	A-C Border	B-C Border	C-D Border
Conventional	0.066*	0.046	0.004	0.006
	(0.027)	(0.041)	(0.005)	(0.007)
Bias-corrected	0.076**	0.048	0.005	0.007
	(0.027)	(0.041)	(0.005)	(0.007)
Robust	0.076*	0.048	0.005	0.007
	(0.033)	(0.049)	(0.006)	(0.008)
Obs.	931	407	10,272	10,871

Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. Local polynomial analysis. The SRDD estimator uses as covariates the sale price, zip code, and year fixed-effects.

Figure C.7: Spatial Regression Discontinuity Design with IPW, for “A - B” and “A - C” Borders, Probability of the Sale Being a Foreclosure

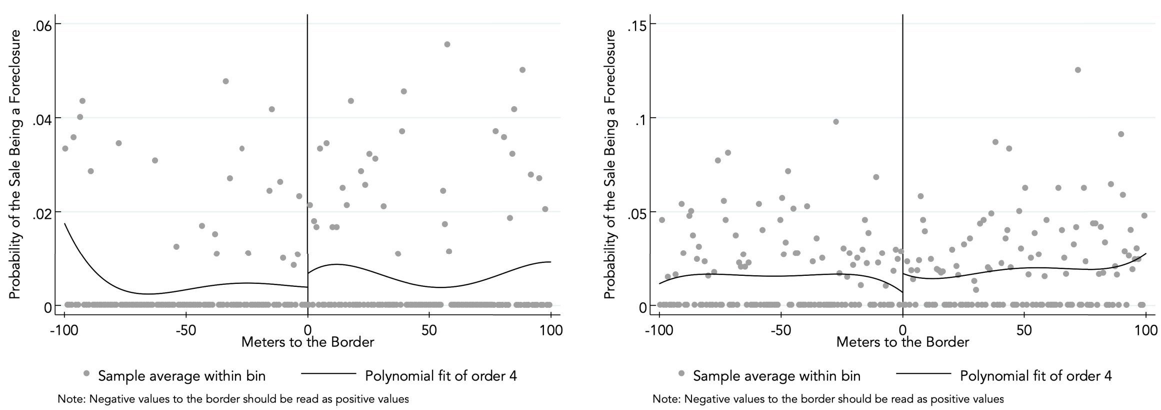
(a) “A - B” Border (b) “A - C” Border



Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and to the right represents the fitted polynomial.

Figure C.8: Spatial Regression Discontinuity Design with IPW, for “B - C” and “C - D” Borders, Probability of the Sale Being a Foreclosure

(a) “B - C” Border (b) “C - D” Border



Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and to the right represents the fitted polynomial.

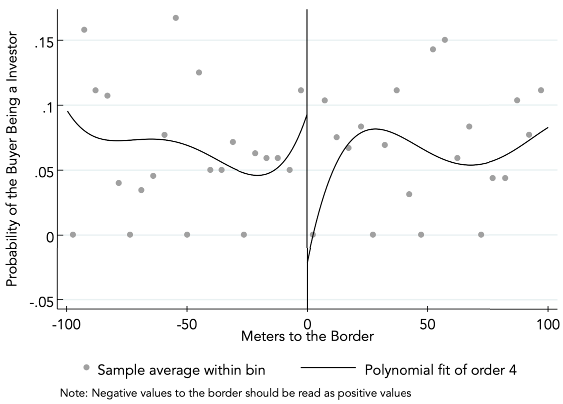
Table C.11: Spatial Regression Discontinuity Design with IPW and Different Estimators, Probability of a Property Transaction Being Bought by an Investor

	A-B Border	A-C Border	B-C Border	C-D Border
Conventional	-0.085	-0.128	-0.023	0.005
	(0.067)	(0.165)	(0.021)	(0.024)
Bias-corrected	-0.119	-0.121	-0.028	0.007
	(0.067)	(0.165)	(0.021)	(0.024)
Robust	-0.119	-0.121	-0.028	0.007
	(0.077)	(0.196)	(0.026)	(0.029)
Obs.	909	400	10,081	10,699

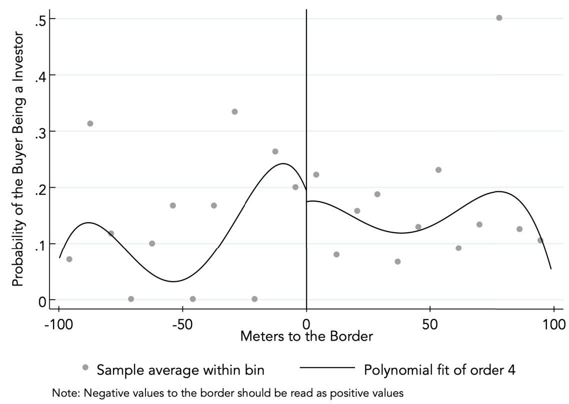
Note: Standard errors (in parentheses) are clustered at the ZIP code level. * p 0.05, ** p 0.01, and *** p 0.001. Local polynomial analysis. The SRDD estimator uses as covariates the sale price, zip code, and year fixed-effects.

Figure C.9: Spatial Regression Discontinuity Design with IPW,
for "A - B" and "A - C" Borders, Probability of the Buyer Being an Investor

(a) "A - B" Border



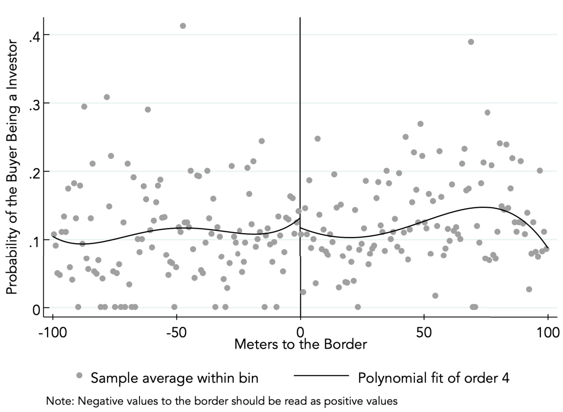
(b) "A - C" Border



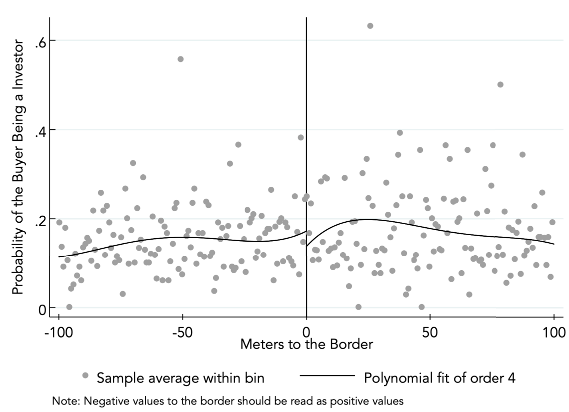
Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and to the right represents the fitted polynomial.

Figure C.10: Spatial Regression Discontinuity Design with IPW,
for "B - C" and "C - D" Borders, Probability of the Buyer Being an Investor

(a) "B - C" Border



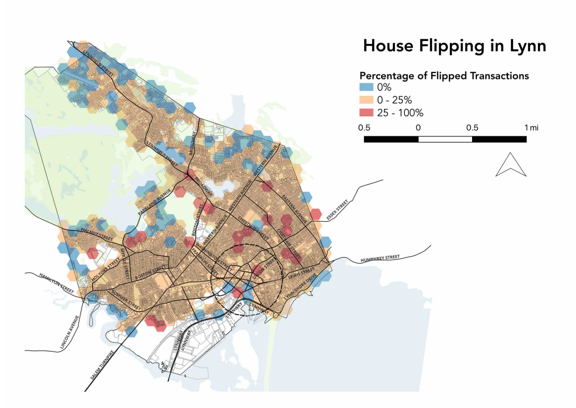
(b) "C - D" Border



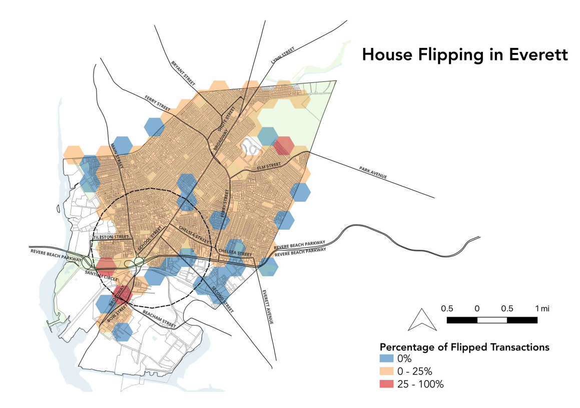
Note: Dots represent average values within the distance shown on the x-axis. The vertical line represents the border. The function to the left and to the right represents the fitted polynomial.

Figure C.11: Raw Maps used during the Collaborative Data Analysis

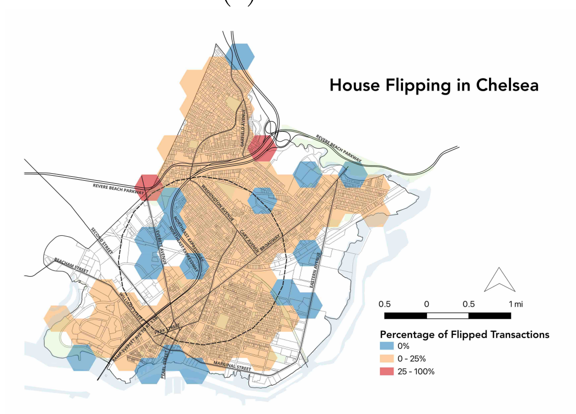
(a) Lynn



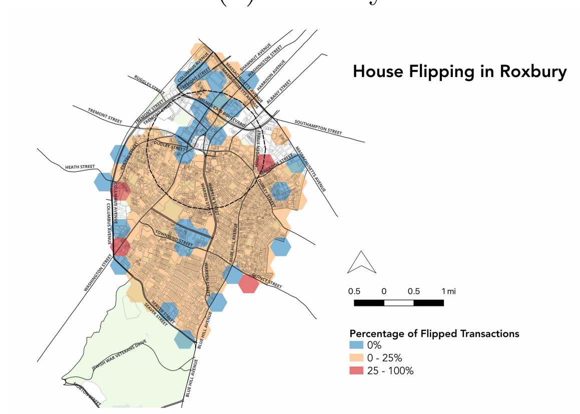
(b) Everett



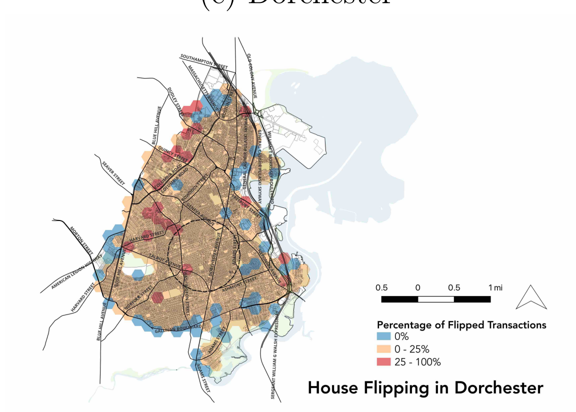
(c) Chelsea



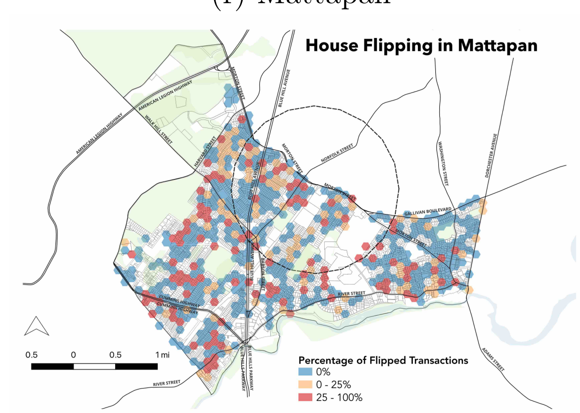
(d) Roxbury



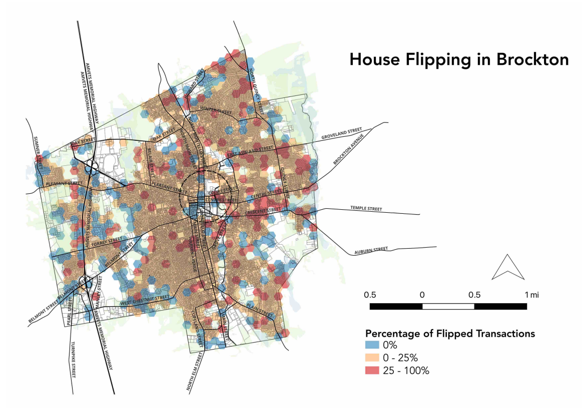
(e) Dorchester



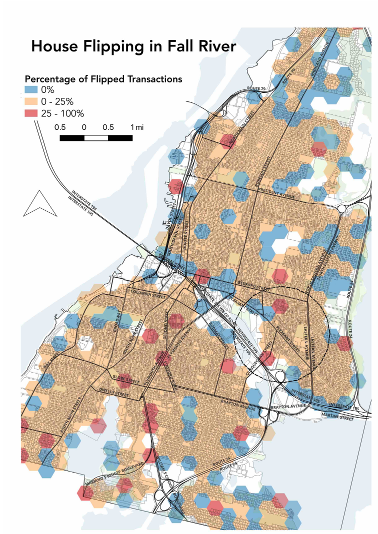
(f) Mattapan



(g) Brockton



(h) Fall River



(i) New Bedford