THREE ESSAYS IN HOUSING

by

Sean Kodama Brunson

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Business Administration

Charlotte

2020

Approved by:

____

Dr. Richard J. Buttimer Jr.

____

Dr. Yongqiang Chu

____

Dr. Kiplan Womack

____

Dr. Thomas Mayock

____

Dr. Craig A. Depken, II

ABSTRACT

SEAN KODAMA BRUNSON. Three essays in housing. (Under the direction of DR. RICHARD J. BUTTIMER JR.)

This dissertation addresses the potential mismeasurement of home price indices and its implications on mortgage default research during the recession and examines the potential drivers of the housing affordability issue since the recession. The first essay provides a rational reason why the distribution of home sellers is drastically different during economic downturns than during normal economic times. I show that when the economy enters into a recession, selection bias may affect observed home transaction prices. Since home price indices use these observed transaction prices, attempts to impute the value of every home in the housing stock is incorrect and will lead to inference issues. The second essay examines the entrance of institutional investors in the single-family housing market. I empirically test whether the entrance of institutional investors contributed to the subsequent increase in home prices after the recession. I find that institutional investors paid a discount of about 8.13%-11.19% per transaction. Additionally, I find that an increase in institutional investor home purchases in the single-family housing market had a positive statistical impact on individual home prices but only a moderate economic impact. The third essay investigates the impact of increased development fees in the residential housing market of York county, South Carolina. Using the increased development fees in the Fort Mill school district located within York county, I test for difference-in-differences in listing prices, closing price, and inventory after two separate increases in development fees. The evidence suggests that increased development fees correspond with a decrease in asking and closing prices and an increase in available inventory in the period immediately after the increase.

DEDICATION

This dissertation is dedicated to my loving wife, Katherine, my parents, Andy and Mie, and my brother, Ken, who continuously pushed and encouraged me to finish this massive project.

ACKNOWLEDGEMENTS

I cannot thank Dr. Richard J. Buttimer Jr. enough for the amount of time spent helping me improve my dissertation. I would also like to acknowledge Dr. Craig A. Depken, II and Dr. Yongqiang Chu for their deep econometric expertise, as well as Dr. Kiplan Womack and Dr. Thomas Mayock for their housing market insights. Finally, I would like to thank Metrostudy and the Carolina Regional Realtors Association for providing the data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Since the Great Recession ended in 2009, the housing market has been in the forefront of both economic research and public policy. Much of this research and the policies implemented have focused on two important areas: mortgage default and housing affordability. The aim of this dissertation is to take a deeper dive into the housing market both during and after the Great Recession. Specifically, this dissertation addresses the potential mismeasurement of home price indices and its implications on mortgage default research during the recession and examines the potential drivers of the housing affordability issue since the recession.

There is currently a large disconnect on the use of home price indices. On one side, papers such as Haurin and Hendershott (1991) and Case and Quigley (1991) argue that indices on illiquid assets are subject to substantial selection bias and that uses of these indices with no correction will lead to inference issues when measuring the value of the aggregate stock of that asset. On the other side, recent papers such as Bhutta et al. (2017) and Laufer (2018) have used these non-corrected indices to measure the aggregate stock value. The disconnect arises because the set of papers that allude to the selection bias are purely empirical and may have measurement errors within the data or methodology. Thus, it is not clear whether the selection bias issue will always affect home price indices.

In addition, decision differences among current homeowners may also be amplified depending on the fluctuations in the aggregate economy. During the mid 1990s and early 2000s, there was a massive surge in mortgage credit. At the same time, the U.S. housing market saw a significant rise in home prices. However, shortly afterward, the U.S. saw an unprecedented level of mortgage defaults as well as a dramatic fall in home prices. The current consensus is that the loosening of credit access in the U.S., especially in the mortgage market, was a contributing factor to the large increase in home prices in the early 2000s as well as a contributing factor to the subsequent fall of home prices during the 2007-2009 recession.

Given the easier access to mortgage credit during the early 2000s, the distributional flow of individuals into and out of the housing market was quite high. During this boom period, observed transaction prices were more likely to reflect the behavior of the population as a whole because the distribution of individuals transacting was larger. However, only a small percentage of homes are sold each period, so a few home sellers can have a large impact on observed market prices. Once the recession occurred, even fewer homeowners voluntarily chose to sell their home. As a result, home price indices during the recession period were only capturing an extremely small subset of all homeowners. This small subset may not be representative of the entire housing market. In fact, since many of the homes selling during the recession came from foreclosures, which are known to sell at a discount, then the resulting home price index might be pushed downward.

The first chapter of this dissertation addresses this home price index mismeasurement. Using a structural equilibrium model of housing as a mixed consumption and investment asset, the first chapter asks the following questions: (1) Is the distribution of home sellers fundamentally different than the distribution of all homeowners? and (2) Does this distributional difference lead to a large separation between observed transaction prices and unobserved home values? The main result from this chapter is the distribution of home sellers is in fact fundamentally different than the distribution of all homeowners. Specifically, the results from this chapter suggest that individuals who voluntarily choose to sell are financially unconstrained relative to other homeowners in the area. Being financially unconstrained allows these individuals to move about freely, which also allows them the ability to take a lower selling price, especially during a recession.

As a result, this chapter also shows that this distributional difference does lead to a large separation between observed transaction prices and unobserved home values. If all homeowners voluntarily chose to sell their home, observed transaction prices would in fact be greater than they actually were. The results of this chapter suggest that home price indices may be biased downwards during recessions, which means that imputed values of homes that did not sell during the recession may also be biased downwards. This will have major implications on default models that take these imputed values as an input.

In the aftermath of the recession, mortgages became hard to obtain, even for well qualified individuals. However, since 2012, home prices have started to increase and have surpassed peak pre-recession prices in many areas across the U.S. Thus, the easier access to mortgage story is now not as applicable in explaining the increase in home prices. In addition, as home prices continue to grow faster than income, housing affordability concerns are back near the top for policy issues in the U.S. For this reason, some media outlets have started to put blame on the large, publicly traded companies (institutional investors) that have entered into the single-family housing market.[1]

These institutional investors, such as Real Estate Investment Trusts (REITs), saw that home prices relative to rents were low. They realized they could purchase single-family homes for the purpose of renting them out to the general population. Furthermore, while many individuals were unable to raise funds to purchase homes, institutional investors already had funds to do so, which meant they did not need to obtain a traditional mortgage for each home purchase. This gave institutional

[1] See `https://www.wsj.com/articles/investors-are-buying-more-of-the-u-s-housing-market-than-ever-before-11561023120`, `https://www.theatlantic.com/technology/archive/2019/02/single-family-landlords-wall-street/582394/`, `https://www.npr.org/2019/06/21/734357279/1st-time-homebuyers-are-getting-squeezed-out-by-investors`, or `https://www.nytimes.com/interactive/2019/06/20/business/economy/starter-homes-investors.html` to name a few.

investors a significant timing advantage in the housing market.

The perception in the popular press has been that these institutional investors entered the housing market and “out-bid" the traditional single-family home buyer. Yet, institutional investors are large, publicly traded companies that want to maximize shareholder value, so they have little incentive to overpay for housing. Furthermore, institutional investors were purchasing at a fragile time in the housing market, which may imply that they had higher bargaining power. Both of these mechanisms would lead to homes being purchased at a discount by institutional investors. However, the net effect that institutional investors have on home prices is unclear. If institutional investors purchased at a discount, they are clearly not raising prices on an individual basis. However, by purchasing homes, they are effectively decreasing the supply of the remaining homes, which would push home prices up on average as long as the remaining homes are being purchased and being purchased at a premium.

The second paper of this dissertation empirically examines the effect institutional investors had on single-family home prices in the Charlotte region between 2005-2017. As expected, the findings in this chapter suggest that institutional investors paid about 8.13%-11.19% less than owner-occupiers per transaction. Additionally, this chapter shows that an increase in institutional investor home purchases in the single-family housing market had a statistically positive impact on home prices but a moderate economic impact. That is, a 1 percentage point increase in the percentage of homes purchased by institutional investors in the previous 6-months increased individual home sale prices by about 0.32%-0.59%. The economic significance is about $4,611.63-$5,099.53 on average for a one standard deviation increase in institutional investor activity in the Charlotte region. Both of these results are consistent across the entire price distribution as well as across various census tract and racial/ethnic groups.

However, the second chapter sheds some concern that this positive price impact may not all be coming directly from institutional investor activity. In fact, the results suggest that a majority of the price increase is coming from a general increase in demand. Given the growing concerns of housing affordability in the U.S., the impact institutional investors have on the market, especially on individuals, is a top priority both economically and politically. Yet, the results of this chapter suggest that any resulting policy should focus on helping individuals rather than keeping institutional investors out of the single-family housing market.

With that being said, some areas have implemented policies that focus on groups other than institutional investors in order to accommodate home price increases and economic and population growth. In fact, some local areas have seen higher economic and population growth than others and must start thinking about infrastructure needs to accompany these growths. The main issue is how do these local governments finance the new infrastructure? One potential option is development fees, which are one time, per-unit fees paid by the developer to build housing and are used for a pre-determined purpose such as parks and recreation, municipal equipment, transportation, fire, or schools. Although, these types of fees are not new, the research on their housing market impact is small.

Additionally, most analysis has focused on areas in which the development fees only increased by at most a few thousand dollars. Since these fees are charged to the developer, which are mostly large companies with deep pockets, a few thousand dollars may not substantially affect their bottom line, nor affect future supply of housing. However, in 2017, the Fort Mill school district within York county, South Carolina conducted a study that suggested that development fees for single family and multi-family housing increase from $2,500 per-unit to $18,158 per-unit and $12,020 per-unit respectively. These development fees were ultimately implemented in 2018.

Given that the magnitude of this increase has not been seen in prior research, it is unclear whether the general findings of the impact of development fees on the housing market are applicable to the Fort Mill school district. In fact, the third chapter of this dissertation finds that prior research on the development fee's impact on existing home prices is not applicable. Specifically, previous literature found that after an increase in development fees, both prices on new and existing homes increased. However, the results of the third chapter show that after the collection of the development fee increases began, only prices on new homes increased, while prices on existing homes decreased. The most likely reason is that the increase in the development fee to $18,158 may have led consumers to think growth will significantly stall as home developers exit the Fort Mill housing market. This would reduce the incoming revenue for the local government so that in response, taxes must increase to pay for the new infrastructure. This is further supported by the results, which show that the demand for existing housing decreased after the fees were implemented.

This dissertation is a collection of three papers surrounding the housing market. The main focus is examining home prices during and after the Great Recession. Although research on the housing market is dense, this dissertation shows that there are ample areas for further study. The rest of the dissertation is as follows: Chapter 2 examines home price index mismeasurement, Chapter 3 examines institutional investors in the single-family housing market, Chapter 4 examines the increase in developments fees in York county, South Carolina, and Chapter 5 includes the closing remarks.

CHAPTER 2: DO HOMEOWNER DECISIONS MATTER? IMPLICATIONS ON HOME PRICE INDICES

2.1 Introduction

There is currently a large disconnect on the use of home price indices. On one side, there is a large literature that argues that indices on illiquid assets are subject to substantial selection bias and that uses of these indices with no correction will lead to inference issues when measuring the value of the aggregate stock of that asset.[1] On the other side, there is a literature that uses these non-corrected indices to measure the aggregate stock value.[2] We bridge this gap by showing whether the self-selection issue has a substantial effect on observed home prices and whether these observed home prices can be used to make inferences about homes that did not sell. Specifically, we provide a theoretical foundation that shows when aggregate conditions in the economy change, the distribution of home sellers is significantly different than homeowners who do not sell, which leads to a large difference between observed and unobserved home values.

The disconnect arises because the set of papers that allude to the selection bias are purely empirical and may have measurement errors within the data or methodology. This has produced mixed results in how certain sellers will affect the indices. For example, "starter homes" transact more frequently, so their observed transactions may make up a larger portion of the data. If the prices of these homes appreciate at a different rate than non-starter homes, then the home price indices will reflect this

[1] Just to name a few: See Haurin and Hendershott (1991), Case and Quigley (1991), Gatzlaff and Haurin (1997), Gatzlaff and Haurin (1998), and Hwang and Quigley (2004) for the housing market and Goetzmann (1993) for the art market.

[2] See Bhutta et al. (2017) and Laufer (2018) to name two of the most recent studies that use home price indices in this way.

difference. However, empirical results for this type of scenario are split.[3] In other words, it is not clear whether the type of home that is being sold, or the type of individual selling, can fundamentally shift home prices. If so, then it is even more unclear whether the magnitudes of these shifts are significant or not.

Homes are the largest asset that individuals will have in their lifetime. Most home buyers are liquidity constrained such that they must use a mortgage to finance their purchase. This mortgage will make up a large portion of an individual's liability. The interaction between homeowner decisions over mortgages (default or not) and housing transaction behavior (sell vs. stay) strongly depends on the value the homeowner places on the home. This value is what we will refer to as the internal value of the homeowner. Note that this internal value may differ from the imputed home value from a transaction based home price index. The imputed value is what the market would deem to be an appropriate transaction price of a given home conditional on that home/homeowner being similar to the homes/homeowners that actually sold in that period. The argument that we make in this study is that although individuals may agree upon the imputed value of homes in an area, it is the internal value that drives their actual decisions. Furthermore, this internal value will vary substantially given the heterogeneous nature of both the individuals and the home types in the market.

Decision differences among current homeowners may also be amplified depending on the fluctuations in the aggregate economy. During the mid 1990s and early 2000s, there was a massive surge in mortgage credit. Much of the economic and financial literature has come to a consensus that credit, especially mortgages, was easier to obtain, which led to a substantial increase in home prices during the 2000s.[4] Landvoigt et al. (2015) shows that selling volume during the boom increased across all home

[3]See Pollakowski et al. (1991) and Smith and Tesarek (1991).

[4]See Mian and Sufi (2009), Mayer et al. (2009), Keys et al. (2010), Landvoigt (2017) to name a few.

prices, suggesting that easier credit access allowed individuals, especially higher risk individuals, to move about more freely leading to an increase in home prices. Thus, it was clear that during this time period, the distributional flow of individuals into and out of the housing market was quite high. During this boom period, observed transaction prices are more likely to reflect the behavior of the population as a whole because the distribution of individuals transacting was larger.

However, only a small percentage of homes are sold each period, so a few home sellers can have a large impact on observed market prices. According to the Census, since 1988, only 8.8% of homes sold annually in the U.S. Once the bust occurred in 2007-2009, transaction volume decreased to under 5% annually as individuals had less incentive to sell their home. Demand from new buyers also decreased as individuals were less willing and able to enter into the housing market. This left a relatively small subset of homeowners who either chose to sell or were forced to put their home on the market, where the latter may have stemmed from foreclosure.

Most theoretical and empirical evidence suggests that negative equity reduces mobility of current homeowners.[5] If home price indices show a decline in home values in an area, then this may imply that current homeowners will be more likely to stay in their home and not sell, further separating the dynamics of sellers with the rest of the homeowner population. Previous research, along with the fact that a higher proportion of homeowners were highly leveraged going into the recession, would suggest that those placing homes on the market are fundamentally different than the rest of the homeowners. Observed market prices may not be a good representation of the unobserved value of the homes that did not sell. For this reason, it is hard to justify the use of transaction-based home price indices to measure anything other than the current period's transaction price unless the index has been adjusted for this self-selection bias.

[5] See Quigley (1987), Stein (1995), Genesove and Mayer (1997), Genesove and Mayer (2001), Chan (2001), Engelhardt (2003), and Ferreira et al. (2010) to name a few.

Our study answers two questions: Is the distribution of home sellers fundamentally different than the distribution of all homeowners? Does this distributional difference lead to a large separation between observed transaction prices and unobserved home values? To answer these questions, we must note that unlike financial assets, people purchase housing-units for both consumption and investment purposes. The market is characterized by various costs, and most home buyers are liquidity constrained such that they must use a mortgage to finance their purchase. The central idea is that as markets rise and fall, these stylized facts mentioned change the distribution of individuals that are in the market to sell their home. We argue that this change alters how observed market data should be used to value the overall housing stock. Specifically, we show that as an economy enters into a recession, homeowners rationally choose not to sell their home. Those that do sell are those that are willing and able to sell at low prices (i.e. those that are not liquidity constrained). We show that this underlying behavior has a profound effect on the appropriate use of housing market data to create indices and a profound effect on the appropriate use of those indices, specifically in measuring mortgage defaults.

To gain insight into homeowner decisions, we construct a structural equilibrium model of housing as a mixed consumption and investment asset. Through our equilibrium model, we analyze how both aggregate and idiosyncratic shocks affect individual's decisions to enter or exit the housing market. We find that when the economy enters into a recession, relatively *unconstrained* homeowners will rationally choose to sell their home. This is the case because during recessions, market prices become volatile, which means that owning a home is risky. Since homeowners in the model are forward-looking, instead of staying in the home during the recession, relatively *unconstrained* homeowners sell just prior to the recession, rent, and purchase homes just prior to when housing demand rebounds.

On the other hand, homeowners that are relatively *constrained* (i.e. those with high mortgage costs or low savings/income) are more likely to stay in their current home. Our model includes two important features to support this rational behavior: a rental market and a long-term mortgage. Agents in our economy can only borrow through a mortgage, which is subject to both a loan-to-value constraint as well as a payment-to-income constraint. If the economy enters into a recession, the equilibrium home price decreases, which lowers the probability that a homeowner will satisfy the loan-to-value constraint. In addition, if the homeowner is also hit with a negative shock to income, then the probability of obtaining a mortgage further decreases. For these reasons, homeowners will have difficulty obtaining a new mortgage to buy a home and will choose to stay in their current home. In addition, the mortgage in the model is a long-term, fully amortizing loan, which means that mortgage payments are spread out over the life of the loan. Unless the homeowner is hit with extremely negative shocks to wealth, the individual will choose to stay in the current home rather than to sell or to default. However, if the homeowner defaults, he/she is typically precluded from buying another home for some period of time and must rent instead. Unlike owner-occupied housing, rental units do not provide the same type of consumption benefits so choosing to rent may decrease the individual's lifetime utility value. Thus, we should expect homeowners with relatively high mortgage costs to stay in their current home rather than default and move to a rental unit.

If a large subset of individuals experience massive negative shocks at the same time, then there will be an initial increase in supply, from home sales out of default, which will decrease equilibrium prices. If there is a continued drop in aggregate income in the subsequent years, housing demand will be low, which will further depress prices leading to the scenario that the U.S. saw during the 2007-2009 period. During this period, mortgage defaults increased, which implies that there were more potentially liquidity constrained homeowners in the market. These homeowners were left with three options: stay in their current home, sell their current home, or default on their mortgage. However, in addition to the argument above, sellers may choose not to default because of high moral aversion to default (Guiso et al. (2013), Bhutta et al. (2017)). This suggests that homeowners will be even more willing to stay in their current home than default or sell.

In line with our study, Stein (1995) suggests that, in general, when home prices decrease, current homeowners will choose not to sell because they may not be able to cover the down payment on a future home. If homeowners do choose to sell, then they will "fish" by setting a listing price high in hopes that a buyer will come in and buy at that high price.[6] We find similar results such that when prices decrease, many relatively *constrained* homeowners will choose not to sell because it will be difficult to obtain a new mortgage. Although we do not explicitly examine the "fishing" behavior, we find that homeowners with a loan under an 80% LTV will have a lower reservation price than homeowners with a loan above an 80% LTV. That is, relatively *unconstrained* homeowners will have a lower internal reservation price that makes them indifferent between selling and staying, so these individuals are more likely the ones voluntarily selling their home. As the LTV goes further into the negative equity territory (LTV $>$ 100%) and the individual becomes more financially constrained, reservation prices start to increase and become higher than the reservation prices of the relatively *unconstrained* homeowners. This is the case because relatively *constrained* homeowners need to sell at a higher price in order to cover the high costs of a mortgage. Since potential buyers would rather purchase at lower prices, relatively *constrained* homeowners have a hard time selling and ultimately leave the market to sell the home. Additionally, negative aggregate shocks to the economy will remove potential buyers from the market making it difficult for current homeowners to sell

[6]Genesove and Mayer (1997), Genesove and Mayer (2001) empirically test this theory and find that homeowners with high LTV loans are more likely to set a higher listing price, sell at a higher price, and be on the market longer than homeowners with lower LTV loans. The authors attribute most of this behavior to nominal loss aversion in which they say sellers want to make up any potential loss from purchase to subsequent sale by setting the listing price above the market price.

at relatively high prices. Thus, the majority of the homes that are on the market to sell will be made up of lower LTV homeowners, who are relatively *unconstrained* homeowners.

In fact, Guren (2018), empirically and theoretically using a search model, shows that sellers cannot set the listing price too high nor too low because it will affect the length of time to sell the home. The reasoning is if the seller sets the listing price too high, then potential buyers will not consider the home because there may be a better deal in the market. However, if the seller sets the listing price too low, then the probability of selling faster increases, but the potential revenue from the sale will decrease. Guren (2018) ultimately shows that this type of interaction produces momentum of home prices that is seen empirically.[7] Thus, rational individuals will be weighing the time costs associated with selling the home. If the homeowner is unable to sell at a later date, then the seller may choose to sell at a lower price.

These dynamics are important because home price indices that gauge the real estate market use these home transactions as inputs. Most of these indices, which are built off a repeat sales method proposed by Bailey et al. (1963) and popularized by Case and Shiller (1987) and Case and Shiller (1989), are meant to measure the average change in home prices in a given area. However, for highly illiquid goods, such as housing, transaction prices may be misleading because there are only a few market participants and a few transactions each period. Any observed price will be heavily driven by the number of transactions and by the distribution of individuals in the market, all of which will depend on the economic environment. If it is the case that homeowners who sell at this time are relatively *unconstrained*, then home price indices will capture a certain distribution of individuals that may not represent the area as a whole. In fact, during economic downturns, these indices might be picking

[7] Case and Shiller (1989), Head et al. (2014), and others have shown that home prices exhibit strong momentum. Capozza et al. (2004), Glaeser et al. (2014), and others have shown that home prices exhibit mean reversion in the long run.

up relatively *unconstrained* individuals selling homes at a low price putting downward pressure on home price indices.

We argue that using a home price index to impute home values with no adjustment for the distribution of individuals transacting may be inappropriate and any attempt to impute the value of every home in the market will be biased. This is especially prevalent due to the increase use of home price indices in this way in a wide range of topics. In recent years, these topics have included home-based equity borrowing, loan origination and lender behavior, and most prominently, mortgage default just to name a few.[8] For example, Bhutta et al. (2017) use home price indices to forecast individual homes for empirical tests on mortgage defaults. They suggest that for "ruthless" defaults to occur, homeowners would need a home price lower than what the option-pricing driven models would imply in order to walk away.[9] However, if one adjusted the home price index to take into account the distribution of individuals transacting, then it may be the case that home prices did not decrease as much as the non-adjusted home price indices would suggest. Bhutta et al. (2017), along with Gerardi et al. (2018), speculate that a potential reason why default rates are not as high as the ones theory predicts is due to the fact that homeowners may have a strong attachment to their home and may have certain optimistic beliefs about the future. This is a valid argument, but we argue that it boils down to an inappropriate use of home price indices to measure default behavior.

This paper is not the first to bring up issues with home price indices.[10] In fact, there is a large literature that purely focuses on index issues and ways to correct

[8]See Mian and Sufi (2011), Garmaise (2015), Adelino et al. (2016), Bhutta et al. (2017), and Favara and Giannetti (2017) for examples in which the authors use home price indices as a measure of home price growth in an area.

[9]Foster and Order (1984), Foster and Order (1985), Kau et al. (1992), and Kau et al. (1995) use option-pricing models to show that as home prices drop below the present value of the remaining mortgage payments, it is optimal for the homeowner to default on the mortgage. This is often referred to in the literature as the "ruthless" default strategy. Using various methods, others have shown that negative equity is an important variable in the default decision (Vandell (1978), Campbell and Dietrich (1983), Deng et al. (2000), Foote et al. (2008), Campbell and Cocco (2015)).

[10]See Cho (1996) for a survey of issues on home price indices.

these issues.[11] However, we believe we are the first to bring to light the issue that homeowner's decisions, especially during economic downturns, can have an effect on empirical research using these indices. In this paper, we use an equilibrium market, life-cycle model to provide a rational basis for this argument. The main reason for this type of model is that it produces a rich cross-sectional distribution of wealth through the combination of savings, housing wealth, and debt. These distributional features are key drivers in housing decisions. In addition, we believe that choices surrounding housing drastically change as the economy moves from a normal/expansion period into a recession/depression period. Thus, an additional feature in our model is the role of aggregate shocks to the economy. If all agents in the economy are hit with a large, negative shock to labor income as well as a tightening of future credit availability, then, as argued above, relatively *unconstrained* homeowners will be the only ones participating in the home selling market. Our model is most similar to Favilukis et al. (2017), Guren et al. (2019), and Kaplan et al. (2019), all of which look at an equilibrium model of the housing market with aggregate uncertainty. However, none of these studies explicitly analyze how the distribution of wealth affects the selling/default tradeoff. A large focus of these papers is on the credit condition changes during the 2000s.

The goal of this paper is to present a theoretical foundation to determine what drives the distribution of home sellers and to determine the price at which these individuals transact. The rest of the paper is as follows: Section 2.2 describes the theoretical model, Section 2.3 discusses the calibration, Section 2.4 discusses the results, and Section 2.5 includes the closing remarks.

[11] Most papers that focus on self-selection issues within asset pricing indices use the standard sample selection model introduced by Heckman (1979).

2.2 Model

2.2.1 Set Up

Time is discrete with yearly intervals indexed by t. The economy is made up of overlapping generations of heterogenous individuals of age $a = 1, 2, ..., A$ who make decisions annually. At age A, the individual dies with certainty. Individuals in our model make many decisions over their lifetime. Each period they choose whether to be a renter or a homeowner, how much to consume and save, and in the case of a homeowner, how much to borrow through a long-term mortgage. These decisions are influenced not only by the idiosyncratic state of the individual, which we will denote by $X_{i,a}$ where i indexes the agent, but also by the vector of aggregate variables in the economy, which we will denote by Ω_t. We assume that there are four aggregate states in the economy: recession period with tight credit, recession period with loose credit, normal economic period with tight credit, and normal economic period with loose credit. We assume that the transitions between a recession period and a normal economic period is governed by an independent Markov process. However, we assume that the transitions between the loose and tight credit state are conditional on the transitions between the recession and normal economic period.

Over the lifetime, an individual can choose either to be a renter or a homeowner. For simplicity, we assume that there is one type of owner-occupied unit and one type of rental unit. The total number of housing units is in fixed supply, and there is no construction. Owner-occupied homes are subject to a per-unit price $p_H(\Omega_t)$, which is endogenously determined and is a function of the aggregate state. Thus, the full price of an owner-occupied housing unit is $p_H(\Omega_t)H$. For simplicity, we assume that there is a risk-neutral rental company that can tear down and replace rental units at a per-unit cost of α.

In addition, owner-occupied housing units are subject to various costs. Homeowners must pay a per-period maintenance cost as well as a per-period property tax on the unit. Both of these are proportional to the price of the home and can be written as $\delta_H p_H(\Omega_t)H$ and $\tau_H p_H(\Omega_t)H$, respectively. Finally, when buying or selling, homeowners are subject to a large transaction cost, which is also proportional to the value of the home. This can be written as $\kappa_H p_H(\Omega_t)H$.

Regardless of the tenure choice, an individual obtains utility from both a numeraire good $C_{i,a}$ and housing $H_{i,a}$ and chooses the optimal bundle between these two goods to maximize lifetime utility. Since owning a home comes with more rights and flexibility than renting, we assume that homeowners receive extra utility from owning a home. That is, we assume that if an individual owns a home, then the housing service flow is $H_{i,a} = 1$, while if the individual rents, then the housing service flow is $H_{i,a} = 0$. The preferences for an individual homeowner can be written as:

$$E_0 \left[\sum_{a=1}^{A} \beta^{a-1} \left(\frac{C_{i,a}^{1-\gamma}}{1-\gamma} + \theta H_{i,a} \right) + \beta^A \mathbf{B}(W_{i,A+1}) \right] \tag{2.1}$$

where i denotes a certain individual, and β is the time discount factor. We assume that the utility function between the numeraire good and housing is a Cobb Douglas function with γ being the curvature parameter and θ being the relative importance of the numeraire good to housing. In addition, we assume that individuals in our model leave a bequest to future offspring. This bequest function $\mathbf{B}$ is a function of the terminal wealth of the agent, $W_{i,A+1}$. We assume that this bequest function takes the following form:

$$\mathbf{B}(W_{i,A+1}) = \psi \frac{W_{i,A+1}^{1-\gamma}}{1-\gamma} \tag{2.2}$$

where ψ is the strength of bequest motivation.

Each individual splits life into a working period and a retirement period. During the working period, individuals receive a stream of income that is subject to both aggregate and idiosyncratic shocks and follows a hump-shaped form over the life cycle.

During the retirement period, the individual receives an annuity equal to a fraction of the income received in the last working period. This annuity is not subject to any shocks. Each period before the retirement age K, homeowners receive a stochastic labor income of $Y_{i,a}$, which is a function of age a. We assume that the log of this before-tax, labor income is:

$$ln\left(Y_{i,a}\right) = \begin{cases} \mathbf{f}(a) + \epsilon_t^{agg} + \epsilon_{i,a}^{id} & a < K \\ \lambda_K Y_{i,K-1} & a \geq K \end{cases} \tag{2.3}$$

where $\mathbf{f}(a)$ is the deterministic income function at age a, ϵ_t^{agg} is the aggregate shock to income at time t, $\epsilon_{i,a}^{id}$ is the idiosyncratic shock to income, and λ_K is the retirement replacement ratio. The deterministic income function and the retirement replacement ratio are common to all agents in the model. The aggregate and idiosyncratic shocks both follow separate, independent Markov processes. Specifically, the idiosyncratic shocks to income, $\epsilon_{i,a}^{id}$, follow:

$$log(\epsilon_{i,a}^{id}) = \rho log(\epsilon_{i,a-1}^{id}) + z_{i,a} \qquad z_{i,a} \sim i.i.d(0, \sigma_z^2) \tag{2.4}$$

All agents in our model can save in a one-period bond at a real interest rate of $r_{b,t}$, which is the realized rate between time t and $t+1$. This rate is exogenous and depends on the aggregate state of the economy (i.e. whether the economy is in a recession period or a normal economic period). However, if agents need to borrow funds, they can obtain a long-term, fixed-rate mortgage but must put up a house as collateral. Additionally, the agent must pay a proportional origination fee, κ_M, on the mortgage. These mortgages are fully-amortizing loans with a mortgage interest rate $\bar{r}_m$. We follow standard convention in the literature by amortizing the mortgage over the life of the borrower.[12] This reduces the number of state variables in the model as we do not have to track an additional variable for time-to-maturity.[13] With the inclusion of this type of loan, agents in our model also have additional decisions to default and refinance.

At origination, the individual can borrow some amount $M_{i,a}$ with a mortgage interest rate exogenously set at the real interest rate $r_{b,t}$ plus some spread, ϕ_t. The mortgage spread is exogenous and depends on the aggregate state of the economy (i.e. whether the economy is in a tight credit period or a loose credit period). This mortgage interest rate, which is set at origination at time t, remains constant over the life of the loan and can be written as:

$$\bar{r}_m = r_{b,t} + \phi_t \tag{2.5}$$

where ϕ_t is the rate spread.

Since this is a fully-amortizing, fixed-rate mortgage, we can find the minimum payment that the individual needs to make in order to pay off the loan by the maturity date. This payment is held constant until the mortgage is fully paid off or until the homeowner defaults, whichever comes first. We can find this minimum, constant payment for an agent at age a by:

$$\mathbf{mp}(M) = M\frac{\bar{r}_m(1+\bar{r}_m)^{A-a}}{(1+\bar{r}_m)^{A-a}-1} \tag{2.6}$$

If the homeowner follows this payment schedule, then the mortgage's outstanding balance next period is: $M_{i,a+1} = M_{i,a}(1+\bar{r}_m) - \mathbf{mp}(M_{i,a})$.

In order to qualify for a mortgage, the homeowner must satisfy two constraints.

[12]See Campbell and Cocco (2003), Campbell and Cocco (2015), Guren et al. (2019), and Kaplan et al. (2019) to name a few

[13]This modeling decision does not alter aggregate or life-cycle moments in the model that we see in the data.

The first constraint is the loan-to-value constraint. That is, the amount borrowed must be less than or equal to some fraction of the home value. Formally this can be written as:

$$M_{i,a} \leq \lambda_{ltv,t} p_H(\Omega_t) H_{i,a} \tag{2.7}$$

where $\lambda_{ltv,t}$ is the maximum loan-to-value ratio, is exogenously determined, and depends on the aggregate state (i.e. whether the economy is in a tight credit period or a loose credit period).

In addition, the individual must satisfy the payment-to-income constraint. That is, the minimum payment must be less than or equal to some fraction of the individual's before-tax income at origination. Formally this can be written as:

$$\mathbf{mp}(M_{i,a}) \leq \lambda_{pti,t} Y_{i,a} \tag{2.8}$$

where $\lambda_{pti,t}$ is the maximum payment-to-income ratio, is exogenously determined, and depends on the aggregate state (i.e. whether the economy is in a tight credit period or a loose credit period).

On top of receiving additional utility for owning over renting, agents in our model also receive financial benefits for borrowing and paying down a mortgage. This is through tax deductibility of the mortgage interest payments. For this reason, we include a progressive tax system within the model similar to the tax system laid out by the U.S. tax code. We will denote this tax function as $\mathbf{T}(Y)$. If the individual does not make an interest payment on a mortgage, then the individual's full income is taxed.

As mentioned earlier, homeowners experience transaction costs when selling, which are a percentage of the current price of the home and makes housing highly illiquid relative to other financial goods. In addition, there are other reasons, outside the specifications in the model, why a homeowner may have to sell and move homes (i.e. family reason, etc.). Since our argument is concerned with the selling behavior of individuals, we include in our model some probability that the homeowner is forced to move. That is, we include an exogenous moving shock. If a homeowner is hit with this shock, then the homeowner can either sell the home or default and move to a rental unit. We denote this moving shock as ξ, which is independent across both time and individuals.

2.2.2 Homeowner's Optimization Problem

In this section we discuss the individual's decisions and their associated value functions. At the beginning of each period, the individual can either be a homeowner or renter. If the individual is a renter, then the individual must choose between renting again or buying a home. If the individual chooses to rent, then the agent will choose the amount to consume and the amount to save. If the individual chooses to purchase a home, then the agent will again choose the amount to consume and the amount to save but will also have the opportunity to borrow in a mortgage.

If the individual enters the period as a homeowner, then the agent has four choices: pay the mortgage and continue living in the same home, refinance the mortgage and continue living in the same home, default on the mortgage and be forced to rent for one period, or sell the home and either buy a new home or rent a home in the following period. If the individual chooses to pay, refinance, or sell the home, then the individual must pay the proportional maintenance and tax cost on the home. In the case of selling the home, the individual must also pay the transaction cost. Once the homeowner sells the home, they then solve the problem as if they had just entered the period as a renter but with the selling proceeds attached to their current savings. However, if the individual defaults, then no owner-occupied housing costs are paid, the individual is forced to rent for one period, and will receive a utility loss of Ξ for defaulting.

We can now write the individual's recursive optimization problem. Moving forward, we drop the i subscript for ease of notation unless it is needed. At the beginning of each aggregate period t, a new generation is born. Aggregate and idiosyncratic shocks are realized such that the agent knows the real interest rate, the current mortgage rate, and the level of income available at the beginning of the period. Agents then make decisions.

Let $\mathbf{V}_a\left(X_a^j, \Omega_t\right)$ be the value function of the individual. X_a^j is the vector of state variables for an aged a individual where $j \in \{o, r\}$ represents whether the individual is entering the period as a homeowner or renter. This is needed because renters cannot enter into the period with a mortgage like homeowners. The vector of individual state variables can be written as:

$$\begin{aligned} X_a^o &= [B_{a-1}, M_{a-1}, \bar{r}_m, Y_a, \xi] \\ X_a^r &= [B_{a-1}, Y_a] \end{aligned} \tag{2.9}$$

From Equation 2.9, we can see that both types of agents at age a will enter the period with previous savings, B, and current income, Y. Agents entering as homeowners also come in with an outstanding loan balance of M, the interest rate on that mortgage from last period, $\bar{r}_m$, and a moving shock ξ.

Since there is aggregate risk in the model that affects both income and credit, we must include the aggregate state of the economy as a state variable, which we will denote as Θ_t. Furthermore, agents in the model interact with each other to buy and sell homes. To do this, they need to accurately forecast all prices within the model. However, this is a function of all individuals' decisions in the economy, which means that the distribution of individual's must also be included as a state variable. We denote this distribution as $\mu\left(X_a^j\right)$. Both of these are aggregate variables in the economy and as such we can further define our vector of aggregate variables as: $\Omega_t = [\Theta_t, \mu\left(X_a^j\right)]$.

Since we have two different agents in the model with different state variables, we can separate their recursive optimization problem. We will first start with the agent that enters the period as a renter at all ages prior to the terminal age A. As mentioned, this individual can either rent or buy and will choose the option that grants the higher utility value. Let $\tilde{\mathbf{V}}_a^{rent}\left(X_a^r, \Omega_t\right)$ and $\tilde{\mathbf{V}}_a^{buy}\left(X_a^r, \Omega_t\right)$ be the intermediate value functions for a renter who chooses to rent or buy respectively. We can then define the ultimate value function for an agent entering the period as a renter as:

$$\mathbf{V}_a^R\left(X_a^r, \Omega_t\right) = max\left[\tilde{\mathbf{V}}_a^{rent}\left(X_a^r, \Omega_t\right), \tilde{\mathbf{V}}_a^{buy}\left(X_a^r, \Omega_t\right)\right] \tag{2.10}$$

If an individual chooses to rent, then the individual will solve:

$$\begin{aligned} \tilde{\mathbf{V}}_a^{rent}\left(X_a^r, \Omega_t\right) = \max_{C_a, B_a} \quad & \mathbf{U}(C_a, H_a) + \beta E_t\left[\mathbf{V}_{a+1}^R\left(X_{a+1}^r, \Omega_{t+1}\right)\right] \\ \text{s.t.} \quad & B_{a-1} + Y_a - \mathbf{T}(Y_a) = C_a + \frac{B_a}{r_{b,t}} + \alpha \\ & B_a \geq 0 \end{aligned} \tag{2.11}$$

If an individual chooses to buy, then the individual will solve:

$$\begin{aligned} \tilde{\mathbf{V}}_a^{buy}\left(X_a^r, \Omega_t\right) = \max_{C_a, B_a, M_a} \quad & \mathbf{U}(C_a, H_a) + \beta E_t\left[\mathbf{V}_{a+1}^O\left(X_{a+1}^o, \Omega_{t+1}\right)\right] \\ \text{s.t.} \quad & B_{a-1} + Y_a - \mathbf{T}(Y_a) + M_a(1-\kappa_M) = C_a + \frac{B_a}{r_{b,t}} + p_H(\Omega_t)H_a \\ & B_a \geq 0 \\ & \bar{r}_m = r_{b,t} + \phi_t \\ & M_a \leq \lambda_{ltv,t} p_H(\Omega_t) H_a \\ & \mathbf{mp}(M_a) \leq \lambda_{pti,t} Y_a \end{aligned} \tag{2.12}$$

Note that the expected future value of an individual choosing to become a buyer is the expected future value function of a homeowner. This value function is made up of the optimal choice of a homeowner, which includes paying, refinancing, selling, and defaulting. This can be formally written as:

$$\mathbf{V}_a^O\left(X_a^o, \Omega_t\right) = max \begin{cases} \tilde{\mathbf{V}}_a^{pay}\left(X_a^o, \Omega_t\right) \\ \tilde{\mathbf{V}}_a^{refi}\left(X_a^o, \Omega_t\right) \\ \tilde{\mathbf{V}}_a^{sell}\left(X_a^o, \Omega_t\right) \\ \tilde{\mathbf{V}}_a^{default}\left(X_a^o, \Omega_t\right) \end{cases} \tag{2.13}$$

If an individual chooses to pay, then the individual will solve:

$$\begin{aligned} \tilde{\mathbf{V}}_a^{pay}\left(X_a^o, \Omega_t\right) = \max_{C_a, B_a, MP_a} \quad & \mathbf{U}(C_a, H_a) + \beta E_t\left[\mathbf{V}_{a+1}^O\left(X_{a+1}^o, \Omega_{t+1}\right)\right] \\ \text{s.t.} \quad & B_{a-1} + Y_a - \mathbf{T}(Y_a - \bar{r}_m M_{a-1}) \\ & \qquad = C_a + \frac{B_a}{r_{b,t}} + (\delta_H + \tau_H) p_H(\Omega_t) H_{a-1} + MP_a \\ & B_a \geq 0 \\ & \bar{r}_m = \bar{r}_m \\ & MP_a \geq \mathbf{mp}(M_{a-1}) \end{aligned} \tag{2.14}$$

If an individual chooses to refinance, then the individual will solve:

$$
\begin{aligned}
\tilde{\mathbf{V}}_a^{refi}\left(X_a^o, \Omega_t\right) = \max_{C_a, B_a, M_a} \quad & \mathbf{U}(C_a, H_a) + \beta E_t \left[\mathbf{V}_{a+1}^O\left(X_{a+1}^o, \Omega_{t+1}\right)\right] \\
& B_{a-1} + Y_a - \mathbf{T}(Y_a - \bar{r}_m^{old} M_{a-1}) + M_a(1 - \kappa_M) \\
\text{s.t.} \quad & \qquad = C_a + \frac{B_a}{r_{b,t}} + (\delta_H + \tau_H) p_H(\Omega_t) H_{a-1} + \\
& \qquad M_{a-1}(1 + \bar{r}_m^{old}) \\
& B_a \geq 0 \\
& \bar{r}_m^{new} = r_{b,t} + \phi_t \\
& M_a \leq \lambda_{ltv,t} p_H(\Omega_t) H_a \\
& \mathbf{mp}(M_a) \leq \lambda_{pti,t} Y_a
\end{aligned}
\tag{2.15}
$$

Note that the tax deduction is based on the mortgage interest rate that the individual came in with at the beginning of the period. However, since the agent is refinancing, the mortgage interest rate on the new mortgage may be different.

If an individual chooses to default, then the individual will solve:

$$
\begin{aligned}
\tilde{\mathbf{V}}_a^{default}\left(X_a^o, \Omega_t\right) = \max_{C_a, B_a} \quad & \mathbf{U}(C_a, H_a) + \beta E_t \left[\mathbf{V}_{a+1}^R\left(X_{a+1}^r, \Omega_{t+1}\right)\right] - \Xi \\
\text{s.t.} \quad & B_{a-1} + Y_a - \mathbf{T}(Y_a) = C_a + \frac{B_a}{r_{b,t}} + \alpha \\
& B_a \geq 0
\end{aligned}
\tag{2.16}
$$

Note that the value function of a defaulter is nearly identical to the value function of a renter, but the defaulter receives a utility loss of Ξ for defaulting.

Finally, if an individual chooses to sell, then the individual will first receive the selling proceeds given by $(1-\delta_H-\tau_H-\kappa_H)p_H(\Omega_t)H_{a-1}-(1+\bar{r}_m)M_{a-1}$ and then will solve Equation 2.10. That is, the home seller will both sell the old home and choose a new home in the same period.

Since the agent's lifetime is finite, the agent must make a terminal decision at age A. Depending on the state variables of the agent, the agent at terminal age A will solve either Equation 2.10 or Equation 2.13. However, within the intermediate functions that govern these two equations, the future expected value is replaced with the bequest function specified in Equation 2.2. That is, the agent will earn value over today's decisions and the discounted value of the bequest action over terminal wealth. Since there are both homeowners and renters within the model, the terminal wealth value will vary. If the individual is a homeowner in the last period, then the terminal wealth will be a combination of savings and the expected future price of the home sale net of all costs. On the other hand, if the individual is a renter in the last period, then the terminal wealth of this individual is only made up of savings.

2.2.3 Equilibrium

A recursive, competitive equilibrium consists of the intermediate value functions $\tilde{\mathbf{V}}_a^{rent}(X_a^r, \Omega_t)$, $\tilde{\mathbf{V}}_a^{buy}(X_a^r, \Omega_t)$, $\tilde{\mathbf{V}}_a^{pay}(X_a^o, \Omega_t)$, $\tilde{\mathbf{V}}_a^{refi}(X_a^o, \Omega_t)$, $\tilde{\mathbf{V}}_a^{sell}(X_a^o, \Omega_t)$, and $\tilde{\mathbf{V}}_a^{default}(X_a^o, \Omega_t)$ and their respective decisions over consumption, savings, housing, mortgage borrowed, and mortgage payments, a per-unit housing price function $p_H(\Omega_t)$, and a law of motion for the aggregate vector Ω_t. Agents optimize their decisions by solving Equations 2.10-2.16. Due to the fixed supply of housing, the housing market clears such that the quantity demanded by home buyers equals the quantity supplied by home sellers and defaulters. If we let $I^{buy}(X_a^{r,o}, \Omega_t)$, $I^{default}(X_a^o, \Omega_t)$, and $I^{sell}(X_a^o, \Omega_t)$ be indicator functions for buying, defaulting, and selling respectively and let $H_a(X_a^{r,o}, \Omega_t)$ and $H_{a-1}(X_a^o, \Omega_t)$ be the housing quantity demanded and supplied respectively, then we can formally write the market clearing condition as:

$$\int I^{buy}\left(X_a^{r,o},\Omega_t\right)H_a\left(X_a^{r,o},\Omega_t\right)d\mu = \int I^{default}\left(X_a^{o},\Omega_t\right)H_{a-1}\left(X_a^{o},\Omega_t\right)d\mu + \int I^{sell}\left(X_a^{o},\Omega_t\right)H_{a-1}\left(X_a^{o},\Omega_t\right)d\mu \tag{2.17}$$

where the left hand side is the quantity demanded for housing and the right hand side is the quantity supplied for housing, which is made up of defaulters and homeowners who sell.

2.2.4 Numerical Computational Method

We now describe how we numerically solve the model and find equilibrium prices. Due to aggregate uncertainty, individuals must track the distribution of individuals in the economy. However, this distribution is an infinite-dimensional object and is not numerically tractable. For this reason, we utilize the insights from Favilukis et al. (2017), Kaplan et al. (2019), and Guren et al. (2019), which builds on the algorithm from Krusell and Smith Jr. (1998) applied to the housing market. This algorithm breaks down the infinite-dimensional object into a few key moments that can be used in order for agents to make accurate forecasts about the future.

Each period, all agents in the model only need to know the per-unit price of owner-occupied housing and need to be able to forecast this price in order to make informed decisions. We conjecture that the per-unit housing price law of motion is an AR(1) model, which is a function of the current price, the current aggregate state, and next period's price conditional on next period's aggregate state. Formally we can write the per-unit price law of motion as:

$$log(p_H(\Omega_{t+1})) = \beta_1(\Omega_t,\Omega_{t+1}) + \beta_2(\Omega_t,\Omega_{t+1})log(p_H(\Omega_t)) \tag{2.18}$$

Using market clearing prices obtained from Equation 2.17 at each time t, we can estimate Equation 2.18 to obtain β_1 and β_2. Since we have four aggregate states, we need to estimate Equation 2.18 for 16 different combinations of the current and future aggregate states. It is worth noting that since agents in our model are only using the current market clearing price to forecast the future price, agents in our model are boundedly (as opposed to fully) rational. However, using another law of motion specification does not significantly alter our results.

With Equation 2.18 at hand, we can now discuss our computational method. This model is numerically intensive, but standard numerical methods can be used. We discretize the grids for B, M, and p_H. To obtain next period's mortgage amount, we use the amortization schedule and the chosen payment amount to linearly interpolate the mortgage value. Finally, given the assumptions laid out above with respect to the aggregate states, we use two values for the one-period saving's rate ($r_{b,t}$), the mortgage spread (ϕ_t), the loan-to-value constraint ($\lambda_{ltv,t}$), the payment-to-income constraint ($\lambda_{pti,t}$), and the aggregate shocks to income (ϵ_t^{agg}).

2.2.5 Solution Procedure

First, we specify an initial vector of coefficients for the law of motion for the per-unit price of housing, which we denote as $\boldsymbol{\beta}^n = \boldsymbol{\beta}^0$, and solve the agent's problem via backward induction and value function iteration for every combination of the state variables starting at age A. With each combination of the state variables, we rearrange the budget constraints so that we only need to optimize over one variable (the saving's amount B). We then plug this optimal saving's amount back into the budget constraint to get the optimal consumption amount of the numeraire good. Any combination of the state variables that violate the budget constraint was given a very large and negative utility value. Since the value of the agent is the current utility plus the discounted expected utility one period in the future, in order to compute the expected utility for points not on the grid, we use linear interpolation. This process is repeated backwards all the way to age $a = 1$.

Once the value functions and their respective choice functions are solved, we simulate a large panel of individuals. That is, for each time period over a long period of time, we simulate 10,000+ agents whose decisions are based off the value and choice functions determined in the previous step. In order to determine the market clearing price in each time period, we must find the aggregate excess demand value for each point in the discretized price grid using Equation 2.17. This gives us a vector of excess demand values at each price point in which we then linearly interpolate to find the market clearing price. We then update the distribution of agents in our economy using this market clearing price. We repeat this process for each period in the simulation, which results in a time series of equilibrium home prices.

Before we calculate the simulated moments and the vector of coefficients for the home price law of motion, we drop the first 100 periods so that the initial distribution does not affect these calculations. We estimate Equation 2.18 on the generated time series of equilibrium prices to get the new coefficients for the home price law of motion, $\tilde{\boldsymbol{\beta}}^{n+1}$. Finally, we check to see if the new coefficients are similar to the previous coefficients. If they are not, we update the coefficients as:

$$\boldsymbol{\beta}^{n+1} = \eta\tilde{\boldsymbol{\beta}}^{n+1} + (1-\eta)\boldsymbol{\beta}^{n} \tag{2.19}$$

where $0 < \eta < 1$. We then repeat the procedure stated above from $n = 0, 1, ...$ until the coefficients between each iteration are significantly similar and the R^2 on the regressions are high.

2.3 Calibration

To gain insights from the model, we calibrate the model to match key moments in the data that occurred before 2000 both at the aggregate level as well as at the idiosyncratic level. We do this through external and internal parameterization. To match key life-cycle moments of individuals, data from the Survey of Consumer Finances in 1998 was used as this was the last year of this survey before the 2000s. All parameters can be found in Table 2.2.

2.3.1 External

2.3.1.1 Aggregate Parameters

As mentioned above, we examine an economy that has four aggregate states: a normal state that can have either loose or tight credit and a recession state that can have either loose or tight credit. When the economy enters into a recession, aggregate income and the real rate of interest decrease. If the economy is also in a period of tight credit, the maximum loan-to-value and maximum payment-to-income ratios decrease, and the mortgage spreads over the real rate of interest increase.

To govern the time spent in the recession and normal economic state as well as the probability of transitioning into the other state, we use the recession dates published by NBER between 1857-1999. These recession dates imply that the U.S. economy spends about an average of 1.48 years in a recession before switching to a normal economic state and about an average of 3.16 years in a normal economic state before switching to a recession. Furthermore, to pin down the transition between the two credit states, we make the assumption that credit changes are rare and on average, only happen once every generation. In addition, we assume that credit will never tighten when the economy is transitioning into a normal economic period nor will it ever loosen when the economy is transitioning into a recession period. Letting state 1 be a recession with tight credit, state 2 be a normal period with tight credit, state 3 by a recession with loose credit, and state 4 be a normal period with loose credit, the assumptions above imply that the aggregate Markov probability matrix is:

$$P^{agg} = \begin{bmatrix} 0.3243 & 0.6486 & 0.0 & 0.0270 \\ 0.3165 & 0.6562 & 0.0 & 0.0273 \\ 0.0130 & 0.0 & 0.3114 & 0.6757 \\ 0.01 & 0.0 & 0.3065 & 0.6835 \end{bmatrix} \tag{2.20}$$

This probability matrix implies that if the economy is currently in a recession with tight credit, then it will stay in this state with a probability of 32.43%. Whereas, if the economy is currently in a normal economic period with loose credit, then it will stay in this state with a probability of 68.35%. This probability matrix also implies that in the long-run, the economy will spend 49.40% of its time in a normal economic state with loose credit and 9.91% of its time in a recession state with tight credit (i.e. the worst case scenario is rare). Finally, this transition matrix implies that normal economic states will be on average 2.14 times longer than recessions, which is consistent with the NBER recession dates.

In order to match the decrease in income and the tightening of credit, we exogenously set ϵ_t^{agg}, $r_{b,t}$, ϕ_t, $\lambda_{ltv,t}$, and $\lambda_{pti,t}$. We set ϵ_t^{agg} so that aggregate income decreases by 4% when the economy goes from a normal economic state to a recession and so that the median income in the stochastic steady state of the model is $42,000. This is consistent with the drop in real gross domestic income from 1980-1999

We assume that when the economy enters into a period of tight credit, both the loan-to-value and payment-to-income ratios decrease. That is, during times of tight credit, the maximum loan-to-value possible at origination is 80%, while during a periods of loose credit, the maximum loan-to-value possible at origination increases to 95%. Similarly, we assume that during times of tight credit, the maximum payment-to-income possible at origination is 36%, while during periods of loose credit, the maximum payment-to-income possible at origination increases to 45%. These credit constraints are largely consistent with the conforming loan limits within the U.S. We also set the savings rate and mortgage spread to match their counterparts in the data. Specifically, we use the average deflated 1-year Treasury bond yield based on the NBER recession dates from 1980-1999 as our one-period saving's rate. This gives us a 2.79% and 3.34% one-period saving's rate in a recession and normal economic state respectively. Similarly, we find the average mortgage spread by taking the difference between the 30-year fixed rate mortgage and the 10-year Treasury bond yield based on the NBER recession dates from 1980-1999. This gives us a 1.92% and 1.59% mortgage spread in a tight and loose credit state respectively.

2.3.1.2 Individual Parameters

We assume that homeowners live for 56 years starting at age 25 and die with certainty at age 80. The homeowner will retire at age 65, and the age-dependent function in the income process as well as the retirement replacement ratio are taken from Cocco et al. (2005), where we use the coefficients from the third order polynomial that fits the life cycle labor income of a college graduate. The income shock process in Equation 2.4 is set with an annual persistence of 0.959 (ρ) and a standard deviation of 0.141 (σ_z), which comes from Hubbard et al. (1995) who estimate the income process using Panel Study of Income Dynamics (PSID) data. Using this setup, the income process will have a hump shape over the life of the homeowner.

We make a few simplifying assumptions about certain costs and preference parameters that are largely consistent with the previous literature. We assume that the curvature parameter γ is 3. We assume that the transaction costs, κ_H, imposed on selling the home is 6%, and the per-period maintenance cost, δ_H, is 2.5% of the value of the home. We assume that the per-period property tax, τ_H is 1%, and the origination fee on a mortgage, κ_M, is 3%. Both are consistent with the aggregate average seen in the U.S.

2.3.1.3 Progressive Tax

As mentioned earlier, individuals may choose to buy a home not only because it is a dual consumption and investment asset but also because it provides financial benefits through tax write-offs on mortgage interest payments. In order to include this benefit, we include an income tax system laid out by the U.S. tax code in 1998. Specifically, this is a progressive tax system in which pre-specified income brackets defined by the federal government are taxed at increasing marginal rates. If the individual pays interest on a mortgage, then the individual may deduct that amount from the taxable income. For example, if the before-tax income of the individual is Y and the individual pays the interest on a mortgage M with interest rate $\bar{r}_m$, then the taxable-income becomes $Y - \bar{r}_m M$. Table 2.1 shows the taxable income brackets with their respective marginal tax rates. These are used for the tax function $\mathbf{T}(Y)$.

2.3.2 Internal

To round out the remaining parameters in the model, we internally calibrate a few parameters in order to minimize the distance between the stochastic moments of the model and their real world data counterparts. We internally calibrate θ, β, ψ, α, ξ, and Ξ. We calibrate θ by matching the average home price-to-income in the data. According to Zillow, the long-run average of this value across the U.S. was about 2.79. β was calibrated so that the percentage of homeowners having a mortgage is roughly 70%, which is in line with the 2000 Census.

In order to have a meaningful amount of savings within the model, we calibrate ψ so that the ratio of median net-worth at age 75 to the median net-worth at age 50 is roughly 1.35. This value is taken from the SCF in year 1998. Since we only have liquid savings, housing, and mortgages in our model, we can not use the SCF pre-defined *networth* variable. Instead, we calculate assets as only liquid assets and housing and calculate debt as only mortgages.[14] α is calibrated to match the per-period average

[14]Specifically, we used the SCF pre-defined variables *liq*, *cds*, *bond*, *retqliq*, *savbnd*, *houses*, and

rent-to-price ratio in the U.S. Using Census data along with indices for home prices and rents, Davis et al. (2008) find that the rent-to-price ratio between 1960-1995 was between 5%-5.5%. However, by 2000, the rent-to-price ratio continuously dropped to 4.6%. For this reason, we choose the midpoint between 4.6% and 5% (i.e. 4.8%).

As mentioned earlier, homeowners may have other reasons that force them to move such that selling is no longer purely endogenous. One could in theory model family dynamics or job changes such that selling becomes purely endogenous, but for simplicity reasons we choose an exogenous approach. With this said, we calibrate ξ so that the per-period selling rate equals 8.8%, which is the average per-period moving rate from 1987-1999. This means that the combination of endogenous and exogenous sales must equal roughly 8.8%.

Finally, we calibrate Ξ by matching the per-period foreclosure rate in the U.S. Since we assume that defaults are immediate in our model and lead to a forfeiture of the home, we choose to match the foreclosure rate rather than the default rate. The quarterly report on household debt and credit from the New York Federal Reserve issues data on the number of consumers with new foreclosures. Unfortunately, this data only goes back to 2003. Nonetheless, we note that in the data, pre-2007 and post-2009 foreclosures were very similar. Given that the U.S. economy had never experienced a housing crisis before the Great Recession, it is likely that the 2003-2006 and 2010-2018 foreclosure behavior is more in line with the long-run foreclosure rate averages. For this reason, we calibrate the foreclosure rate using these two time periods. To find the foreclosure rate, we divide the number of foreclosures by the estimated number of total owner-occupied housing units in the U.S. in 2003-2006 and 2010-2018. This gives us an average per-period foreclosure rate of about 1%.

---

mrthel

2.4 Model Results

2.4.1 Model Fit

Table 2.3 shows the model fit to its empirical counterpart. Overall, the model seems to fit the data quite well. The homeownership rate, the percentage of homeowners with a mortgage, the sell rate, and the default rate are all very similar to the data. However, the price-to-income ratio, rent-to-price ratio, and the median wealth ratio are significantly different than the data.

Figure 2.1 shows a few lifecycle moments for the agents in the model. Homeownership fits well with the data and almost perfectly at older ages. However, the model undershoots homeownership in the early ages. This is the case because when agents first enter the model at age 25, they do so as renters with little wealth on average. As they progress through life, these individuals slowly build up their savings so that they can put a down payment on a house. Furthermore, the economy spends nearly 70% of the time in a normal/expansion phase. As will be clearer later, during this time, equilibrium home prices are relatively high. These high prices push lower wealth individuals, such as the young, out of the market and into the rental market. Following this logic, we can see that at younger ages, networth in the model undershoots its empirical counterpart. As individuals age, their networth explodes such that it greatly overshoots its empirical counterpart. This significant difference is again due to the fact that the economy experiences high equilibrium prices roughly 70% of the time. Since housing wealth is included in networth, the majority of individuals have extremely large networth in the model as opposed to the data.

Finally, Figure 2.2 shows the equilibrium prices when simulated through a boom-bust cycle. The first 8 periods have aggregate shocks associated with a normal/loose credit phase, the next 3 periods have aggregate shocks associated with a recession/tight credit phase, and the last 4 periods have shocks again associated with a normal/loose credit phase. All prices are relative to time 0, which is the last period before the bust. The model seems to do a decent job replicating the bust portion of the cycle. Prices are roughly constant throughout the initial normal/loose credit phase, then drop by about 42.64% when the recession shocks hit the economy at time 1. This is roughly in line with the peak-to-trough drop in home price indices. After the recession period is over, equilibrium prices revert back to pre-recession levels. This is again similar to the information that home price indices convey. That is, by the beginning of 2017, the Case-Shiller National Home Price Index has passed its pre-recession maximum level. Thus, overall, the model seems to do a decent job matching the data.

2.4.2 Does the Distribution of Sellers Change in a Recession?

The first question that this paper addresses is: Is the distribution of home sellers fundamentally different than the distribution of all homeowners? As pointed out earlier, relatively *unconstrained* homeowners should have an easier time trying to sell their home because they do not need to ask for an extremely high price. Homeowners can still ask for a high price, but this will decrease the probability that they will successfully sell the home. Furthermore, as the economy enters into a recession, potential buyers will be less willing and able to enter into the housing market. Thus, only sellers who can afford to sell at low prices will do so.

Figure 2.3 shows the characteristics of homeowners in a normal/loose credit state and in a recession/tight credit state. The figure further breaks homeowners down by the choices they make. Specifically, it breaks it down on whether a homeowner defaults, sells, or stays in the same home. We will refer to these individuals as defaulters, sellers, and stayers respectively. It is clear that during normal periods, sellers and stayers are quite similar in their characteristics. Both have similar loan-to-income and loan-to-value ratios. Sellers have a slightly higher savings-to-income ratio relative to stayers, but the difference is small and insignificant. As for defaulters, it is clear that they are much different than sellers and stayers. Defaulters are the most debt burdened, low income/savings homeowners.

When the economy is in a recession, the distribution makes a visible change. Sellers during recessions seem to be relatively *unconstrained* to stayers. Sellers on average have higher savings-to-income ratio than stayers and slightly lower loan-to-income ratios than stayers, but this difference is quite small. Nonetheless, the difference in income and savings between sellers and stayers is quite large, which makes it easier for sellers to sell and move to another home.

Finally, since we solve for equilibrium prices in the model, this also gives us the ability to find the reservation price of each homeowner. This price is the minimum prices that homeowners would be willing to accept in order to be indifferent between their actual optimal choice given the equilibrium price and selling. For example, given the equilibrium price, if a homeowner decides to stay in the home by paying the mortgage, then the reservation price is found by equating the value from this paying decision and the value from selling. By assumption, we assume that if the homeowner's optimal choice is to sell given the equilibrium price, then the equilibrium price is the reservation price for sellers. We then create the "reservation ratio," which is the reservation price divided by the equilibrium price in a given period. Again, by assumption, the seller's reservation price is equal to one.

From Figure 2.3, we can see that on average, during normal periods, all homeowners have similar reservation ratios. However, when the economy shifts to a recession, the reservation ratios among the different homeowners shifts substantially. To no surprise, defaulters have the highest reservation ratio among all homeowners. In order for defaulters to have the ability to sell their home, the equilibrium home price needs to be roughly 1.77 times greater. This is the case because defaulters on average have the highest loan amount and lowest income and savings. Since they are extremely cost burden, they need a higher price in order to financially sell their home. The reservation ratio of stayers also increases when going from a normal economic state to a recession period. Specifically, during a normal economic state, stayers have an average reservation price that is 1.32 times greater than the market price. However, when the economy moves to a recession state, this average reservation ratio increases to 1.69, which means that in order for stayers to be incentivized to sell their home, the equilibrium price must be 1.69 times greater. Note that even though these reservation ratios are largely driven by the financial constraints of the homeowner, this value gives us an indication of the internal price at which the homeowner would want to sell the home because it takes in both the consumption and financial aspects of housing when constructing this value.

Thus, results in Figure 2.3 suggest that, on average, the distribution of sellers and stayers is roughly similar when the the economy is a normal state. However, when the economy shifts to a recession, the distribution drastically shifts such that sellers are those that are relatively *unconstrained.* Since sellers are relatively *unconstrained*, they do not demand as high of a price than the rest of the homeowners, which means that they are willing and able to sell at a lower price point. This pushes equilibrium home prices down, which further separates the distribution of sellers from the distribution of stayers and defaulters.

2.4.3 Does the Distribution of Homeowners Affect Prices?

The second question this paper addresses is: Does this distributional difference lead to a large separation between observed transaction prices and unobserved home values? The previous section suggests that during normal economic states, the distribution between sellers and stayers is similar. In some dimensions, even the distribution of defaulters is similar to the rest of the homeowners. However, if the economy is in a recession, the distribution of these individuals drastically shifts, such that sellers become visibly less constrained relative to stayers and defaulters. Since sellers are less financially constrained, they have more flexibility in the price at which they are willing and able to sell. If there are enough sellers who are financially unconstrained and are willing to sell at lower prices, then this subset of individuals has the ability to decrease equilibrium prices.

However, this low price may not reflect the price at which stayers and defaulters would transact. As was seen in Figure 2.3, stayers and defaulters during recession states have large and significant reservation ratios (i.e. they need to sell at a price above the observed equilibrium price). This large reservation ratio suggests that if these individuals were able to get the price at which they would want to sell, then equilibrium prices should be higher than what was actually observed. In addition, these large reservation values suggest that the remaining distribution of homeowners are willing to sell at a price higher than the current sellers. Note that during normal economic states, the reservation ratio for all homeowners is close to one, while during recessions, the reservation ratio for stayers and defaulters is greater than one. If the entire homeowner distribution were to place their home on the market, then transaction prices should be higher, but only during recession periods. However, since buying and selling happen in the same period and owning is more beneficial than renting, stayers may want to sell at a high price, but they also want to buy at a low price. Furthermore, since the market must clear each period and there is a fixed amount of supply, it is unclear what the price may look like when the entire distribution transacts.

Figure 2.4 plots this scenario. That is, we compare the equilibrium prices from the model to the average reservation price of each stayer and mover. Similar to Figure 2.2, the economy enters into a recession at time one and stays in a recession for three periods. We see that during normal economic states, the equilibrium price and the average reservation price of all homeowners is quite similar. However, once the recession occurs at year 1, we can see large differences between the equilibrium price and the reservation prices of the homeowners. Specifically, we can see that stayers want a significantly higher price than the observed transaction price and as a result, choose not to transact. Movers on the other hand have reservation prices closer to the observed equilibrium prices. The biggest difference among the mover group occurs at year 1 when there are a high amount of defaulters. As mentioned above, defaulters have a higher reservation price because they need to pay off their high mortgage. This drives up the average reservation price of movers at year 1, but the remaining years of the recession (year 2 and year 3) show average reservation prices of movers very similar to the observed transaction price. This suggest that the distribution of homeowners does seems to matter. If we take into consideration the price at which homeowners would like to transact, then prices would be slightly higher during recessions than what is actually observed.

2.5 Conclusion

This paper explores the dynamics of housing decisions and how these decisions affect observed transaction prices. Homeowners need to make many decisions and have to take into consideration not only their own circumstances, but the circumstances of the surrounding area. From the model laid out in the paper, we show that the distribution of sellers is different than the remaining distribution of homeowners. During normal economic states, the distribution between sellers and stayers is similar, but when the economy transitions into a recession state, the distribution between sellers and stayers shifts. That is, during recessions, sellers are relatively *unconstrained* when compared with homeowners who default or those that stay in the same home. Since sellers are relatively *unconstrained*, they do not demand as a high of a price when selling. This pushes out other homeowners who want to sell but cannot because they need to sell at a higher price. If the entire distribution of homeowners put their home on the market to sell, which takes into account the reservation price of the underlying homeowners, then transaction prices should be higher. Thus, on average, the observed transaction price during recessions may not be an accurate portrayal of the unobserved home value, which is evident from the results laid out in this paper.

REFERENCES

Adelino, M., Schoar, A., and Severino, F. (2016). Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class. *The Review of Financial Studies*, 29(7):1635–1670.

Bailey, M. J., Muth, R. F., and Nourse, H. O. (1963). A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association*, 58(304):933–942.

Bhutta, N., Dokko, J., and Shan, H. (2017). Consumer Ruthlessness and Mortgage Default During the 2007 to 2009 Housing Bust. *Journal of Finance*, 72(6):2433–2466.

Campbell, J. Y. and Cocco, J. F. (2003). Household Risk Management and Optimal Mortgage Choice. *Quarterly Journal of Economics*, 118(4):1449–1494.

Campbell, J. Y. and Cocco, J. F. (2015). A Model of Mortgage Default. *Journal of Finance*, 70(4):1495–1554.

Campbell, T. S. and Dietrich, J. K. (1983). The Determinants of Default on Insured Conventional Residential Mortgage Loans. *Journal of Finance*, 38(5):1569–1581.

Capozza, D. R., Hendershott, P. H., and Mack, C. (2004). An Anatomy of Price Dynamics in Illiquid Markets: Analysis and Evidence from Local Housing Markets. *Real Estate Economics*, 32(1):1–32.

Case, B. and Quigley, J. M. (1991). The Dynamics of Real Estate Prices. *Review of Economics and Statistics*, 73(1):50–58.

Case, K. E. and Shiller, R. J. (1987). Prices of Single-Family Homes Since 1970: New Indexes for Four Cities. *New England Economic Review*, (851):45–56.

Case, K. E. and Shiller, R. J. (1989). The Efficiency of the Market for Single-Family Homes. *American Economic Review*, 79(1):125–137.

Chan, S. (2001). Spatial Lock-In: Do Falling House Prices Constrain Residential Mobility? *Journal of Urban Economics*, 49(3):567–586.

Cho, M. (1996). House Price Dynamics: A Survey of Theoretical and Empirical Issues. *Journal of Housing Research*, 7(2):145–172.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and Portfolio Choice Over the Life Cycle. *The Review of Financial Studies*, 18(2):491–533.

Davis, M. A., Lehnert, A., and Martin, R. F. (2008). The Rent-Price Ratio for the Aggregate Stock of Owner-Occupied Housing. *Review of Income and Wealth*, 54(2):279–284.

Deng, Y., Quigley, J. M., and Order, R. V. (2000). Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options. *Econometrica*, 68(2):275–307.

Engelhardt, G. V. (2003). Nominal Loss Aversion, Housing Equity Constraints, and Household Mobility: Evidence from the United States. *Journal of Urban Economics*, 53(1):171–195.

Favara, G. and Giannetti, M. (2017). Forced Asset Sales and the Concentration of Outstanding Debt: Evidence from the Mortgage Market. *Journal of Finance*, 72(3):1081–1118.

Favilukis, J., Ludvigson, S. C., and Nieuwerburgh, S. V. (2017). The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium. *Journal of Political Economy*, 125(1):140–223.

Ferreira, F., Gyourko, J., and Tracy, J. (2010). Housing Busts and Household Mobility. *Journal of Urban Economics*, 68(1):34–45.

Foote, C., Gerardi, K., and Willen, P. S. (2008). Negative Equity and Foreclosure: Theory and Evidence. *Journal of Urban Economics*, 64(2):234–245.

Foster, C. and Order, R. V. (1984). An Option-Based Model of Mortgage Default. *Housing Finance Review*, 3(4):351–372.

Foster, C. and Order, R. V. (1985). FHA Terminations: A Prelude to Rational Mortgage Pricing. *Real Estate Economics*, 13(3):273–291.

Garmaise, M. J. (2015). Borrower Misreporting and Loan Performance. *Journal of Finance*, 70(1):449–484.

Gatzlaff, D. H. and Haurin, D. R. (1997). Sample Selection Bias and Repeat-Sales Index Estimates. *Journal of Real Estate Finance and Economics*, 14(1-2):33–50.

Gatzlaff, D. H. and Haurin, D. R. (1998). Sample Selection Bias and Biases in Local House Value Indices. *Journal of Urban Economics*, 43(2):199–222.

Genesove, D. and Mayer, C. J. (1997). Equity and Time to Sale in the Real Estate Market. *American Economic Review*, 87(3):255–269.

Genesove, D. and Mayer, C. J. (2001). Loss Aversion and Seller Behavior: Evidence from the Housing Market. *Quarterly Journal of Economics*, 116(4):1233–1260.

Gerardi, K., Herkenhoff, K. F., Ohanian, L. E., and Willen, P. S. (2018). Can't Pay or Won't Pay? Unemployment, Negative Equity, and Strategic Default. *The Review of Financial Studies*, 31(3):1098–1131.

Glaeser, E. L., Gyourko, J., Morales, E., and Nathanson, C. G. (2014). Housing Dynamics: An Urban Approach. *Journal of Urban Economics*, 81:45–56.

Goetzmann, W. N. (1993). Accounting for Taste: Art and the Financial Markets Over Three Centuries. *American Economic Review*, 83(5):1370–1376.

Guiso, L., Sapienza, P., and Zingales, L. (2013). The Determinants of Attitudes Toward Strategic Default on Mortgages. *Journal of Finance*, 68(4):1473–1515.

Guren, A. M. (2018). House Price Momentum and Strategic Complementarity. *Journal of Political Economy*, 126(3):1172–1218.

Guren, A. M., Krishnamurthy, A., and McQuade, T. J. (2019). Mortgage Design in an Equilibrium Model of the Housing Market. *Working Paper*.

Haurin, D. R. and Hendershott, P. H. (1991). House Price Indexes: Issues and Results. *Real Estate Economics*, 19(3):259–269.

Head, A., Lloyd-Ellis, H., and Sun, H. (2014). Search, Liquidity, and the Dynamics of House Prices and Construction. *American Economic Review*, 104(4):1172–1210.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161.

Hubbard, R. G., Skinner, J., and Zeldes, S. P. (1995). Precautionary Saving and Social Insurance. *Journal of Political Economy*, 103(2):360–399.

Hwang, M. and Quigley, J. M. (2004). Selectivity, Quality Adjustment, and Mean Reversion in the Measurement of House Values. *Journal of Real Estate Finance and Economics*, 28(2-3):161–178.

Kaplan, G., Mitman, K., and Violante, G. L. (2019). The Housing Boom and Bust: Model Meets Evidence. *Working Paper*.

Kau, J. B., Keenan, D. C., Muller, W. J., and Epperson, J. F. (1992). A Generalized Valuation Model for Fixed-Rate Residential Mortgages. *Journal of Money, Credit, and Banking*, 24(3):279–299.

Kau, J. B., Keenan, D. C., Muller, W. J., and Epperson, J. F. (1995). The Valuation at Origination of Fixed-Rate Mortgages with Default and Prepayment. *Journal of Real Estate Finance and Economics*, 11:5–36.

Keys, B. J., Mukherjee, T., Seru, A., and Vig, V. (2010). Did Securitization Lead to Lax Screening? Evidence from Subprime Loans. *Quarterly Journal of Economics*, 125(1):307–362.

Krusell, P. and Smith Jr., A. A. (1998). Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy*, 106(5):867–896.

Landvoigt, T. (2017). Housing Demand During the Boom: The Role of Expectations and Credit Constraints. *The Review of Financial Studies*, 30(6):1865–1902.

Landvoigt, T., Piazzesi, M., and Schneider, M. (2015). The Housing Market(s) of San Diego. *American Economic Review*, 105(4):1371–1407.

Laufer, S. (2018). Equity Extraction and Mortgage Default. *Review of Economic Dynamics*, 28:1–33.

Mayer, C., Pence, K., and Sherlund, S. M. (2009). The Rise in Mortgage Defaults. *Journal of Economic Perspectives*, 23(1):27–50.

Mian, A. and Sufi, A. (2009). The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis. *Quarterly Journal of Economics*, 124(4):1449–1496.

Mian, A. and Sufi, A. (2011). House Prices, Home Equity-Based Borrowing, and the U.S. Household Leverage Crisis. *American Economic Review*, 101(5):2132–2156.

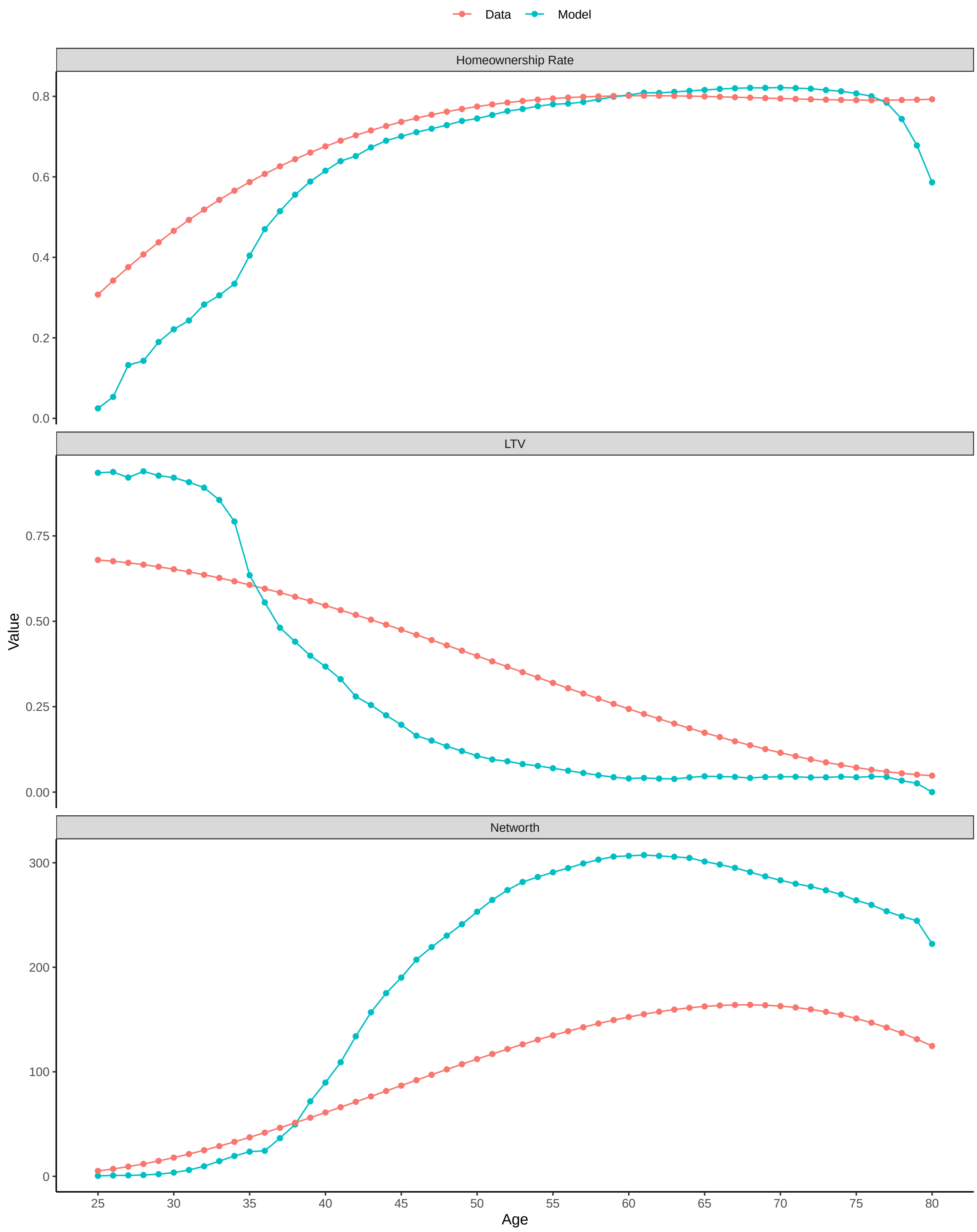
Pollakowski, H. O., Stegman, M. A., and Rohe, W. (1991). Rates of Return on Housing of Low and Moderate-Income Owners. *Real Estate Economics*, 19(3):417–425.

Quigley, J. M. (1987). Interest Rate Variations, Mortgage Prepayments, and Household Mobility. *Review of Economics and Statistics*, 69(4):636–643.

Smith, B. A. and Tesarek, W. P. (1991). House Prices and Regional Real Estate Cycles: Market Adjustments in Houston. *Real Estate Economics*, 19(3):396–416.

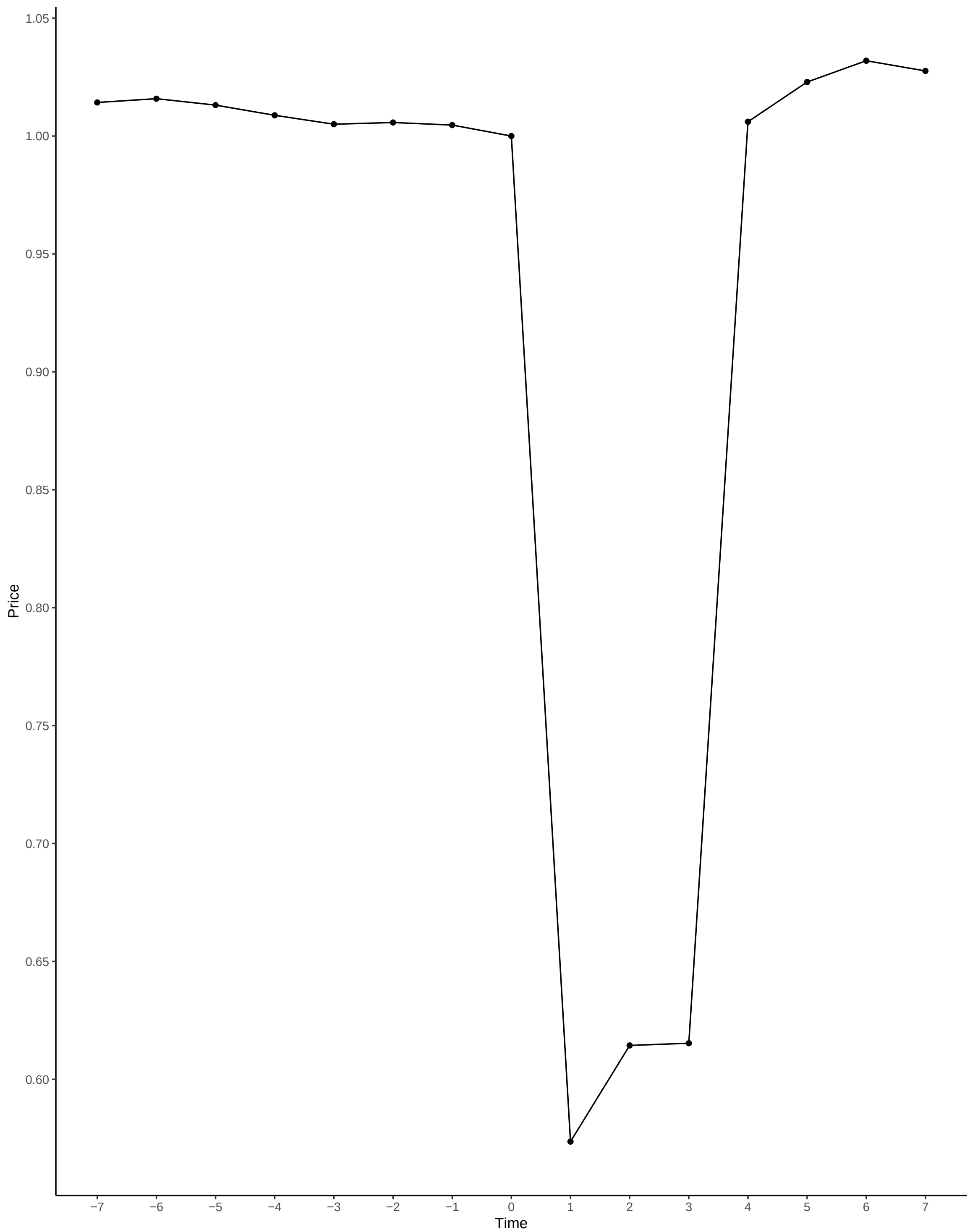
Stein, J. C. (1995). Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects. *Quarterly Journal of Economics*, 110(2):379–406.

Vandell, K. D. (1978). Default Risk Under Alternative Mortgage Instruments. *Journal of Finance*, 33(5):1279–1296.



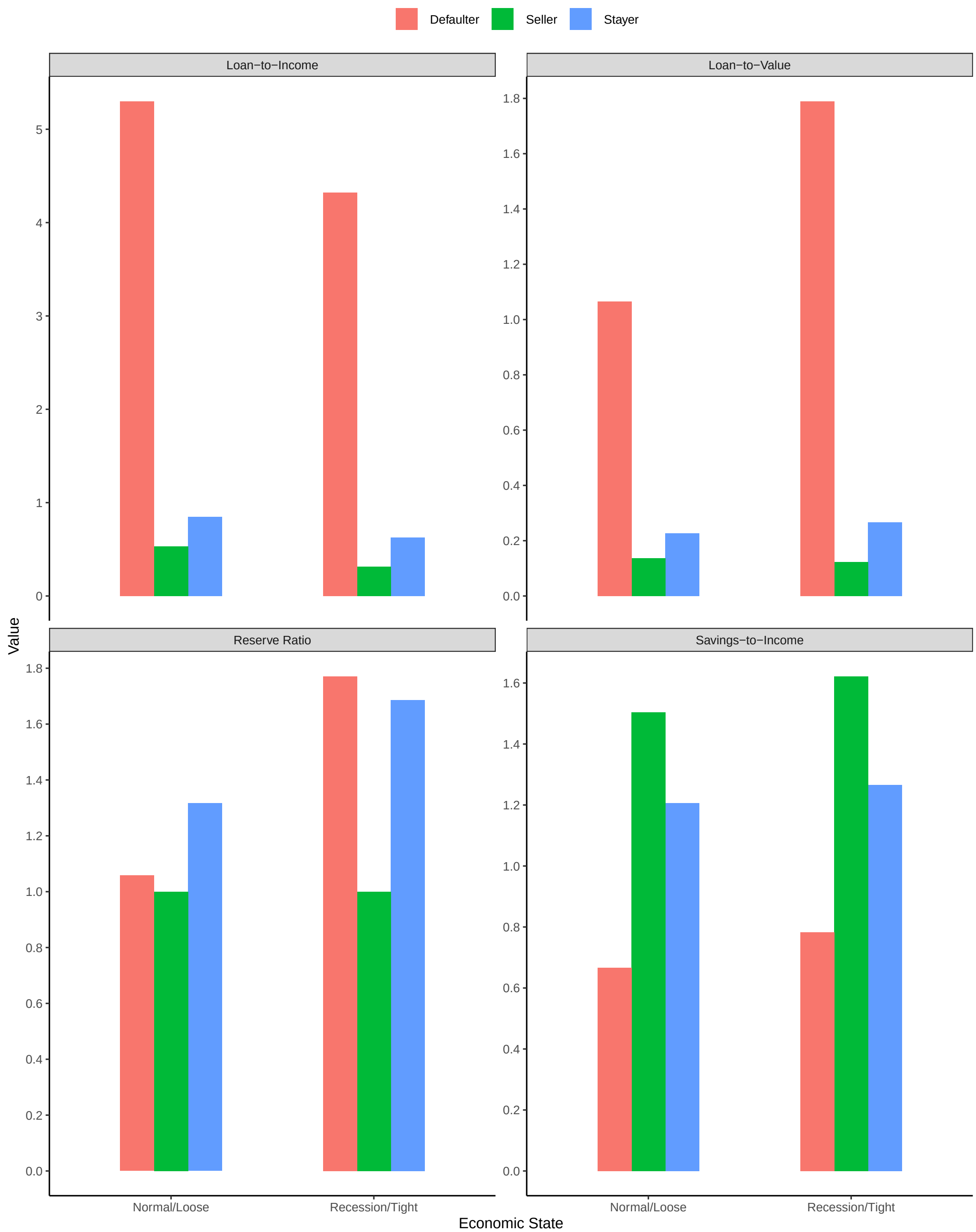
Results obtained from simulation using the policy functions obtained from the model. The first 100 periods are dropped to remove initial distribution assumptions. Data comes from the 1998 SCF. Networth is in the $1,000s.

Figure 2.1: Lifecycle Moments



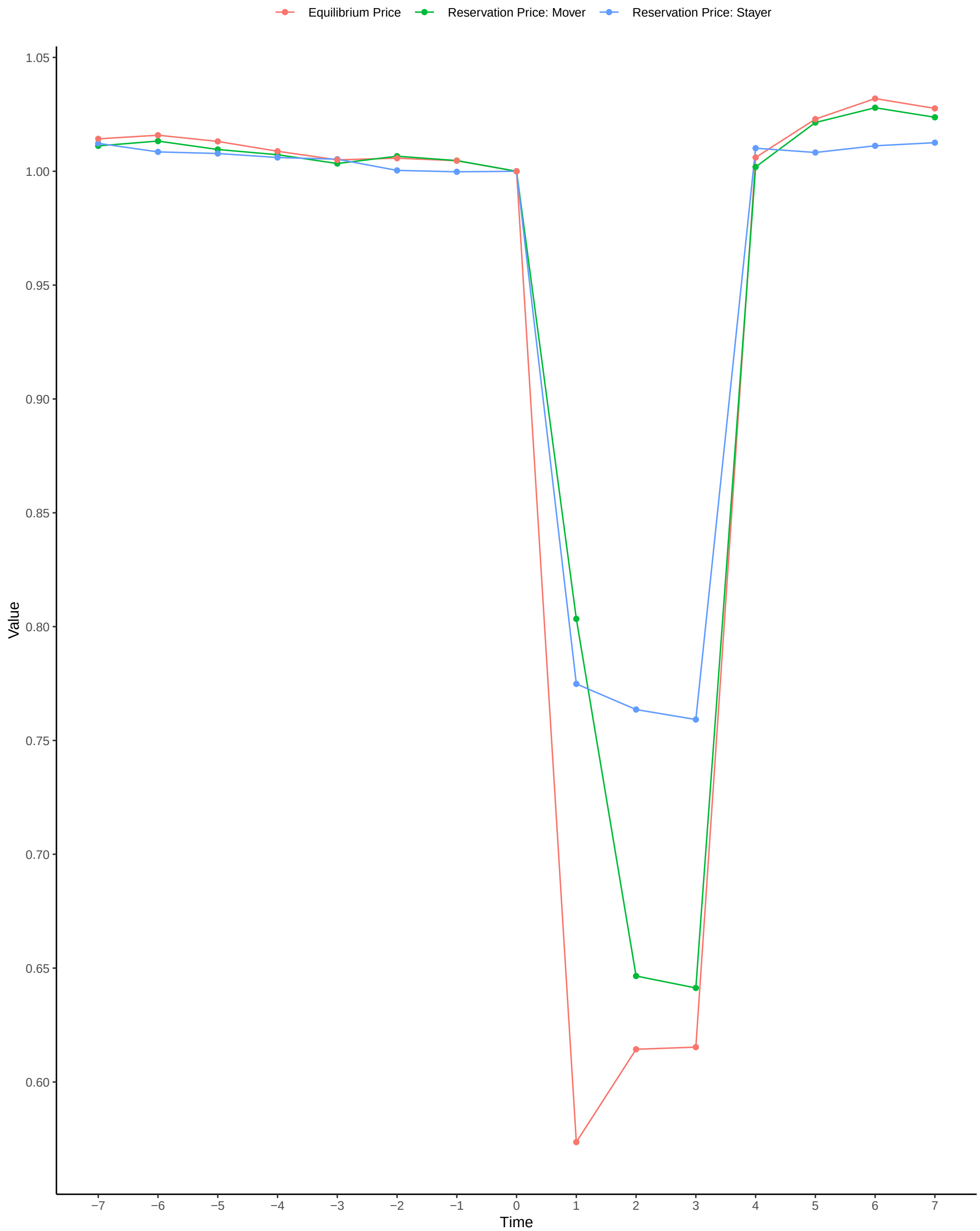
Results obtained from simulation in which equilibrium prices are solved for using Equation 2.17. The first 100 periods are dropped to remove initial distribution assumptions. The first 8 periods represent an economy in a Normal/Loose credit phase. At time 0, the economy enters into a Recession/Tight credit phase for three periods before reverting back to a Normal/Loose credit phase. Prices are relative to time 0.

Figure 2.2: Equilibrium Price Relative to Year 0



Results obtained from simulation using the policy functions obtained from the model. The first 100 periods are dropped to remove initial distribution assumptions. Reserve ratio is the reservation price of the homeowner divided by the equilibrium price in a given period. Home value and income are in $1,000s

Figure 2.3: Comparison Between Homeowners



Results obtained from simulation in which equilibrium prices are solved for using Equation 2.17. The reservation price is the minimum price that homeowners would be willing to accept in order to be indifferent between their actual optimal choice given the equilibrium price and selling. The first 100 periods are dropped to remove initial distribution assumptions. The first 8 periods represent an economy in a Normal/Loose credit phase. At time 0, the economy enters into a Recession/Tight credit phase for three periods before reverting back to a Normal/Loose credit phase. All prices are relative to time 0.

Figure 2.4: Prices Relative to Year 0: Equilibrium Price vs Reservation Price

Table 2.1: U.S. Progressive Tax System Function

Marginal Tax Rate	Taxable Income Bracket
15%	\$0-\$25,350
28%	\$25,350-\$61,400
31%	\$61,400-\$128,100
36%	\$128,100-\$278,450
39.6%	$>$ \$278,450

Note:
This table presents the income tax system laid out by the U.S. tax code in 1998. This is a progressive tax system in which pre-specified income brackets defined by the federal government are taxed at increasing marginal rates.

Table 2.2: Parameters

Description	Parameter	Value	External
Time and Preference			
Time Discount Factor	β	0.95	N
Relative Preference Weight for Housing	θ	2	N
Relative Risk Aversion	γ	3	Y
Bequest Motive	ψ	200	N
Default Utility Penalty	Ξ	0.01	N
Terminal Time	T	80	Y
Retirement Age	K	65	Y
Starting Age		25	Y
Income			
Persistence	ρ	0.959	Y
Standard Deviation of Innovations	σ_z	0.141	Y
Aggregate Income Shocks	ϵ^{agg}	0.149 / 0.215	Y
Mortgage and Savings			
Saving's Rate	r_b	2.79% / 3.34%	Y
Mortgage Spread	ϕ	1.92% / 1.59%	Y
Loan-to-Value Constraint	λ_{ltv}	80% / 95%	Y
Payment-to-Income Constraint	λ_{pti}	36% / 45%	Y
Mortgage Origination Cost	κ_M	3.0%	Y
Housing			
Housing Transaction Cost	κ_H	6.0%	Y
Housing Maintenance Cost	δ_H	2.5%	Y
Housing Tax Cost	τ_H	1.0%	Y
Rental Price	α	0.16	N
Probability of Exogenous Moving Shock	ξ	8.0%	N

Note:
This table presents the parameters used in this paper. Some parameters can take on two values. These are the values affected by the aggregate states. The left value is the value during a recession/tight credit state, and the right value is the value during a normal/loose credit state. The top panel holds the time and preference parameters for the individual. The second panel holds the income parameters. The third panel holds the costs associated with mortgages and savings. The fourth panel holds the costs associated with housing.

Table 2.3: Model Fit

Moment	Model	Data
Homeownership Rate	65.00%	65.00%
Has Mortgage	66.98%	70.00%
Sell	8.17%	8.80%
Default	1.00%	1.00%
Price-to-Income Ratio	4.83	2.79
Rent-to-Price Ratio	3.03%	4.80%
Median Wealth Ratio	1.04	1.35

Note:
Results based on simulation using the policy functions obtained from the model. The first 100 periods are dropped to remove initial distribution assumptions. The median wealth ratio is the median wealth at age 75 divided by the median wealth at age 50.

CHAPTER 3: WALL STREET, MAIN STREET, YOUR STREET: HOW INVESTORS IMPACT THE SINGLE-FAMILY HOUSING MARKET

3.1 Introduction

Since the early 2000s, the U.S. housing market has seen a significant rise, a dramatic fall, and a subsequent rebound in home prices. The current consensus is that the loosening of credit access in the U.S., especially in the mortgage market, was a contributing factor to the large increase in home prices in the early 2000s as well as a contributing factor to the subsequent fall of home prices during the 2007-2009 recession. In this paper, we examine the driving forces in the subsequent rebound in home prices after the 2007-2009 period. Specifically, we look at the emergence of institutional investors in the single-family housing market and empirically test if they contributed to the increase in home prices and if they contributed to the housing affordability issue in the U.S.

Following the Great Recession, credit became harder to obtain even for well-qualified individuals. This was the result of many factors, including: depository institutions exiting the mortgage lending business, depository institutions being capital constrained, and regulatory uncertainty before and after the passage of Dodd-Frank and the Qualified Mortgage standard. In addition, the introduction of HARP and other government programs (which operated through financial institutions) resulted in increased demand for re-financing, which affected the mortgage origination pipelines of lenders. Meanwhile, home prices continued to fall.

However, since 2012, national home prices have risen and have surpassed peak pre-recession prices. At the same time, the average national homeownership rate between 2010 and 2017 was 64.87%, which is below the average national home ownership rate of 68.29% between 2000 and 2006. A contributing factor to this low homeownership rate is the fact that many single-family housing units have been converted into rental units. Although these are national level numbers, many of these trends can be seen at a lower level. For example, each of these trends for the Charlotte region can be found in Figure 3.1. The combination of these three trends has put housing affordability concerns back near the top for policy issues in the U.S.[1]

One of the key differences between the pre-recession and post-recession periods is the increased presence of institutional investors in the single-family housing market, which has been well documented.[2] Institutional investors, such as Real Estate Investment Trusts (REITs), saw that home prices relative to rents were low. They realized they could purchase single-family homes for the purpose of renting them out to the general population. Furthermore, while many individuals were unable to raise funds to purchase homes, institutional investors already had funds to do so, which meant they did not need to obtain a traditional mortgage for each home purchase. This gave institutional investors a significant timing advantage in the housing market.

The perception in the popular press has been that these institutional investors entered the housing market and "out-bid" the traditional single-family home buyer. In essence they are arguing two things: institutional investors were "crowding out" traditional owner-occupied buyers, and institutional investors were willing and able to pay more for a house than non-institutional investors. However, what is not immediate clear is from whom were institutional investors purchasing? The single-family housing market has always had some mix of owner-occupiers and investors participating at any given time. Historically, these investors were the "mom and pop" investors who

[1]See American Housing and Economic Mobility Act of 2019 and the Executive Order Establishing a White House Council on Eliminating Regulatory Barriers to Affordable Housing.

[2]See `https://www.wsj.com/articles/investors-are-buying-more-of-the-u-s-housing-market-than-ever-before-11561023120`, `https://www.theatlantic.com/technology/archive/2019/02/single-family-landlords-wall-street/582394/`, `https://www.npr.org/2019/06/21/734357279/1st-time-homebuyers-are-getting-squeezed-out-by-investors`, or `https://www.nytimes.com/interactive/2019/06/20/business/economy/starter-homes-investors.html` to name a few.

purchased and rented out these homes. It is entirely possible that institutional investors came into the single-family housing market and purchased directly from these "mom and pop" investors and not from owner-occupiers themselves. If this is the case, then one could argue that institutional investors were in a completely different market.

If institutional investors were purchasing a majority of their homes directly from owner-occupiers, then concerns of potential outbidding may be warranted. However, Mills et al. (2019) show that buy-to-rent investors only made up to 0-15% of total home purchases each year between 2012-2014, which is small relative to the overall market. What they do not show is where on the price distribution these institutional investors were purchasing. For example, institutional investors could have segmented the market by identifying homes that are better suited for renting than for owning. These homes might have been of lower quality on average, so the demand for these homes may not have been as high before the entrance of institutional investors. When institutional investors started buying these homes, they boosted demand and left only the homes that were more appropriate for owner-occupiers. If traditional owner-occupiers wanted to buy these lower quality homes, then they would need to pay a premium in order to out-bid these institutional investors. This would lead to an increase in the lower part of the price distribution and leave only higher priced homes in the market, which would naturally boost the average price in an area.

On the other hand, institutional investor managers are professionals who are buying relatively large numbers of homes in the market, so we would expect that they are, on average, more informed about the market than individuals that are buying for owner-occupied use. Thus, institutional investors would be less likely to overpay than owner-occupied purchasers. Furthermore, institutional investors were purchasing at a fragile time in the housing market, which may imply that they had higher bargaining power. Both of these mechanisms would lead to homes being purchased at a discount.

Thus, the net effect that institutional investors have on home prices is unclear. If institutional investors purchased at a discount, they are clearly not raising prices on an individual basis. However, by purchasing homes, they are effectively decreasing the supply of the remaining homes, which would push home prices up on average as long as the remaining homes are being purchased and being purchased at a premium.

In this paper, we empirically examine the effect institutional investors had on home prices at the transaction level. Using transaction level data for Charlotte, North Carolina, between the years 2005-2017, we find that institutional investors paid a discount of about 8.13%-11.19% per transaction. This discount is independent of the discount obtained from purchasing a real-estate owned (REO) home. Additionally, we find that an increase in institutional investor home purchases in the single-family housing market had a statistically positive impact on home prices but a moderate economic impact. That is, a 1 percentage point increase in the percentage of homes purchased by institutional investors in the previous 6-months increased individual home sale prices by about 0.32%-0.59%. The economic significance is about $4,611.63-$5,099.53 for a one standard deviation increase in institutional investor activity in the Charlotte region. Both of these results are consistent across the entire price distribution as well as across various census tract and racial/ethnic groups.

Our results suggest that although investors pay a discount on an individual basis, their aggregate activity increased home prices across the entire price distribution. As more homes continue to be purchased, traditional home buyers (owner-occupiers and non-institutional investors) have to bid higher prices for the lower quality homes, have to buy higher quality homes, or have to leave the single-family housing market. This creates a competition effect and in combination with the tight access to credit, hinders low income/wealth buyers from buying homes. In line with this thought, our results show that individuals are more likely to pay a premium relative to the asking price of the home.

However, we shed some concern that this positive price impact may not all be coming directly from institutional investor activity. In fact, we find that a majority of the price increase is coming from a general increase in demand. Using a difference-in-difference based test, we find that there is no direct price impact on owner-occupier bought homes within 0.1 miles of an institutional investor bought home. Nonetheless, we find that owner-occupier homes further away from institutional investor purchased homes were positively impacted. In other words, homes purchased by owner-occupiers saw an increase in price of about 0.28%-0.34% for each home purchased by an institutional investor within 0.1-0.25 miles. This effect, combined with the increased likelihood owner-occupiers have been paying a premium, suggests that competition for housing among owner-occupiers has increased.

Given the recent surge of institutional investor activity in the single-family housing market, the literature on their effect is small but growing. At the transaction level, our results are most similar to Allen et al. (2018) and Smith and Liu (2018). Both of these papers find that investors paid a discount at the transaction level. We extend their work by examining the competition effect among the remaining single-family home buyers as well as examining the effect across different quantiles of the price distribution. The aforementioned papers only examine the effect on average prices, and in the case of Smith and Liu (2018), only focus on individual effects rather than effects from aggregate investor activity.

In terms of aggregate investor activity, our results are in line with the results of Garriga et al. (2019), Lambie-Hanson et al. (2019), and Mills et al. (2019), in which each find that an increase in investor activity increases home prices at the aggregate level. However, each of these studies only look at the aggregate level and do not look at the transaction level. For this reason, their focus is not on the competition effect among the different types of buyers within the single-family housing market. Thus, they cannot speak on the impact institutional investors have on traditional single-family home purchasers. Given the growing concern of housing affordability in the U.S., the impact institutional investors have on the market, especially on individuals, is a top priority both economically and politically.

The rest of the paper is as follows: Section 3.2 discusses the data, Section 3.3 discusses the main methodology used, Section 3.4 discusses the results, and Section 3.5 includes the closing remarks.

3.2 Data

The main data to study the effects of institutional investors on the housing market comes from Metrostudy, which maintains one of the most comprehensive U.S. housing databases. Metrostudy is a leading provider in residential construction activity across the nation, but they complement this data with transaction level history for a wide range of property types. This transaction level housing data comes from CoreLogic, a major provider of housing data for the U.S, as well as from deed and tax assessor data from individual counties. Since housing is highly localized, we focus our attention to a single area: the Charlotte region between the years 2005-2017.[3] In addition, given that townhomes and condominiums have had a larger history with investors, we restrict our study to the detached single-family housing market.

Each observation contains the sale price, sale date, address, sale type, loan amount (if applicable), and other housing characteristics. Most importantly, each observation also contains the name of the seller, purchaser, and lender. Any observation missing a variable was dropped from the analysis. Furthermore, any home sale considered to be part of a "bulk" sale was also dropped. We define a bulk sale as three or more home sales with the same sale price, sale date, and purchaser name. Using the names of the seller and purchaser, we can identify investors and non-investors in the Charlotte housing market. We define investors as non-individuals that are not banks,

[3]The Charlotte region covers Cabarrus, Gaston, Iredell, Lincoln, Mecklenburg, and Union counties.

mortgage/credit lenders, relocation companies, building companies, nor government entities. This leaves us with companies we believe have purchased homes for investment purposes. Many of the company names include words such as "investment" or "asset management." However, to ensure the accuracy of the data, we inspect each name to make sure they are not an individual or entity listed above.

Using this definition of an investor, we are able to identify 14,850 homes purchased by an investor in the Charlotte region. These observations include homes bought by institutional investors as well as homes bought by non-institutional investors. Given that the main focus of this study is to examine the effect institutional investors had on the single-family housing market, we create a strict definition for what we consider an institutional investor. We define an institutional investor as an investor that has filed as a publicly traded company, that has filed as a REIT with the SEC, or that has filed a Form D with the SEC.[4] Using this definition of an institutional investor, we can identify 7,365 homes purchased by an institutional investor.[5]

Table 4.1 provides the summary statistics for the entire dataset. There are 336,878 home sales with no missing attributes across the 6 counties.[6] A majority of these sales comes from regular resales of existing homes, about a quarter of the sales comes from new construction, and the remaining home sales come from REO sales. 82.55% finance the home purchase with a mortgage, which means that 17.45% paid with only cash-on-hand. The average price in the Charlotte region is $256,231.98.

Table 3.2 breaks the summary statistics down by investors and non-investors. Investors typically do not finance the home purchase with a mortgage and instead pay

[4] Form D is used to file a notice of an exempt offering of securities with the SEC. For more information see: `https://www.sec.gov/smallbusiness/exemptofferings/formd`

[5] These institutional investors include: American Homes 4 Rent, American Housing REIT, Amherst Holdings, Cerberus, Colony Starwood Homes, Conrex, Equity One, Front Yard Residential, Harrison Street, Home Partners of America, Invitation Homes, Main Street Renewal, Progress Residential, Silver Bay Realty Trust, Tamina Homes, Tricon, and West Granite Homes. Many of these institutions are similar to those used in Mills et al. (2019) and Smith and Liu (2018).

[6] Although Metrostudy has foreclosure observations, all of these observations were missing the sales price. For this reason, all foreclosure observations were dropped from the analysis.

with cash-on-hand because only 12.42% of investors obtained a mortgage. In addition, the average loan amount among investors is also relatively small at $19,381.16, whereas non-investors have an average loan of about $190,208.84. However, investors are also buying smaller homes and are more likely to purchase an REO home relative to non-investors. Both of these are typically associated with lower home prices, which is clear given that investors pay an average price of $167,710.56, while non-investors pay an average price of $260,314.06.

The summary statistics in Table 3.2 clearly show that there is a difference in activity and eventual outcome between investors and non-investors. What it does not show is how much of the investor activity is attributed to institutional investors and to non-institutional investors. Table 3.3 shows the difference between these two types of investors. One can see that more than half of the investor observations comes from institutional investors, and 0.00% of these home purchases were financed with a mortgage. Furthermore, institutional investors were more likely to buy non-REO homes. Thus, we would expect that on average, institutional investors will pay a slightly higher price than non-institutional investors, which is the case.

One of the most important aspects of the data is the time period in which institutional investors entered the housing market. Prior to the 2007-2009 recession, home prices were at an all-time high. Once the recession occurred, many individuals began to default on their mortgage. This led to mortgage lenders reducing the number of new mortgage originations and led to policy makers enacting legislation to tighten financial regulation. The outcome of these events was a reduction in home purchases, especially among traditional single-family home buyers. The number of transactions in the Charlotte region can be found in Figure 3.2. We can see that following the beginning of the recession in 2007, home purchases decreased steadily and hit the lowest point in 2011 at about 13,230 transactions. However, since 2011, home transactions have steadily increased.

As competition in the housing market was low, institutional investors realized an opportunity. Figure 3.3 shows the number of homes purchased, the percentage of homes purchased relative to the total market, and the total dollar value of homes purchased by year by investor type. From the figure, we can see that between 2005-2011, there were no homes purchased by institutional investors. The market was entirely made up of traditional single-family home purchasers. Starting in 2012, institutional investors entered the Charlotte housing market, and between 2013-2015, institutional investors purchased nearly double the amount of homes purchased by non-institutional investors. The amount spent by these institutional investors is also significant as they spent between \$49.43-\$332.12 million per year since 2012.

Given that aggregate changes in housing trends in the Charlotte region began to change in 2012 (Figure 3.1), informal conclusions may suggest that these aggregate changes occurred due to the entrance of institutional investors. However, as mentioned earlier, it is entirely possible that institutional investors were purchasing directly from "mom and pop" investors and not from owner-occupiers themselves. Figure 3.4 plots the percentage of single-family homes bought by investors directly from owner-occupiers by year. This figure shows that institutional investors purchased about 61.67%-83.17% of their homes directly from owner-occupiers. This suggest that institutional investors were actively involved in the single-family market and may be directly competing with owner-occupiers.

However, similar to Mills et al. (2019), our data shows that institutional investors only made up about 2.11%-7.40% of total home purchases per year since 2012. This suggests that although they contributed to demand, they did not contribute by an extreme amount. Nonetheless, it is not necessarily the total number of homes purchased that matters but the number of homes purchased in a given price range that matters. For example, if institutional investors were buying extremely low priced homes that no other buyers wanted, then institutional investors will not "crowd out" traditional home buyers. If institutional investors were buying the so called "starter home," then there is a higher likelihood that institutional investors will "crowd out" traditional home buyers, especially owner-occupiers.

Figure 3.5 plots the percentage of homes bought by each purchaser by sale price since 2012. Figure 3.5a compares institutional investors with non-institutional investors, and Figure 3.5b compares institutional investors with owner-occupiers. One can see that both types of investors purchased the majority of their homes below $250,000. Most interestingly, non-institutional investors purchased the majority of their homes at the lowest price level. Non-institutional investors purchased about 54.85% of their homes below $100,000, while institutional and owner-occupiers only purchased 13.62% and 10.48% of their homes under $100,000 respectively. As one moves along the sale price range, non-institutional investors purchased less homes relative to their total purchases, while institutional investors and owner-occupiers purchased slightly more homes. Specifically, institutional investors purchased about 81.94% of their homes between $100,000 and $250,000, while non-institutional and owner-occupiers purchased 38.29% and 50.36% of their homes in the same range respectively. The key takeaway is that as sales price increases and enters into the "starter home" price range, institutional investors conduct a majority of their activity. This further suggests that institutional investors may be directly competing with owner-occupiers.

On top of that, institutional investors are less likely to sell homes relative to traditional single-family buyers. Figure 3.6 shows the number of homes sold and the percentage of homes sold relative to the total market by year by investor type. One can see that institutional investors rarely sold homes in their possession as opposed to non-institutional investors. By holding these homes, institutional investors reduced the supply of single-family housing. In addition, institutional investors were more likely to sell homes to other investors (Figure 3.7a), while non-institutional investors were more likely to sell homes to owner-occupiers (Figure 3.7b).

3.3 Methodology

The summary statistics laid out in Section 3.2 suggest that the entrance of institutional investors in the Charlotte housing market had a pronounced effect on home prices. The main issue with summary statistics is that it does not take into account all the factors that can drive home prices. For example, many papers, such as Anenberg and Kung (2014), Campbell et al. (2011), and Clauretie and Daneshvary (2009), have shown that homes that were REO sales or that sold near a foreclosed home, sold for a significant discount. Given that investors were more likely to purchase REO homes than non-investors, it is hard to conclude from summary statistics that investors increased prices.

In order to tease out the various factors and extract the marginal effect investors had on home prices, we utilize a hedonic regression approach. Given that both types of investors conduct most of the their home buying activity at the lower end of the price distribution and were less likely to obtain traditional mortgage financing, we would expect investors to have a different effect at the lower end than the higher end. Specifically, we expect that investors will have a larger discount at the lower end of the price distribution relative to the higher end.

Since there was high uncertainty surrounding mortgage financing after the Great Recession, it was not immediately clear to sellers whether potential buyers would be able to satisfy the requirements to obtain a mortgage. This increase in uncertainty increased the risk that sellers would not be able to sell their home. Given that investors, on average, had enough funds to pay for the house without a mortgage (and could provide proof of the funds), investors were able to reduce the credit uncertainty issue for the seller. Thus, sellers would be more willing to accept a discount from investors

However, as one moves up the price distribution, the likelihood for single-family purchasers obtaining a mortgage increases, particularly for individuals and non-institutional investors. From Table 3.3, we know that institutional investors never used a traditional mortgage. In addition, Figure 3.5a shows that more non-institutional investors participated in the higher end of the market relative to institutional investors. Therefore, the higher end of the market is made up mostly of non-institutional investors and owner-occupiers, all of whom are more likely to purchase with a mortgage. We would still expect a discount from investors as a whole at the top of the price distribution, but this discount will be less than their discount from the lower end of the market. Thus, as prices increase, the discount obtained from an investor will decrease.

These dynamics also suggest that a purchase with just cash-on-hand will have a negative effect on price. Again, since purchasers are more likely to use just cash-on-hand at the low end of the market relative to the high end, we would expect that as prices increase, the discount from cash should decrease in magnitude. The remaining issue is that it is unclear which of the discounts will have a larger relative magnitude.

In order to examine the effects at various price points, we depart from the standard approach to the hedonic regression by using a quantile approach rather than OLS. Unlike OLS, which examines the marginal effect an explanatory variable can have on the *average* value of the response variable, quantile regressions examine the marginal effect on different points of the response variable distribution. We find this to be a more appropriate approach because we are more interested in how investors, especially institutional investors, affect different parts of the price distribution and not just how they affect the average price. This is important because there are concerns that investors are affecting first-time homebuyers, who mostly concentrate their search efforts at the lower end of the price distribution.

Furthermore, quantile regressions are appealing for a few reasons. Analogous to OLS, quantile regressions minimize a set of residuals. Instead of minimizing the sum of squared residuals like OLS, quantile regressions minimize the sum of absolute deviations. Additionally, quantile regressions make no assumptions about the distribution of the response variable. Finally, quantile regressions use the entire data set and not just certain sub-samples, thus avoiding the sample selection issue laid out in Heckman (1979).

Our main specification for house i, at time t, is as follows:

$$p_{i,t} = \beta_0 + \beta \mathbf{X}_{i,t} + \delta_1 PurchasedBy_{i,t} + \delta_2 REO_{i,t} + \delta_3 NoMortgage_{i,t} + \delta_4 RIA_{C,[t-1,t-6]} + ZC_i + QY_t + \epsilon_{i,t} \quad (3.1)$$

where $p_{i,t}$ is the log sale price, $\mathbf{X}_{i,t}$ are housing characteristics such as number of bedrooms, number of bathrooms, size of lot (acres), finished square footage, and age of the home, $REO_{i,t}$ indicates whether the home was an REO sale, $NoMortgage_{i,t}$ indicates whether the home was purchased without a mortgage or with a mortgage, $RIA_{C,[t-1,t-6]}$ is recent institutional investor activity expressed as a percentage of homes bought by institutional investors in the past six months within county C relative to the total number of homes purchased in the past six months within county C, and ZC_i and QY_t are zip code and quarter-year fixed effects respectively.

For the $PurchasedBy_{i,t}$ variable, we use two separate definitions. The first definition is binary such that $PurchasedBy_{i,t}$ equals 1 if the home was purchased by any type of investor and 0 otherwise. The second definition is categorical such that $PurchasedBy_{i,t}$ represents purchases made by institutional investors, non-institutional investors, or owner-occupiers. One can view the categorical version of this variable as two dummy variables: an institutional investor dummy and a non-institutional investor dummy. These dummies equal 1 if the home was purchased by an institutional investor or non-institutional investor respectively. Note that when both of these dummies equal 0, then the home was purchased by an owner-occupier.

With this specification, we can capture the effect investors have on home prices through δ_1. As mentioned above, at the transaction level, investors are large entities that can quickly and easily purchase homes. In addition, given the slow recovery following the recession, investors may have had significant borrowing power relative to non-investors. Thus, we should expect the sign on δ_1 to be negative, such that investors will pay less than non-investors on average.

In order to assess the effect institutional investors have on home prices as a whole, we include the recent institutional investor activity from the past 6-months.[7] That is, we include the percentage of homes bought by institutional investors relative to the total number of homes purchased from time $t-6$ to $t-1$. If the popular press is correct, then we should see a statistically positive and economically large δ_4 coefficient. However, given that institutional investors only make up a small amount of purchases and are expected to pay a discount, we believe we should not expect a large coefficient. Although, it is entirely possible that the coefficient remains positive as it still represents a demand factor in the Charlotte region.

3.4 Results

3.4.1 Investor's Effect on Home Sale Prices

Table 3.4 shows the results from our main specification. We present the results from the standard hedonic regression using OLS in the first few columns and the results from the quantile in the remaining columns. We report the 25th, 50th, and 75th percentile of the sales price for the quantile regression. We consider the OLS results as our baseline. Panel A shows the results without the recent investor activity variable, and Panel B shows the results with the recent investor activity variable.

From Table 3.4 Panel A, we can see that the baseline model suggests that investors pay a 19.5% discount on an average home in the Charlotte market. This discount decreases to 9.5% when the control for $NoMortgage_{i,t}$ is included. As we look at the

[7] We also use 3, 9, and 12 months, all of which give similar results.

quantile results, we can see that the investor effect is about 14.24%-19.97%. That is, an investor will pay roughly 14.24%-19.97% less than an owner-occupier purchaser. Moreover, when we control for whether the purchase was made with a mortgage or not, the investor discount effect decreases in magnitude to 10.29%-12.99%, but still remains statistically and economically significant.

These results are line with what we would expect from an investor. On average and at various points on the sale price distribution, all types of investors pay a discount relative to owner-occupiers. This discount is both statistically and economically significant. Furthermore, as we move up the price distribution, the investor discount effect has a smaller magnitude when we do not control for a mortgage. When we do control for a mortgage, the investor discount effect increases as we move up the price distribution. This occurs because single-family purchasers at the high end of the market are more likely to use a mortgage. As mentioned before, this creates uncertainty for the seller. Although non-institutional investors still obtain mortgages, they obtain much smaller ones than owner-occupiers because non-institutional investors have a larger amount of available funds to purchase a home. The combination of these two will make investors seem more attractive to the seller. This gives investors an upper hand relative to owner-occupiers with a mortgage. Thus at the high end of the market, investors have a stronger discount effect, which is evident in Table 3.4 Panel A.

When we use our categorical definition for investors (i.e. individual indicator variables for whether the home was purchased by an institutional investor, non-institutional investor, or owner-occupier), we see that non-institutional investors purchase an average home at a price about 29.28% less than owner-occupiers. Institutional investors purchase homes at a price on average about 11.44% less than owner-occupiers. Looking at the quantile regressions, we see that non-institutional investors purchase the 25th, 50th, and 75th percentile home at a price about 35.24%, 30.9%, and 23.22% less than owner-occupiers respectively. Similarly, we see that institutional investors purchase the 25th, 50th, and 75th percentile home at a price about 8.12%, 9.7%, and 11.25% less than owner-occupiers respectively.

These results are consistent with our argument from Section 3.3. That is, non-institutional investors will have a higher discount at the low end of the market relative to their discount at the high end of the market. This is the case because they are more likely to use a mortgage, which will increase uncertainty for the seller. Therefore, non-institutional investors lose some bargaining power when they take on a traditional mortgage. On the other hand, institutional investors never use a traditional mortgage and always use just cash-on-hand. For this reason, their discount remains fairly stable throughout the price distribution.

Institutional investors also produce a significantly lower discount compared to non-institutional investors. The reason being is that non-institutional investors are mostly local, "mom and pop" investors, while institutional investors are non-local, large investors. When it comes to house price dynamics, location is important. Since these non-institutional investors are mostly local, they have a comparative advantage as they know the area better than institutional investors. This may give non-institutional investors the ability to close home sales more quickly relative to institutional investors.

When we control for the recent institutional investor activity in the Charlotte region from the previous 6-months, all the above mentioned effects still hold. This can be seen in Table 3.4 Panel B. Results suggest that investors pay a discount both on average and at various locations on the price distribution. These discounts are both statistically and economically significant. The most important point to note though, is the effect from recent investor activity. We can see that across all specifications, the coefficient on this variable is statistically positive. This means that a 1 percentage point increase in the percentage of homes purchased by institutional investors in the previous 6-months increased individual home sale prices by about 0.32%-0.59%.

Although this is statistically significant, it only has a moderate economic impact on average. That is, for a one standard deviation increase in institutional investor activity in the Charlotte region in the previous 6-months, the average home price sale increased by about $4,611.63-$5,099.53. The impact is even smaller at the lower end of the price distribution. For home prices at the bottom 25th percentile, a one standard deviation increase in institutional investor activity in the previous 6-months increased individual home sale prices by about $1,385.56-$2,673.98.

Note that this is the price increase that owner-occupiers pay. This positive impact does not offset the large discounts that investors gain. Furthermore, this recent investor activity does not explain the large increase in home prices in the Charlotte region since 2012. The home price indices in Figure 3.1 show that since 2012, home prices in the Charlotte region have increased about 5.61%-5.75% per year. However, the results in Table 3.4 Panel B suggest that an increase in investor activity only increased home prices by about 0.32%-0.59%.

3.4.1.1 Investor's Effect Across Groups

One area of concern in terms of housing affordability is that some groups may be more effected by institutional investor purchases than other groups. Specifically, those with low income or those belonging to a racial minority may be more effected. Historically, minorities have been less likely to own homes, and there has been some evidence that discrimination may occur in the housing market.[8] Thus, with the increased presence of institutional investors in the housing market, it is important to investigate whether these minority groups or those with low income have been pushed out of the market.

In order to do this, we group our data by census tract median income and by census tract minority population percentage. We use the Census 5-Year ACS estimates between 2007-2011 to identify median income for each of the census tracts in our

[8] See Ondrich et al. (2003) and Ahmed and Hammarstedt (2008) to name a couple.

data and to identify the census tracts that have a minority population greater than 50%. We create six different census tract groups: census tracts with a minority population greater than 50%, census tracts with a minority population less than or equal to 50%, census tracts with median income less than \$35,100, census tracts with median income between \$35,100-\$44,600, census tracts with median income between \$44,600-\$58,000, and census tracts with median income greater than \$58,000.

We then rerun our main specification (Equation (3.1)) on these six groups. To get a meaningful interpretation of our RIA variable, we change it so that it now represents the group of interest rather than the county. That is, RIA for this test is the percentage of homes bought by institutional investors in the past six months within group G relative to the total number of homes purchased in the past six months within group G, where group is the census tract group of interest.

The results for this test can be found in Table 3.5. We can see that similar to our main results institutional investors purchase at a discount of about 6.34%-18.35% and that institutional investor activity had a similar impact across all six groups. Although, the discount effect in high minority areas was slightly lower, these results suggest that regardless of race or income of the individuals in a given area, institutional investors did not outbid owner-occupiers. These results also suggest that no group was more adversely effected than another group.[9]

3.4.2 Investor's Effect on Home Listing Prices

On the surface, our results suggest that investors did not *directly* contribute to rising home prices but that they contributed *indirectly*, although only slightly. The next question we address is through which avenue did institutional investors affect other single-family purchasers. The most natural approach to answer this question is did institutional investors increase competition within the single-family market? In

[9]Note that running this specification by area group may still cover up the micro-level heterogeneity at the individual level. In the appendix, we rerun the main specification using the race of the individual. We find results consistent with Table 3.5.

order to answer this question, we first need to know whether owner-occupiers paid a premium relative to the list price of the home and if they were more likely to pay a premium once institutional investors entered the market.

The results from the previous section do not tell us if owner-occupiers were paying a premium relative to the list price of the home. The results only tell us that owner-occupiers paid more than investors, yet home prices, in general, have been increasing substantially. It could be the case that the entrance of investors led homeowners to pay a larger amount over the list price in order to "outbid" investors. This process would still affect home prices *indirectly*. That is, as more investors entered into the market, homeowners were more likely to pay a higher amount over the list price. This would lead to an increase in the final sales price. If investors do indeed have a significant impact on the *likelihood* of owner-occupiers paying a premium relative to the list price of the home, then concerns of investors creating a housing affordability issue may be warranted.

To test this conjecture, we utilize data from the MLS for the Charlotte region over the 2005-2017 time period since Metrostudy does not have list price data. Each observation in the MLS contains the sale price, sale date, address, and other housing characteristics. Most importantly, each observation also contains the list price as well as the listing contract date. The latter is used to construct the number of days on the market. Unfortunately, the MLS does not contain the name of the seller nor the name of the purchaser. In order to obtain this information, we match the MLS data with the Metrostudy data by sale date and address. Only observations found in both data sets were kept. Summary statistics for this merged data can be found in Table 3.6.

The largest issue with this merge is that any home not sold by a real estate agent gets dropped from the data set. Thus, the final merged data set consists of 244,453 observations, which is about 72.56% of the full Metrostudy data set used in the previous section. Nonetheless, we are able to preserve 10,694 purchases by investors, which is 72.01% of the number of investor purchases in the full Metrostudy data. Given that there are less investors in this merged data, we should expect a slightly higher sales price. The sales price in the merged data is \$260,723.86, which is slightly higher than the \$256,231.98 sales price in the Metrostudy data. Finally, we note that the housing characteristics are on average very similar between the two sets of data.

The key piece of this merged data is the days on the market (DOM) and list price variable. Table 3.6 shows that the average DOM and average list price in the Charlotte region between 2005-2017 was 118 and \$268,707.65 respectively. That is, it takes roughly 118 days to sell a home and on average, homes list for a higher price than the final sales price. Since 2011, the number of days it takes a home to sell has drastically decreased and has dipped below the long run average. At the same time, both the list price and the sale price have increased, and the list price has consistently been larger than the sale price. These trends can be seen in Figure 3.8. It is also worth mentioning that since 2011, the gap between the sales price and the list price has shrunk. These statistics suggest that, on average, single-family home purchasers are paying higher prices, are paying closer to the list price, although only marginally, and are closing transactions quicker. Furthermore, these trends hold for owner-occupiers as well as for institutional investors, which makes it difficult to interpret whether the entrance of institutional investors led to owner-occupiers paying closer to the list price.

To formalize this finding, we estimate a logit model to determine the likelihood that a purchaser pays more than the list price of the home.

$$Pr(SalePrice > ListPrice)_{i,t} = \Lambda(\beta\mathbf{X}_{i,t} + \delta_1 PurchasedBy_{i,t} + \delta_2 REO_{i,t} + \delta_3 NoMortgage_{i,t} + \delta_4 RIA_{C,[t-1,t-6]} + ZC_i + QY_t) \tag{3.2}$$

where $Pr(SalePrice > ListPrice)_{i,t}$ is the probability that the sale price is larger than the list price for transaction i at time time t, and Λ is the logistic function. The remaining variables are similar to Equation (3.1).

Table 3.7 shows the results for this logistic test. We can see that across each specification, the coefficient on the *PurchasedBy* variable is negative and statistically significant, while the coefficient on the *RIA* variable is positive and statistically significant. The marginal effects associated with the *RIA* variable suggests that a one standard deviation increase in the number of homes purchased by institutional investors relative to owner-occupiers within the past 6-months, increases the probability that an owner-occupier will pay more than the list price of an average home by about 1.49%-1.58%. Consistent with previous results, this suggests that institutional investors had an indirect effect on prices through the list price channel. As more institutional investors entered the single-family market, owner-occupiers were more likely to pay a higher value for a home, possibly to "outbid" investors.

3.4.3 The Bias with Home Listing Prices

One major issue with the previous result is that listing prices of homes may be strategically chosen by the seller. There is a long literature that shows that both theoretically and empirically, a higher listing price will lead to a longer time-on-the-market. However, at the same time, home sellers do not want to set too low of a listing price because they will lose out on potential capital gains.[10] Therefore, current home sellers may be choosing a listing price to fit their own needs rather than choosing a listing price that is comparable to a fair market value. This may bias our results, but it is difficult to determine in which direction the bias occurs.

In order to limit the bias issues, we must remove the strategical component of the home seller. However, this is difficult because we do not know the reason why

[10] See Arnold (1999), Knight (2002), Cheng et al. (2008), Haurin et al. (2010), and Carrillo (2013) to name a few.

the homeowner is selling. All we know is that home sellers are balancing time-on-the-market and capital gains, which may imply that the listing price home sellers ultimately set may be above or below the fair market value of the home. Thus, we need to estimate a listing price that is consistent with a fair market value of the home because in the end, we are concerned whether the increase in institutional investor home purchases increased the likelihood that owner-occupiers paid more than the market value of the home.

To accomplish this task, we need to estimate each home price, $p_{i,t}$, in our sample. We use a hedonic regression to carry out this estimation. However, using prices at and beyond time t as well as using home prices significantly further back historically may confound this estimated price. Thus, we use a 6-month rolling hedonic regression from time $t-6$ to time $t-1$ to estimate home price i at time t. Our hedonic regression includes only housing characteristics and zip code fixed effects. We then replace the list price of home i at time t with the imputed value from our rolling regression and re-estimate Equation (3.2).

The result from this re-estimation can be found in Table 3.8. We can see that similar to the previous results, the coefficient on the *PurchasedBy* variable is negative and statistically significant, while the coefficient on the *RIA* variable is positive and statistically significant. The marginal effects are slightly larger in magnitude for both variables but tell a consistent story: institutional investors may have had an indirect effect on prices through the list price channel. That is, owner-occupiers were more likely to pay higher than the market value of the home relative to institutional investors and were more likely to pay higher than the market value of the home as more homes were purchased by institutional investors.

3.4.4 Robustness

3.4.4.1 Propensity Score Analysis

The major caveat with our results thus far is the fact that there is an ample self selection issue surrounding investors. Institutional investors came into the single-family market because they realized that home prices relative to rents were low. Additionally, these investors are large corporations that want to maximize their return. We would expect investors to buy certain *types* of homes that are relatively cheap but can be rented at or above the market rent. For example, American Homes 4 Rent, one of the largest institutional investors in our data, stated in their 2016 annual report exactly what type of home in which they want to buy:

> "We are focused on acquiring homes with a number of key property characteristics, including: (i) construction after 1990; (ii) three or more bedrooms; (iii) two or more bathrooms; (iv) a range of $100,000 estimated minimum valuation to $350,000 maximum bid price; and (v) estimated renovation costs not in excess of 25% of estimated value. We target areas with above average median household incomes, well-regarded school districts and access to desirable lifestyle amenities."

Given that institutional investors choose homes based off some type of observables, we know that this sample is not random leading to our prior results being biased. Specifically, it could be that institutional investors choose some *type* of home that they believe can be obtained at a low value. If this is the case, then institutional investors are not necessarily paying a discount but instead, are finding low values homes relative to owner-occupiers. This will inflate the discounts that we uncovered in the previous sections. Thus, one could view our discount results as an upper limit.

In order to mitigate this endogeneity issue, we employ a Propensity Score Matching (PSM) algorithm to match all of the treated units with all of the untreated units based on a set of observables. By matching units based on a set of observables, we effectively remove any observable difference between treated and untreated units such that any difference in the outcome will be due to the treatment effect. In our case, a treated unit is a home bought by an institutional investor, and an untreated unit is a home bought by an owner-occupier.

In order to implement the PSM algorithm, we again use the full Metrostudy similar to Section 3.4.1. For each zip code, year pair, we run a logistic regression in which the dependent variable is equal to 1 if the home was bought by an institutional investor and 0 if the home was bought by an owner-occupier. The independent variables include only observable housing characteristics such as number of bedrooms, number of bathrooms, size of lot (acres), finished square footage, and age of the home. From this logistic regression, we obtain the predicted probability (propensity score) of the home being a treated or untreated home. We then use a nearest neighbor matching algorithm to match each of the treated homes with an untreated home based on the propensity score. This results in a balanced sample of treated and untreated homes based on observable characteristics. We do three separate matches: institutional investor home matched with any owner-occupier home, institutional investor home matched with an owner-occupier home with a mortgage, and institutional investor home matched with an owner-occupier home without a mortgage.

Summary statistics for the three matched samples can be found in Table 3.9. As expected, for each of the three samples, the average observable home characteristics are similar between institutional investor bought homes and owner-occupier bought homes. Most importantly, owner-occupiers pay about $21,199.74-$23,857.21 more than institutional investors for a similar home. This supports prior results that institutional investors pay a discount relative to owner-occupiers.

However, note that the difference in the average sale price does not take into account the different physical locations of the home nor the different time periods in which each home sold. In order to control for this, we re-estimate a modified Equation (3.1). The modified specification includes all the same independent variables except for the housing characteristics because by definition all the homes in each sample have similar housing characteristics. We can see that as expected, when we control for the selection bias, our estimates decrease in magnitude. The most striking result though is the estimate for the *RIA*. The estimate for this variable is again positive, but it is no longer statistically significant. This suggests that although there may be some induced competition among owner-occupiers after the entrance of institutional investors into the single-family housing market, this effect is fairly weak.

3.4.4.2 Difference-in-Difference Analysis

The results up to now strongly suggest that institutional investors did pay lower than owner-occupiers both across the price distribution as well as across census tract areas. The results also show that as more institutional investors entered the single-family housing market, owner-occupiers paid more for housing. However, the propensity score analysis suggests this may not be the case, at least not for starter homes. In this section, we attempt to investigate whether this price increase from institutional investors was directly from institutional investors or just from a general increase in demand. The issue with our previous measure for institutional investor activity (*RIA*) is that it could just be picking up on a general increase in demand in the housing market rather than a direct price effect from institutional investors. That is, prices could have already been increasing before institutional investors entered the market.

In order to test the potential spillover effect from institutional investors, we utilize the difference-in-difference framework of Linden and Rockoff (2008), which was subsequently used in the foreclosure literature by Campbell et al. (2011) and Anenberg and Kung (2014). This methodology compares the price of a home by an owner-occupier before and after the purchase of a home by an institutional investor. This comparison is done with institutional investor purchased homes close to (within 0.1 miles) and further away from (within 0.25 miles) the owner-occupier purchased home. The main assumption is that at the hyper-local level, price trends in housing are similar, but there should be varying effects depending on the distance from the institutional investor purchased home.

Formally, this can be written as:

$$\begin{aligned} p_{i,t} = \beta_0 + \beta \mathbf{X}_{i,t} + \delta_{Close,Before} N_{i,t}^{Close,Before} + \delta_{Close,After} N_{i,t}^{Close,After} + \\ \delta_{Far,Before} N_{i,t}^{Far,Before} + \delta_{Far,After} N_{i,t}^{Far,After} + \delta_2 REO_{i,t} + \\ \delta_3 NoMortgage_{i,t} + CTY_i + \epsilon_{i,t} \end{aligned} \tag{3.3}$$

where $N_{i,t}^{Close,Before}$ and $N_{i,t}^{Close,After}$ are the number of homes purchased by an institutional investor within 0.1 miles of home i both before and after the purchase of home i respectively, and $N_{i,t}^{Far,Before}$ and $N_{i,t}^{Far,After}$ are the number of homes purchased by an institutional investor within 0.25 miles of home i both before and after the purchase of home i respectively. *Before* are all home sales that occurred one or two years before home sale i, and *After* are all home sales that occurred one or two years after home sale i. CTY_i is the census tract-year fixed effect. The estimated spillover effect comes from the difference between $\delta_{Close,Before}$ and $\delta_{Close,After}$. Any difference between $\delta_{Far,Before}$ and $\delta_{Far,After}$ estimates general price changes in the area.

The results for the difference-in-difference test can be found in Table 3.11. The first two columns represent institutional investor home purchases within one year before and after the purchase of home i. The last two columns represent institutional investor home purchases within two years before and after the purchase of home i. Both sets of columns report the difference between $\delta_{Close,Before}$ and $\delta_{Close,After}$ and the difference between $\delta_{Far,Before}$ and $\delta_{Far,After}$.

When only institutional investor homes purchased within 0.1 miles are being controlled, we can see that the difference between $\delta_{Close,Before}$ and $\delta_{Close,After}$ is positive. This suggest that an additional home purchased by an institutional investor within 0.1 miles increases owner-occupier prices by about 0.18%-0.49%, however only the estimate from the two year window is statistically significant. This is expected because the two year window will have more institutional investor home purchases, therefore representing a larger demand increase. Nonetheless, the difference in "close" coefficients suggest that there may be a spillover effect from institutional investor home purchases.

However, there still could be a general demand shock in this hyper-local area. To control for any price increases in the area, we include institutional investor home purchases within 0.25 miles of home i. We can see that once this variable is controlled for, all effects from "close" homes goes to 0, while institutional investor homes purchased between 0.1 and 0.25 miles away absorb all the positive price effects. This suggest that all positive price impacts may be coming from a general increase in home prices within the area and that there is no real spillover effect from institutional investor home purchases.

3.5 Conclusion

Using various tests, we empirically find that institutional investors paid a discount of about 8.13%-11.19% per transaction. Our results also suggest that as more institutional investors bought single-family homes relative to owner-occupiers, owner-occupiers paid about 0.32%-0.59% higher for single-family homes than institutional investors. However, this effect is fairly weak both statistically and economically. We find that this positive price impact may be coming entirely from a general increase in demand and not necessarily from institutional investors themselves. That is, we find no real spillover effect of institutional investor home purchases.

Given the growing concern of increased home prices, we find little evidence that the entrance of institutional investors had a significant effect on rising home prices. Nonetheless, institutional investors target specific types of homes, especially the "starter" home, which reduces the supply of these homes in the area. Therefore, owner-occupiers must search for other types of homes in the area or stay out of the market. Given that the number of total home sale transactions is below the pre-recession peak, it may be the case that there is not an adequate *supply* of "starter" homes in the area, which will push lower income individuals out of the market. In addition, the remaining homes are aggressively being purchased at or above the listing price as more institutional investors enter into the market. These dynamics would magnify any housing affordability issue.

REFERENCES

Ahmed, A. M. and Hammarstedt, M. (2008). Discrimination in the Rental Housing Market: A Field Experiment on the Internet. *Journal of Urban Economics*, 64(2):362–372.

Allen, M. T., Rutherford, J., Rutherford, R., and Yavas, A. (2018). Impact of Investors in Distressed Housing Markets. *Journal of Real Estate Finance and Economics*, 56(4):622–652.

Anenberg, E. and Kung, E. (2014). Estimates of the Size and Source of Price Declines Due to Nearby Foreclosures. *American Economic Review*, 104(8):2527–2551.

Arnold, M. A. (1999). Search, Bargaining and Optimal Asking Prices. *Real Estate Economics*, 27(3):453–481.

Bayer, P., Ferreira, F., and Ross, S. L. (2018). What Drives Racial and Ethnic Differences in High-Cost Mortgages? The Role of High-Risk Lenders. *Review of Financial Studies*, 31(1):175–205.

Campbell, J. Y., Giglio, S., and Pathak, P. (2011). Forced Sales and House Prices. *American Economic Review*, 101(5):2108–2131.

Carrillo, P. E. (2013). To Sell or Not to Sell: Measuring the Heat of the Housing Market. *Real Estate Economics*, 41(2):310–346.

Cheng, P., Lin, Z., and Liu, Y. (2008). A Model of Time-on-Market and Real Estate Price Under Sequential Search with Recall. *Real Estate Economics*, 36(4):813–843.

Clauretie, T. M. and Daneshvary, N. (2009). Estimating the House Foreclosure Discount Corrected for Spatial Price Interdependence and Endogeneity of Marketing Time. *Real Estate Economics*, 37(1):43–67.

Consumer Financial Protection Bureau (2014). Using Publicly Available Information to Proxy for Unidentified Race and Ethnicity.

Diamond, R. and McQuade, T. (2019). Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development. *Journal of Political Economy*, 127(3):1063–1117.

Diamond, R., McQuade, T., and Qian, F. (2019). The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco. *American Economic Review*, 109(9):3365–3394.

Elliott, M. N., Morrison, P. A., Fremont, A., McCaffrey, D. F., Pantoja, P., and Lurie, N. (2009). Using the Census Bureau's Surname List to Improve Estimates of Race/Ethnicity and Associated Disparities. *Health Services and Outcomes Research Methodology*, 9(69).

Garriga, C., Gete, P., and Tsouderou, A. (2019). The Dynamic Effects of Investors in Housing Markets. *Working Paper*.

Haurin, D. R., Haurin, J. L., Nadauld, T., and Sanders, A. (2010). List Prices, Sale Prices and Marketing Time: An Application to U.S. Housing Markets. *Real Estate Economics*, 38(4):659–685.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161.

Knight, J. R. (2002). Listing Price, Time on Market, and Ultimate Selling Price: Causes and Effects of Listing Price Changes. *Real Estate Economics*, 30(2):213–237.

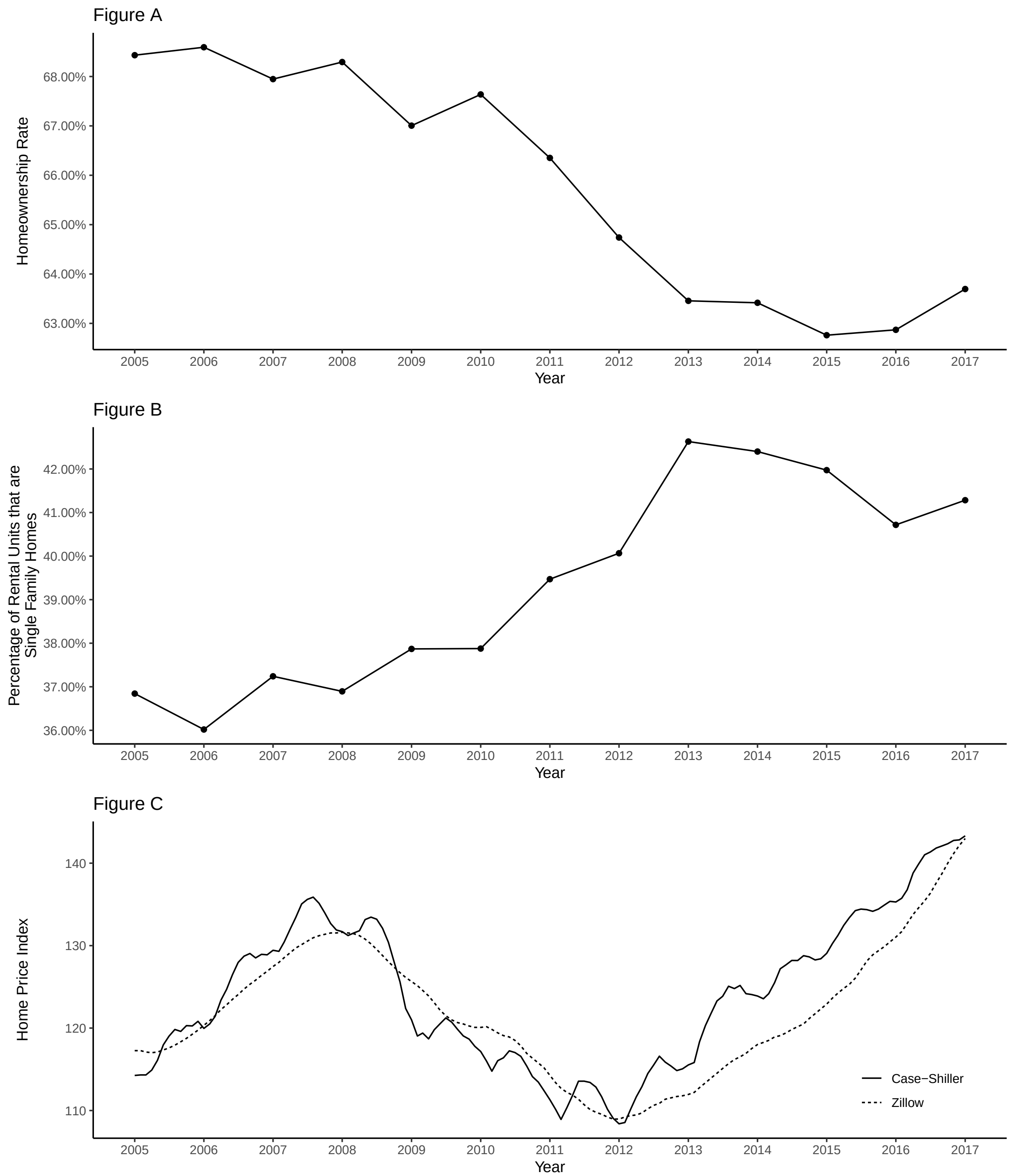
Lambie-Hanson, L., Li, W., and Slonkosky, M. (2019). Leaving Households Behind: Institutional Investors and the U.S. Housing Recovery. *Working Paper*.

Linden, L. and Rockoff, J. E. (2008). Estimates of the Impact of Crime Risk on Property Values from Megan's Laws. *American Economic Review*, 98(3):1103–1127.

Mills, J., Molloy, R., and Zarutskie, R. (2019). Large-Scale Buy-to-Rent Investors in the Single-Family Housing Market: The Emergence of a New Asset Class. *Real Estate Economics*, 47(2):399–430.

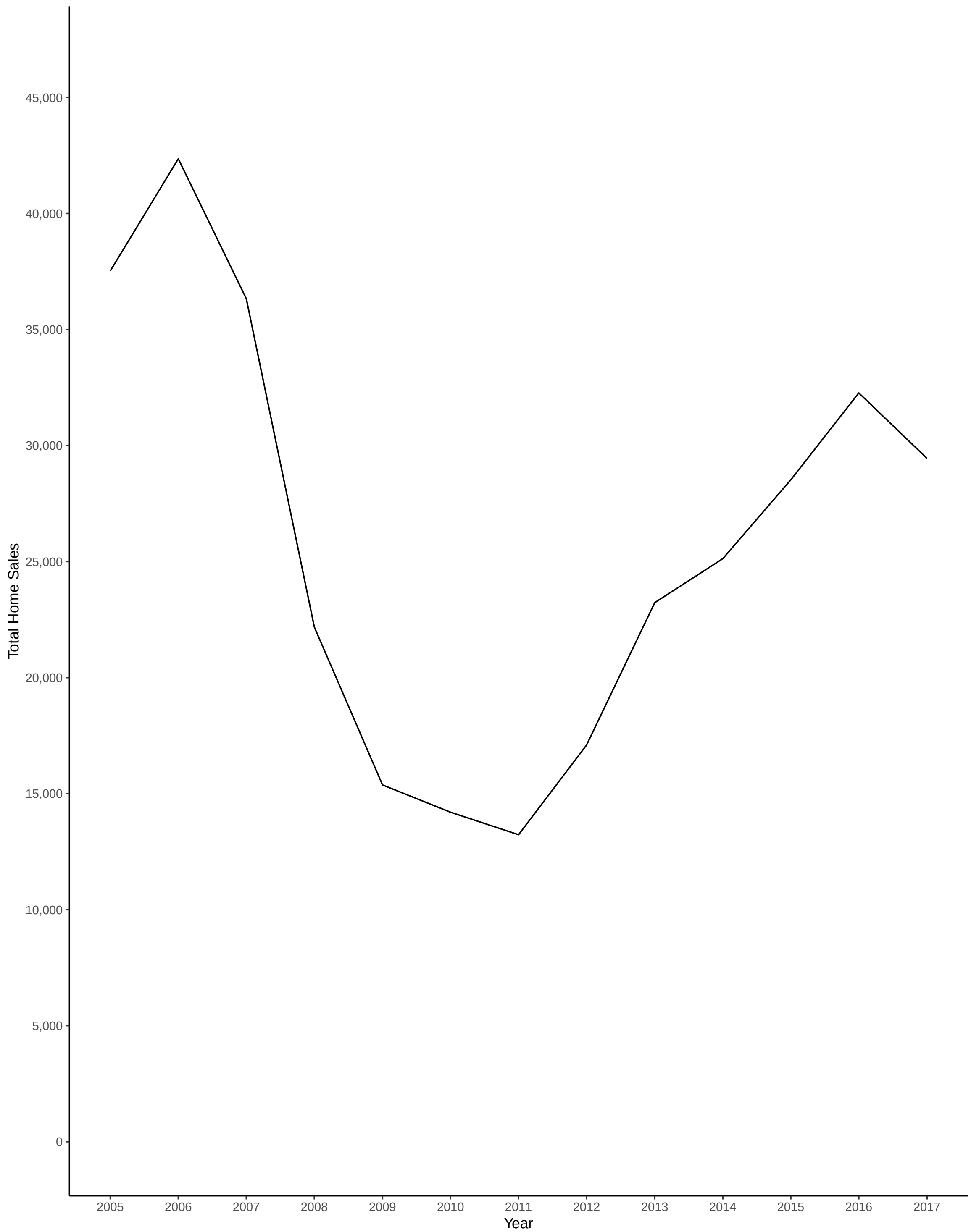
Ondrich, J., Ross, S., and Yinger, J. (2003). Now You See It, Now You Don't: Why Do Real Estate Agents Withhold Available Houses from Black Customers? *Review of Economics and Statistics*, 85(4):854–873.

Smith, P. S. and Liu, C. H. (2018). Institutional Investment, Asset Illiquidity and Post-Crash Housing Market Dynamics. *Real Estate Economics*, 00(0):1–37.



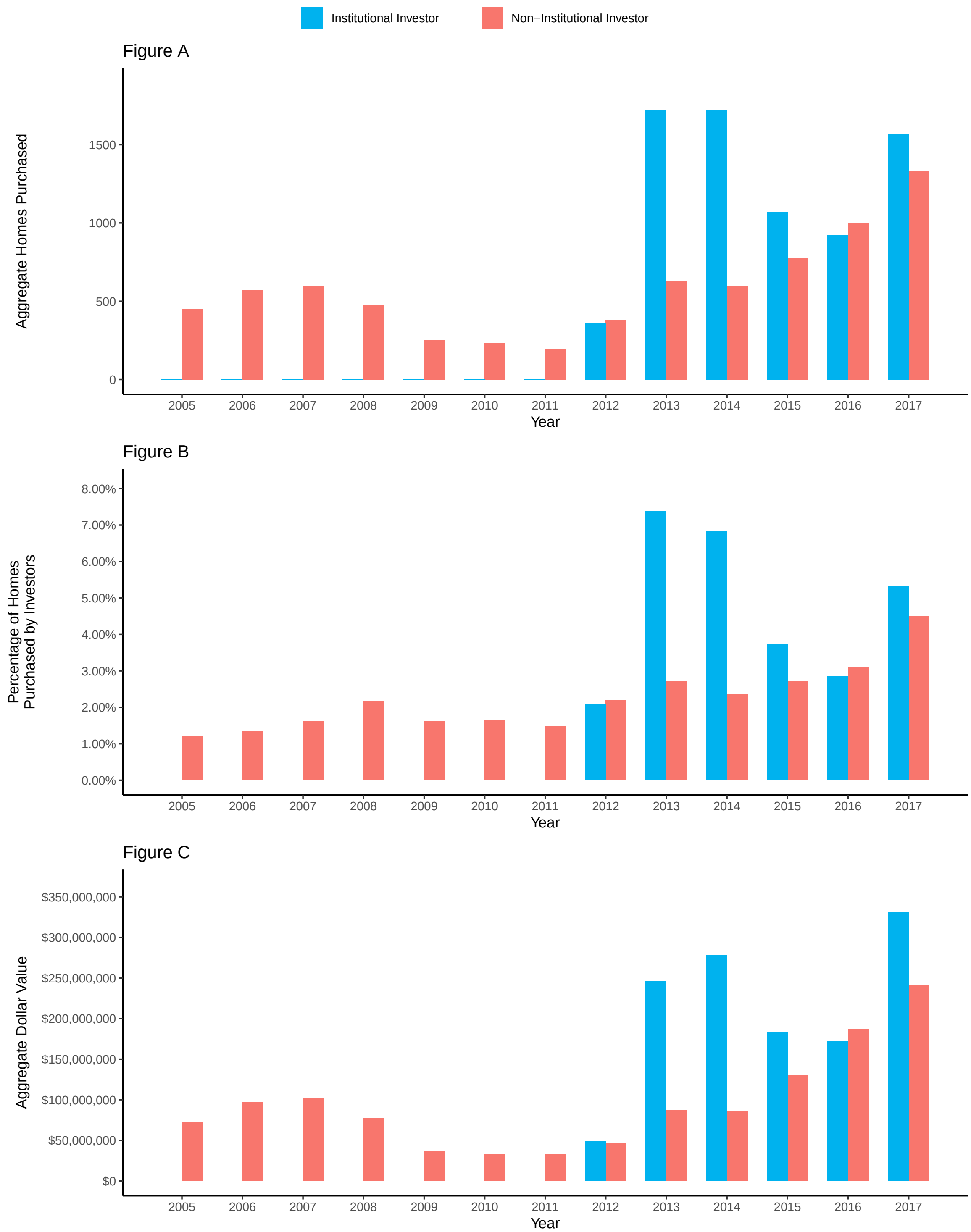
This figure shows the homeownership rate from the 1−year ACS estimates, the percentage of rental units that are single−family units from the 1−year ACS estimates, and the home price index from Case−Shiller and Zillow. The homeownership rate and the percentage of single−family units covers Cabarrus, Gaston, Iredell, Lincoln, Mecklenburg, and Union counties. Single−family rental units include detached and attached units, mobile homes, and units such as RVs and boats. The stock of rental units includes all occupied rental units plus vacant units for rent and units rented but not occupied. Figure 1b is calculated by dividing single−family rental units by the stock of rental units. The Case−Shiller home price index covers Anson Cabarrus, Gaston, Mecklenburg, Union, and York (SC) counties, while the Zillow home price index covers the Charlotte−Gastonia−Concord MSA. The base year for the indices is 2000.

Figure 3.1: Homeownership Rate, Single-Family Rentals, and Home Prices



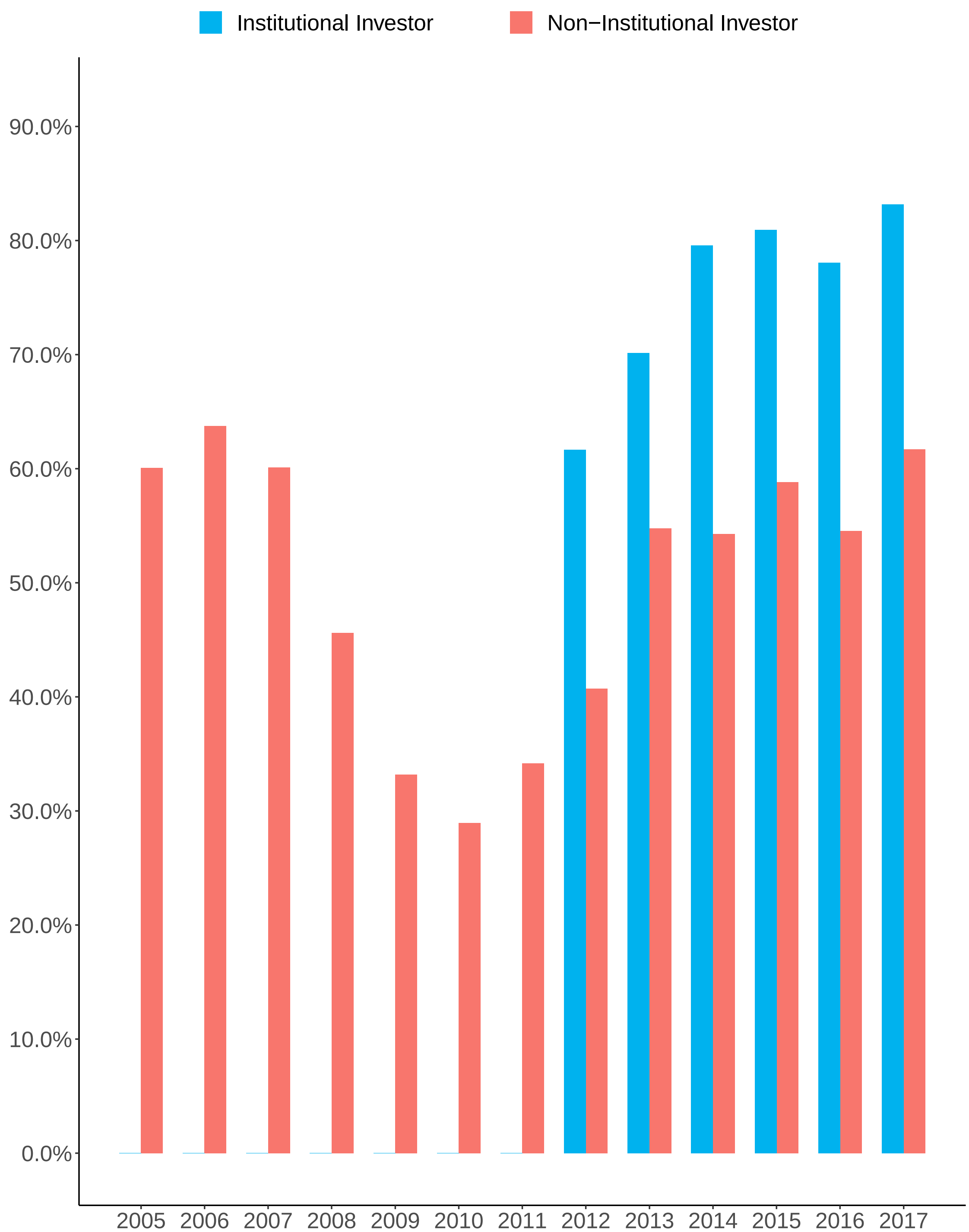
This figure shows the total number of single−family home sales in the Charlotte region. The Charlotte region covers Cabarrus, Gaston, Iredell, Lincoln, Mecklenburg, and Union counties from 2005−2017. Data is from Metrostudy.

Figure 3.2: Total Transactions in the Charlotte Region



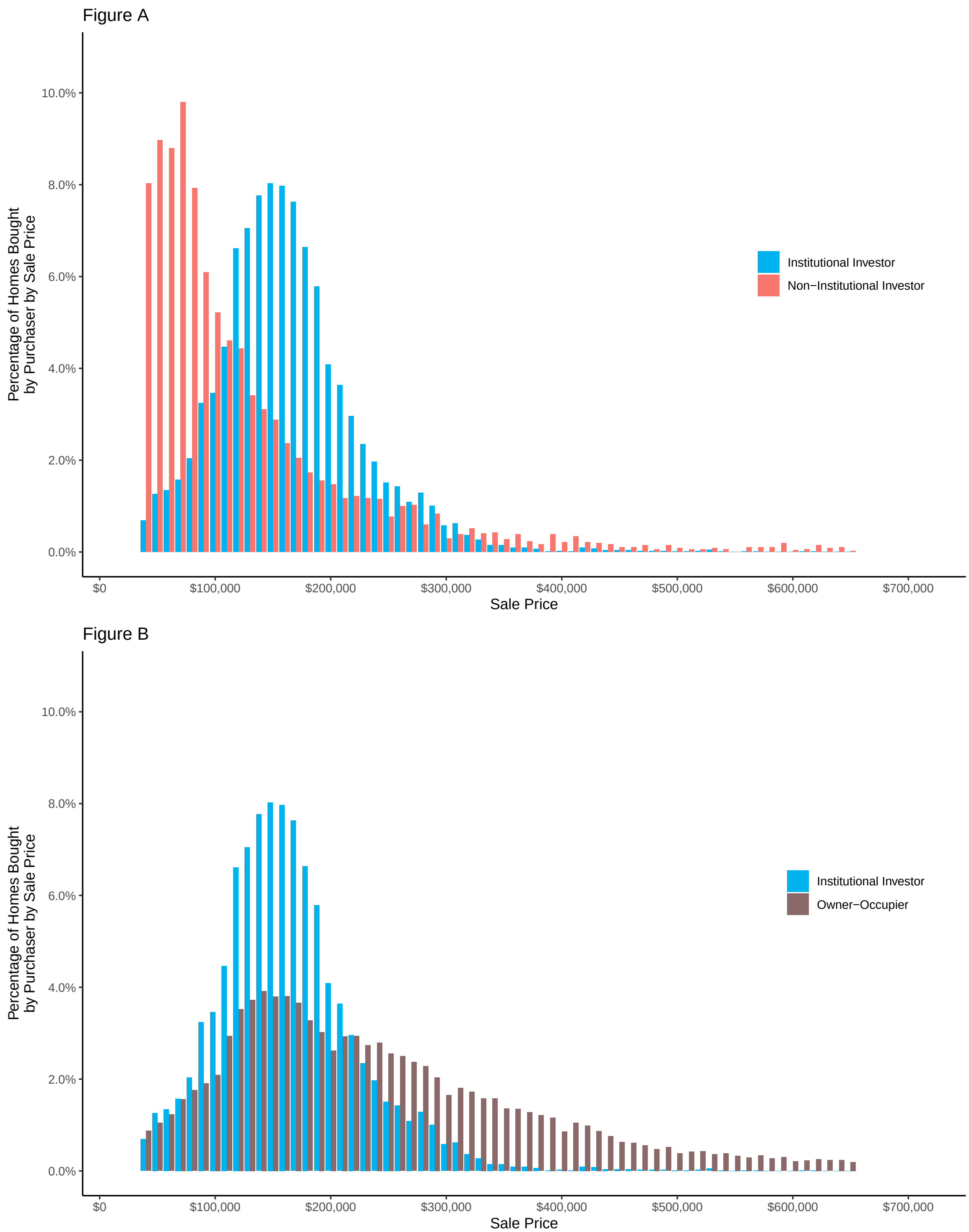
This figure shows the total number of single−family home purchases made by investors, the percentage of homes purchased by investors, and the total dollar value of homes purchased by investors. Investors are defined as non−individuals that are not banks, mortgage/credit lenders, relocation companies, building companies, nor government entities. Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

Figure 3.3: Homes Purchased by Investors



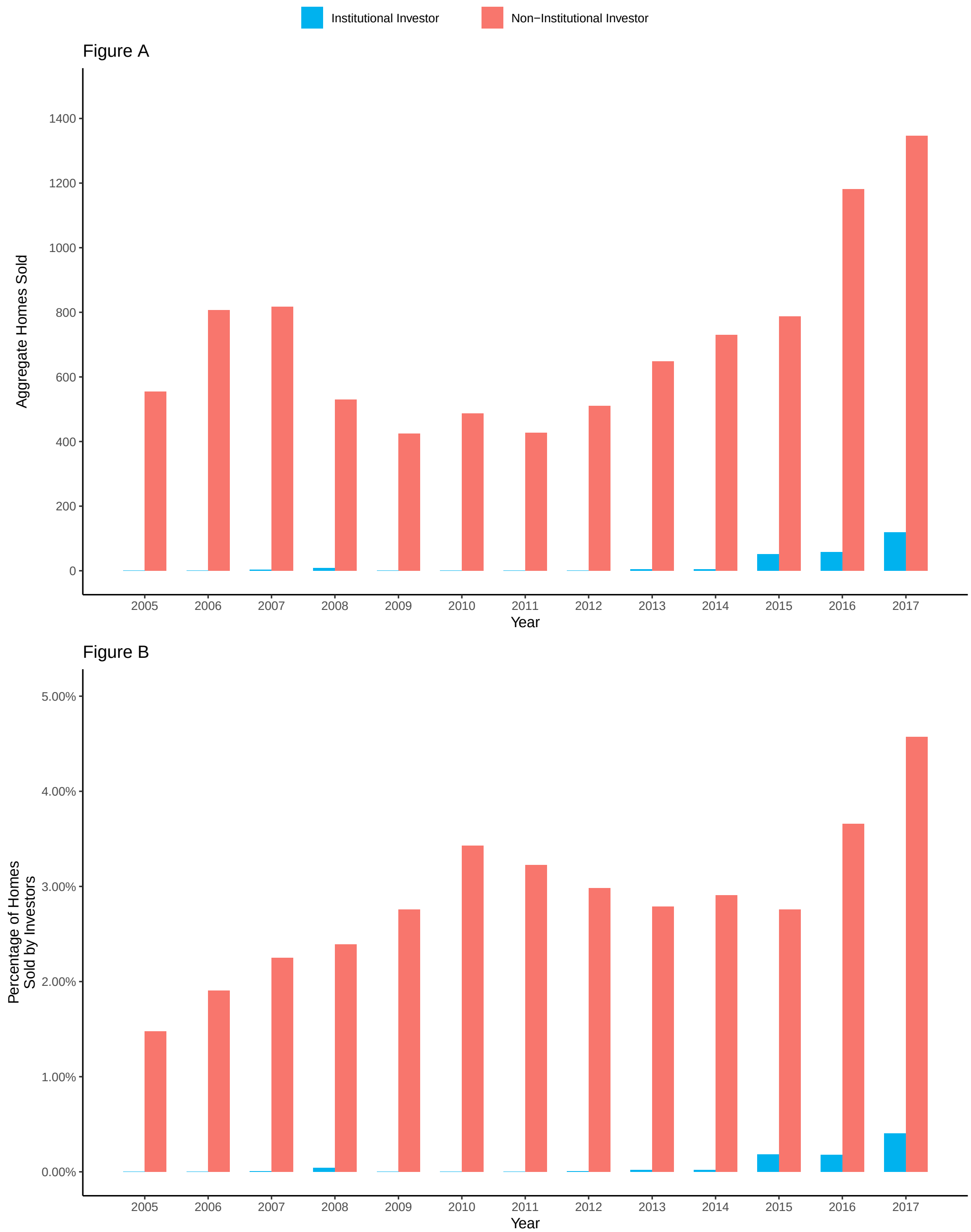
This figure shows the percentage of single−family homes bought by investors directly from owner−occupiers by year. Investors are defined as non−individuals that are not banks, mortgage/credit lenders, relocation companies, building companies, nor government entities. Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

Figure 3.4: Homes Purchased by Investors from Owner-Occupiers



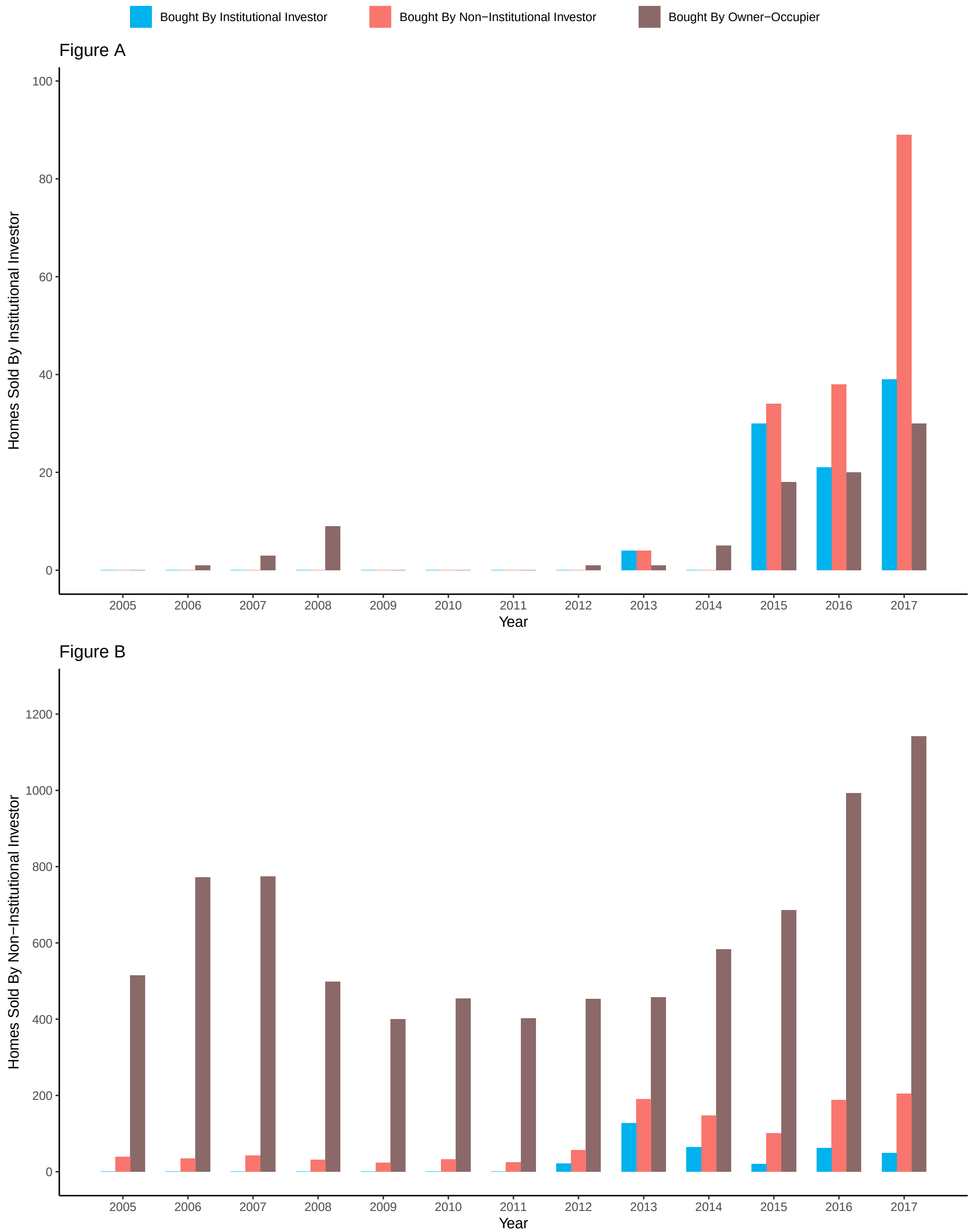
This figure shows the percentage of homes bought by each purchaser by sales price since 2012. Each bar represents a sale price range of $10,000. Investors are defined as non−individuals that are not banks, mortgage/credit lenders, relocation companies, building companies, nor government entities. Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

Figure 3.5: Homes Purchased by Investors By Sale Price



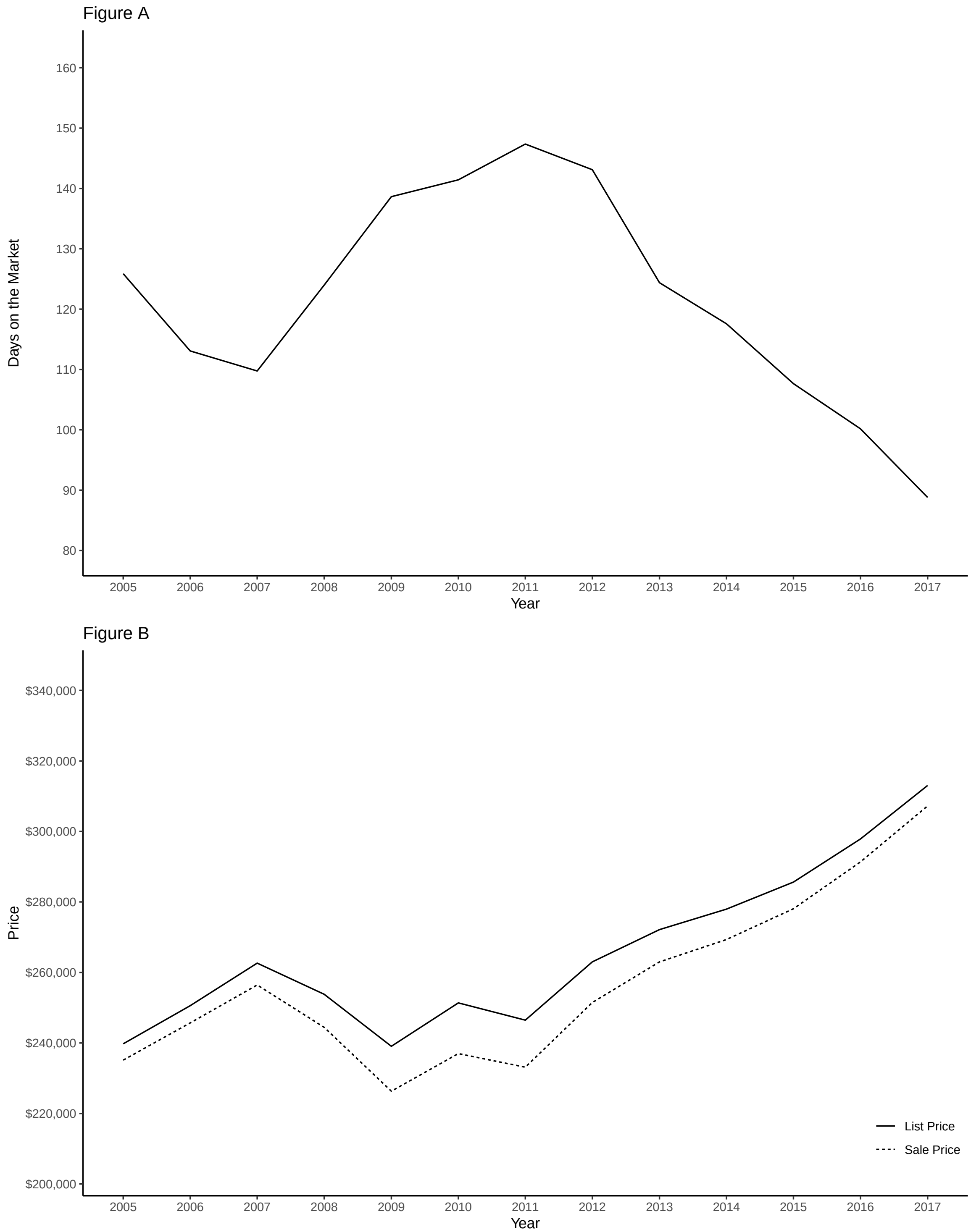
This figure shows the total number of single−family home sales by investors and the percentage of homes sold by investors. Investors are defined as non−individuals that are not banks, mortgage/credit lenders, relocation companies, building companies, nor government entities. Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

Figure 3.6: Homes Sold by Investors



This figure shows the total number of single−family homes that were sold by institutional investors and the total number of single−family homes that were sold by non−institutional investors. The total number is broken down by the group that subsequently bought the home. These groups are: institutional investors, non−institutional investors, and non−investors. Investors are defined as non−individuals that are not banks, mortgage/credit lenders, relocation companies, building companies, nor government entities. Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

Figure 3.7: To Whom Did Investors Sell Homes?



This figure shows the average days on the market, the average listing price, and the average sales price for single−family homes in the Charlotte region. The data used was the matched data between Metrostudy and MLS. This match is a unique match between Metrostudy and MLS by sale date and address. Only observations found in both data sets were kept.

Figure 3.8: Days on the Market, List Price, and Sale Price by Year

Table 3.1: Summary Statistics

Variables	Mean	Median	Min	Max	Std. Dev
Total Observations	336,878				
Has Mortgage	82.55%				
Purchased By Investor	14,850				
Type of Sale					
New	82,321				
Regular Resale	230,655				
REO Sale	23,902				
Sale Characteristics					
Sale Price	$256,231.98	$195,000.00	$30,000.00	$8,350,000.00	$223,763.99
Age	18.67	11.00	0.00	238.00	22.04
Bathrooms	2.40	2.50	1.00	10.50	0.84
Bedrooms	3.48	3.00	1.00	20.00	0.81
Lot Acres	0.49	0.28	0.01	20.15	0.95
Sqft Finished	2,343.47	2,118.00	336.00	15,391.00	1,068.77
Loan Characteristics					
Loan Amount	$221,303.80	$179,900.00	$5,000.00	$5,000,000.00	$158,639.71

Note:
Data set covers Cabarrus, Gaston, Iredell, Lincoln, Mecklenburg, and Union county from 2005-2017. Investors are defined as non-individuals that are not banks, mortgage/credit lenders, relocation companies, building companies, nor government entities. Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

Table 3.2: Investor Purchases vs Non-Investor Purchases

	Investor					Non-Investor				
Variables	Mean	Median	Min	Max	Std. Dev	Mean	Median	Min	Max	Std. Dev
Total Observations	14,850					322,028				
Has Mortgage	12.42%					85.78%				
Type of Sale										
New	1.65%					25.49%				
Regular Resale	81.46%					67.87%				
REO Sale	16.89%					6.64%				
Sale Characteristics										
Sale Price	$167,710.56	$140,000.00	$40,500.00	$8,350,000.00	$210,156.86	$260,314.06	$199,500.00	$30,000.00	$7,836,833.00	$223,527.90
Age	25.68	15.00	0.00	164.00	23.14	18.34	10.00	0.00	238.00	21.94
Bathrooms	2.17	2.50	1.00	9.00	0.73	2.41	2.50	1.00	10.50	0.85
Bedrooms	3.25	3.00	1.00	10.00	0.70	3.49	3.00	1.00	20.00	0.81
Lot Acres	0.36	0.23	0.01	19.56	0.79	0.50	0.29	0.01	20.15	0.95
Sqft Finished	1,909.81	1,757.50	480.00	11,977.00	819.90	2,363.47	2,140.00	336.00	15,391.00	1,074.65
Loan Characteristics										
LTV	11.99%	0.00%	0.00%	198.70%	32.90%	75.80%	90.00%	0.00%	200.00%	34.16%
Loan Amount	$19,381.16	$0.00	$0.00	$5,000,000.00	$102,729.27	$190,208.84	$161,413.00	$0.00	$4,000,000.00	$165,349.56

Note:
Data covers Cabarrus, Gaston, Iredell, Lincoln, Mecklenburg, and Union county from 2005-2017. Data used is the full MetroStudy data. All cash sales were kept. The summary statistics in the table are for all home sale observations. These calculations include both observations with a mortgage and observations with cash. Investors are defined as non-individuals that are not banks, mortgage/credit lenders, relocation companies, building companies, nor government entities.

Table 3.3: Institutional Investor Purchases vs Non-Institutional Investor Purchases

	Institutional Investor					Non-Institutional Investor				
Variables	Mean	Median	Min	Max	Std. Dev	Mean	Median	Min	Max	Std. Dev
Total Observations	7,365					7,485				
Has Mortgage	0.00%					24.65%				
Type of Sale										
New	1.94%					1.36%				
Regular Resale	86.83%					76.18%				
REO Sale	11.23%					22.46%				
Sale Characteristics										
Sale Price	$171,191.88	$164,000.00	$40,500.00	$2,400,500.00	$69,007.95	$164,285.06	$96,500.00	$40,500.00	$8,350,000.00	$287,958.59
Age	13.85	12.00	0.00	116.00	10.46	37.33	36.00	0.00	164.00	26.10
Bathrooms	2.33	2.50	1.00	5.50	0.48	2.01	2.00	1.00	9.00	0.88
Bedrooms	3.40	3.00	1.00	8.00	0.63	3.11	3.00	1.00	10.00	0.74
Lot Acres	0.22	0.19	0.04	4.64	0.14	0.50	0.27	0.01	19.56	1.09
Sqft Finished	2,059.06	1,986.00	675.00	5,547.00	602.72	1,762.96	1,507.00	480.00	11,977.00	965.85
Loan Characteristics										
LTV	0.00%	0.00%	0.00%	0.00%	0.00%	23.78%	0.00%	0.00%	198.70%	43.20%
Loan Amount	$0.00	$0.00	$0.00	$0.00	$0.00	$38,451.61	$0.00	$0.00	$5,000,000.00	$142,145.75

Note:
Data covers Cabarrus, Gaston, Iredell, Lincoln, Mecklenburg, and Union county from 2005-2017. Data used is the full MetroStudy data. All cash sales were kept. The summary statistics in the table are for all home sale observations. These calculations include both observations with a mortgage and observations with cash. Investors are defined as non-individuals that are not banks, mortgage/credit lenders, relocation companies, building companies, nor government entities. Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

Table 3.4: Hedonic Results: Quantile Regression

	Baseline			$\tau = 0.25$			$\tau = 0.50$			$\tau = 0.75$		
Variable	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Panel A: Without Recent Investor Activity Dummy												
Purchased By Investor	-0.195***	-0.095***		-0.200***	-0.103***		-0.160***	-0.114***		-0.142***	-0.130***	
	(-72.908)	(-33.320)		(-46.766)	(-18.269)		(-56.343)	(-34.178)		(-67.075)	(-45.472)	
Purchased By Non-Institutional Investor			-0.293***			-0.352***			-0.309***			-0.232***
			(-80.358)			(-59.232)			(-57.309)			(-39.539)
Purchased By Institutional Investor			-0.114***			-0.081***			-0.097***			-0.112***
			(-30.779)			(-25.206)			(-39.160)			(-53.205)
No Mortgage		-0.143***			-0.144***			-0.058***			-0.015***	
		(-93.454)			(-48.476)			(-28.619)			(-9.464)	
Intercept	10.959***	10.987***	10.974***	10.758***	10.788***	10.781***	11.064***	11.076***	11.057***	11.153***	11.155***	11.157***
	(353.364)	(358.791)	(358.113)	(111.433)	(126.810)	(113.102)	(216.870)	(230.214)	(221.806)	(297.820)	(293.506)	(283.296)
R^2	0.767	0.773	0.773	0.542	0.548	0.549	0.582	0.583	0.584	0.608	0.608	0.608
N	336,878	336,878	336,878	336,878	336,878	336,878	336,878	336,878	336,878	336,878	336,878	336,878
Panel B: With Recent Investor Activity Dummy												
Purchased By Investor	-0.193***	-0.091***		-0.195***	-0.096***		-0.158***	-0.110***		-0.141***	-0.128***	
	(-71.337)	(-31.398)		(-61.779)	(-18.774)		(-89.917)	(-101.694)		(-106.161)	(-71.106)	
Purchased By Non-Institutional Investor			-0.291***			-0.350***			-0.308***			-0.233***
			(-78.627)			(-65.573)			(-76.828)			(-69.865)
Purchased By Institutional Investor			-0.114***			-0.081***			-0.098***			-0.112***
			(-30.762)			(-26.137)			(-37.761)			(-49.252)
No Mortgage		-0.146***			-0.148***			-0.060***			-0.016***	
		(-92.745)			(-40.849)			(-40.832)			(-10.077)	
Recent Investor Activity	0.006***	0.006***	0.006***	0.004***	0.004***	0.003***	0.005***	0.005***	0.005***	0.006***	0.006***	0.006***
	(9.896)	(10.327)	(9.326)	(5.646)	(5.405)	(4.360)	(10.305)	(10.838)	(9.256)	(11.407)	(11.710)	(10.905)
Intercept	11.058***	11.087***	11.069***	10.950***	10.931***	10.975***	11.124***	11.130***	11.123***	11.192***	11.208***	11.221***
	(348.368)	(353.919)	(353.000)	(108.801)	(149.634)	(142.700)	(212.777)	(242.498)	(219.482)	(303.242)	(323.216)	(285.251)
R^2	0.768	0.775	0.774	0.544	0.550	0.550	0.583	0.584	0.585	0.607	0.607	0.608
N	319,423	319,423	319,423	319,423	319,423	319,423	319,423	319,423	319,423	319,423	319,423	319,423

Note:
*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] The log sale price is the independent variable in each regression. Each regression also includes housing characteristics, transaction type, zip code fixed effects, and quarter-year fixed effects. t-statistics are in parenthesis.

[b] Baseline is the hedonic regression using ordinary least squares. τ represents the τth quantile of the log sales price.

[c] Purchased By Investor is a dummy variable that equals 1 if the home is purchased by a non-individual that is not a bank, mortgage/credit lender, relocation company, building company, nor government entity and 0 otherwise. Purchased By Non-Institutional Investor is a dummy variable that equals 1 if the home is purchased by a non-institutional investor and 0 otherwise. Purchased By Institutional Investor is a dummy variable that equals 1 if the home is purchased by an investor that has filed as a publicly traded company, that has filed as a REIT with the SEC, or that has filed a Form D with the SEC and 0 otherwise.

[d] Recent Investor Activity is the percentage of homes purchased by institutional investors in the previous 6 months within each county. For example, if a home was bought in August of 2013 in Mecklenburg county, then recent investor activity would be the percentage of homes purchased by institutional investors in Mecklenburg county from January to July of 2013 relative to the total number of homes purchased in Mecklenburg county in that same time frame.

Table 3.5: Hedonic Results: By Area Subset

	Income				Minority Population	
Variable	< $35,100	$35,100-$44,600	$44,600-$58,000	$58,000 <	Low Minority Population	High Minority Population
Purchased By Non-Institutional Investor	-0.284***	-0.265***	-0.264***	-0.252***	-0.251***	-0.302***
	(-31.686)	(-35.019)	(-37.977)	(-42.158)	(-49.676)	(-61.365)
Purchased By Institutional Investor	-0.184***	-0.110***	-0.097***	-0.121***	-0.131***	-0.063***
	(-7.233)	(-9.268)	(-14.672)	(-26.969)	(-28.067)	(-11.307)
Recent Investor Activity	-0.002	0.010**	0.004	0.004**	0.005**	0.005**
	(-0.197)	(2.261)	(1.278)	(2.199)	(2.104)	(2.133)
Intercept	10.888***	10.899***	11.222***	9.976***	11.092***	10.883***
	(139.428)	(202.242)	(324.547)	(34.116)	(351.668)	(148.293)
R^2	0.607	0.656	0.678	0.749	0.764	0.665
N	22,165	35,215	68,713	193,241	244,669	74,665

Note:

*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] The log sale price is the independent variable in each regression. Each regression also includes housing characteristics, transaction type, zip code fixed effects, and quarter-year fixed effects. t-statistics are in parenthesis.

[b] The income groups are based on census tract median incomes split into quartiles. The median income comes from the ACS 5-Year estimates between 2007-2011. The minority population groups are based on census tracts with a minority population greater than 50% and census tracts with a minority population less than or equal to 50%. The minority population comes from the ACS 5-Year estimates between 2007-2011.

[c] Purchased By Investor is a dummy variable that equals 1 if the home is purchased by a non-individual that is not a bank, mortgage/credit lender, relocation company, building company, nor government entity and 0 otherwise. Purchased By Non-Institutional Investor is a dummy variable that equals 1 if the home is purchased by a non-institutional investor and 0 otherwise. Purchased By Institutional Investor is a dummy variable that equals 1 if the home is purchased by an investor that has filed as a publicly traded company, that has filed as a REIT with the SEC, or that has filed a Form D with the SEC and 0 otherwise.

[d] Recent Investor Activity is the percentage of homes purchased by institutional investors in the previous 6 months within each area. For example, if a home was bought in August of 2013 in a High Minority Population, then recent investor activity would be the percentage of homes purchased by institutional investors from January of 2013 to July of 2013 in census tracts with a minority population greater than 50% relative to the total number of homes purchased in that same time frame in census tracts with a minority population greater than 50%.

Table 3.6: Metrostudy vs Metrostudy+MLS

Variables	MetroStudy	MetroStudy+MLS
Total Observations	336,878	244,453
Has Mortgage	82.55%	82.62%
Purchased By Investor	14,850	10,694
Type of Sale		
New	24.44%	16.65%
Regular Resale	68.47%	73.70%
REO Sale	7.10%	9.64%
Sale Characteristics		
DOM		117.58
List Price		$268,707.65
Sale Price	$256,231.98	$260,723.86
Age	18.67	19.67
Bathrooms	2.40	2.42
Bedrooms	3.48	3.49
Lot Acres	0.49	0.50
Sqft Finished	2,343.47	2,325.89
Loan Amount	$221,303.80	$226,970.21

Note:

Each data set covers Cabarrus, Gaston, Iredell, Lincoln, Mecklenburg, and Union county from 2005-2017. "MetroStudy+MLS" is the unique merge between MetroStudy and MLS. Only observations found in both data sets were kept. Investors are defined as non-individuals that are not banks, mortgage/credit lenders, relocation companies, building companies, nor government entities. Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

Table 3.7: Logit Results: With List Price

	Model 1		Model 2		Model 3	
Variable	Estimate	Marginal Effect	Estimate	Marginal Effect	Estimate	Marginal Effect
Purchased By Investor	-1.149***	-0.188	-0.791***	-0.129		
	(-35.793)		(-22.878)			
Purchased By Non-Institutional Investor					-0.553***	-0.081
					(-12.459)	
Purchased By Institutional Investor					-1.624***	-0.175
					(-34.881)	
No Mortgage			-0.478***	-0.078		
			(-28.179)			
Recent Investor Activity	0.029***	0.005	0.029***	0.005	0.031***	0.005
	(4.783)		(4.793)		(5.075)	
Intercept	-1.194**		-1.132**		-1.169**	
	(-2.476)		(-2.339)		(-2.423)	
Pseudo-R^2	0.066		0.070		0.068	
N	233,694		233,694		233,694	

Note:
*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] The probability of a home selling for more than the list price is estimated in each logit regression. Each regression also includes housing characteristics, transaction type, zip code fixed effects, and quarter-year fixed effects. z-statistics are in parenthesis.

[b] Purchased By Investor is a dummy variable that equals 1 if the home is purchased by a non-individual that is not a bank, mortgage/credit lender, relocation company, building company, nor government entity and 0 otherwise. Purchased By Non-Institutional Investor is a dummy variable that equals 1 if the home is purchased by a non-institutional investor and 0 otherwise. Purchased By Institutional Investor is a dummy variable that equals 1 if the home is purchased by an investor that has filed as a publicly traded company, that has filed as a REIT with the SEC, or that has filed a Form D with the SEC and 0 otherwise.

[c] Recent Investor Activity is the percentage of homes purchased by institutional investors in the previous 6 months within each county. For example, if a home was bought in August of 2013 in Mecklenburg county, then recent investor activity would be the percentage of homes purchased by institutional investors in Mecklenburg county from January to July of 2013 relative to the total number of homes purchased in Mecklenburg county in that same time frame.

Table 3.8: Logit Results: With Estimated Home Price

	Model 1		Model 2		Model 3	
Variable	Estimate	Marginal Effect	Estimate	Marginal Effect	Estimate	Marginal Effect
Purchased By Investor	-1.678***	-0.324	-1.644***	-0.317		
	(-39.769)		(-37.385)			
Purchased By Non-Institutional Investor					-1.458***	-0.207
					(-22.932)	
Purchased By Institutional Investor					-1.832***	-0.235
					(-32.517)	
No Mortgage			-0.043***	-0.008		
			(-2.781)			
Recent Investor Activity	0.038***	0.007	0.038***	0.007	0.038***	0.007
	(7.030)		(7.032)		(7.097)	
Intercept	-0.718**		-0.711**		-0.712**	
	(-2.262)		(-2.239)		(-2.245)	
Pseudo-R^2	0.067		0.067		0.067	
N	217,512		217,512		217,512	

Note:
*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] The probability of a home selling for more than the estimated home price is estimated in each logit regression. The estimated home price of home i at time t is an imputed price from a 6-month rolling hedonic regression from time *t-6* to time *t-1.* Each logit regression also includes housing characteristics, transaction type, zip code fixed effects, and quarter-year fixed effects. z-statistics are in parenthesis.

[b] Purchased By Investor is a dummy variable that equals 1 if the home is purchased by a non-individual that is not a bank, mortgage/credit lender, relocation company, building company, nor government entity and 0 otherwise. Purchased By Non-Institutional Investor is a dummy variable that equals 1 if the home is purchased by a non-institutional investor and 0 otherwise. Purchased By Institutional Investor is a dummy variable that equals 1 if the home is purchased by an investor that has filed as a publicly traded company, that has filed as a REIT with the SEC, or that has filed a Form D with the SEC and 0 otherwise.

[c] Recent Investor Activity is the percentage of homes purchased by institutional investors in the previous 6 months within each county. For example, if a home was bought in August of 2013 in Mecklenburg county, then recent investor activity would be the percentage of homes purchased by institutional investors in Mecklenburg county from January to July of 2013 relative to the total number of homes purchased in Mecklenburg county in that same time frame.

Table 3.9: Propensity Score Matched Samples

	Matched With All Individuals		Matched With Individuals With Mortgage		Matched With Individuals Without Mortgage	
Variable	Institutional Investor	Owner-Occupier	Institutional Investor	Owner-Occupier	Institutional Investor	Owner-Occupier
Total	7,365	7,365	7,365	7,365	4,121	4,121
Age	13.85	13.74	13.85	13.81	14.61	19.38
Bathrooms	2.33	2.32	2.33	2.32	2.27	2.21
Bedrooms	3.40	3.38	3.40	3.40	3.42	3.34
Lot Acres	0.22	0.22	0.22	0.23	0.24	0.29
Sale Price	$171,191.88	$192,391.62	$171,191.88	$194,766.73	$175,347.54	$199,204.75
Sqft Finished	2,059.06	2,046.87	2,059.06	2,054.61	2,062.24	2,044.04

[a] Each institutional investor observation is matched with a home purchased by an individual by year by zip code. The match was obtained using Propensity Score Matching with a nearest neighbor matching algorithm. Three matches were made: institutional investor home matched with any owner-occupier home, institutional investor home matched with an owner-occupier home with a mortgage, and institutional investor home matched with an owner-occupier home without a mortgage.

[b] Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

Table 3.10: Hedonic Results: With Propensity Score Matching

Variable	Matched With All Individuals	Matched With Individuals With Mortgage	Matched With Individuals Without Mortgage
Panel A: Without Recent Investor Activity Dummy			
Purchased By Institutional Investor	-0.078***	-0.102***	-0.016*
	(-14.996)	(-20.265)	(-1.823)
Intercept	11.811***	11.735***	11.575***
	(52.978)	(54.885)	(28.653)
R^2	0.436	0.444	0.425
F-Statistic	128.547	133.007	69.410
N	14,730	14,730	8,242
Panel B: With Recent Investor Activity Dummy			
Purchased By Institutional Investor	-0.079***	-0.102***	-0.016*
	(-15.047)	(-20.239)	(-1.754)
Recent Investor Activity	0.002	0.004	-0.004
	(0.587)	(1.526)	(-0.899)
Intercept	11.815***	11.744***	11.567***
	(52.972)	(54.909)	(28.632)
R^2	0.436	0.444	0.426
F-Statistic	126.983	131.514	68.865
N	14,730	14,730	8,242

Note:
*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] Each institutional investor observation is matched with a home purchased by an individual by year by zip code. The match was obtained using Propensity Score Matching with a nearest neighbor matching algorithm. Three matches were made: institutional investor home matched with any owner-occupier home, institutional investor home matched with an owner-occupier home with a mortgage, and institutional investor home matched with an owner-occupier home without a mortgage.

[b] The log sale price is the independent variable in each regression. Each regression also includes zip code fixed effects and quarter-year fixed effects. t-statistics are in parenthesis.

[c] Purchased By Institutional Investor is a dummy variable that equals 1 if the home is purchased by an investor that has filed as a publicly traded company, that has filed as a REIT with the SEC, or that has filed a Form D with the SEC and 0 if purchased by an individual.

[d] Recent Investor Activity is the percentage of homes purchased by institutional investors in the previous 6 months within each county. For example, if a home was bought in August of 2013 in Mecklenburg county, then recent investor activity would be the percentage of homes purchased by institutional investors in Mecklenburg county from January to July of 2013 relative to the total number of homes purchased in Mecklenburg county in that same time frame.

Table 3.11: Spillover Effect on Owner-Occupier Home Purchases

	Within 1-Year Window		Within 2-Year Window	
Variable	Model 1	Model 2	Model 1	Model 2
$\delta_{Close,Before} - \delta_{Close,After}$	0.002	-0.002	0.005***	0.000
	(0.916)	(-0.850)	(3.161)	(0.151)
$\delta_{Far,Before} - \delta_{Far,After}$		0.003**		0.003***
		(2.340)		(3.921)
R^2	0.608	0.608	0.609	0.609
F-Statistic	47,049.686	39,224.878	47,080.090	39,255.989
N	308,449	308,449	308,449	308,449

Note:
*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] The log sale price is the independent variable in each regression. Only homes purchased by owner-occupiers are included within the regression. Each regression also includes housing characteristics, transaction type, and census tract-year fixed effects. Standard errors are clustered at the census tract-year level. t-statistics are in parenthesis.

[b] Close represents the number of homes purchased by institutional investors within 0.1 miles of an owner-occupier home one/two year before and one/two year after the sale. $\delta_{Close,Before} - \delta_{Close,After}$ represents the effect of an additional home purchased by an institutional investor within 0.1 miles before the owner-occupier home sale. Far represents the number of homes purchased by institutional investors within 0.25 miles of an owner-occupier home one/two year before and one/two year after the sale. $\delta_{Far,Before} - \delta_{Far,After}$ represents the effect of an additional home purchased by an institutional investor within 0.25 miles before the owner-occupier home sale.

[c] Institutional investors are investors that have filed as a publicly traded company, that have filed as a REIT with the SEC, or that have filed a Form D with the SEC.

APPENDIX

Imputed Race

As a robustness check for the hedonic results by area in Table 3.5, we include in our main specification the race/ethnicity of the purchaser. We do this because our results by area may miss on some of the heterogeneity within the area. Since the race/ethnicity of the purchaser is not originally included within our dataset, we must impute this value. We use the methodology laid out by Elliott et al. (2009) and subsequently used by the Consumer Financial Protection Bureau (2014) and Diamond et al. (2019). This methodology uses Bayes Theorem to impute the race of the individual based on the individual's surname and census tract location.

Specifically, for each individual purchaser with surname n in census tract ct, we estimate the probability of race r for each of the six categories defined by the Census Bureau. These six categories are: White, Black, Asian and Native Hawaiian and Other Pacific Islander, American Indian and Alaska Native, Two or More Races, and Hispanic or Latino Origin. We use the Summary File 1 from the 2010 Census to identify the percentage of the population of race/ethnicity r in a given census tract. Also from the 2010 Census, we obtain the frequency of each surname by race/ethnicity r. Using Bayes Theorem, we can obtain the probability of race/ethnicity r given the surname n and census tract ct:

$$Pr(r|ct, n) = \frac{Pr(r|n)Pr(ct|r)}{\sum_{r' \in R} Pr(r'|n)Pr(ct|r')} \tag{3.4}$$

where R is the set of six race/ethnic categories. The main assumption that we must impose is that the probability of living in a specific location, given one's race, is independent of the individual's surname. Similar to Diamond et al. (2019), we assign a final probability of race/ethnicity to be the maximum posterior probability if that probability is greater than 80%. Otherwise, the individual's race is left unclassified.

In addition, to ensure that our results are not driven by this imputation process, we merge our full Metrostudy data with the Home Mortgage Disclosure Act (HMDA) data. HMDA data is collected by the U.S. government and is the most comprehensive data set on mortgage originations. The most important aspect of the HMDA data is that it includes the applicant's race and ethnicity. Following Bayer et al. (2018) and Diamond and McQuade (2019), we merge our Metrostudy data with HMDA based on the sale year, loan amount, lender name, state, county, and census tract. This procedure results in a high level match between the two data sets.

We run our main specification (Equation (3.1)) on the imputed data set as well as the merged data set. These results can be seen in Table 3.12 and are consistent with the results in Table 3.5. That is, institutional investors still purchase at a discount regardless of race/ethnicity, and a one percentage point increase in institutional investors leads to about a 0.26%-0.62% increase in the price owner-occupiers pay for a single-family home.

Table 3.12: Hedonic Results: By Race Subset

Variable	Minorities with Mortgage	All Minorities	Non-Minorities with Mortgage	All Non-Minorities
Panel A: Full Metrostudy Data				
Purchased By Institutional Investor	-0.064***	-0.037***	-0.183***	-0.173***
	(-17.937)	(-9.806)	(-53.013)	(-45.863)
Recent Investor Activity	0.005***	0.006***	0.005***	0.004***
	(3.142)	(3.848)	(7.442)	(6.139)
Intercept	11.038***	11.040***	11.034***	11.010***
	(86.499)	(79.351)	(324.513)	(322.086)
R^2	0.829	0.816	0.817	0.793
N	35,904	40,032	143,877	163,456
Panel B: Merged Metrostudy and HMDA Data				
Purchased By Institutional Investor	-0.085***		-0.178***	
	(-27.187)		(-56.097)	
Recent Investor Activity	0.004***		0.003***	
	(3.055)		(3.587)	
Intercept	11.339***		11.007***	
	(77.002)		(294.917)	
R^2	0.822		0.825	
N	39,656		107,257	

Note:

*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] The log sale price is the independent variable in each regression. Each regression also includes housing characteristics, transaction type, zip code fixed effects, and quarter-year fixed effects. t-statistics are in parenthesis.

[b] Minorities uses the 2010 Census demographic definitions and includes Blacks, Asians and Native Hawaiians and Other Pacific Islanders, American Indians and Alaska Natives, Two or More Races, and Hispanic or Latino Origins. The top panel uses Bayes Theorem to impute the race of a purchaser given the individual's surname and census tract location. The bottom panel uses the race/ethnicity in the merged Metrostudy-HMDA data.

[c] Purchased By Institutional Investor is a dummy variable that equals 1 if the home is purchased by an investor that has filed as a publicly traded company, that has filed as a REIT with the SEC, or that has filed a Form D with the SEC and 0 otherwise.

[d] Recent Investor Activity is the percentage of homes purchased by institutional investors in the previous 6 months within each county. For example, if a home was bought in August of 2013 in Mecklenburg county, then recent investor activity would be the percentage of homes purchased by institutional investors in Mecklenburg county from January to July of 2013 relative to the total number of homes purchased in Mecklenburg county in that same time frame.

CHAPTER 4: THE IMPACT OF INCREASED DEVELOPMENT FEES ON EXISTING HOUSING: EVIDENCE FROM SOUTH CAROLINA

4.1 Introduction

Since the Great Recession, some local areas have seen higher economic and population growth than others and as a result, must start thinking about infrastructure needs to accompany these growths. The main issue is how do these local governments finance the new infrastructure? In this paper, we investigate one such method: development fees, which are one time, per-unit fees paid by the developer to build housing and are used for a pre-determined purpose such as parks and recreation, municipal equipment, transportation, fire, or schools. We are concerned about the impact that these fees have on the housing market. Using the significantly large development fee increase in York county, South Carolina as our quasi-natural experiment, we empirically find that the increased development fees correspond with a decrease in asking and closing prices and an increase in available inventory in the period immediately after the increase.

Development fees are an increasingly popular choice among local governments because they can pass the fee directly to the developer instead of passing a fee directly to the residents, such as what happens with a property tax. Thus, development fees may be easier to implement because they do not put an immediate monetary cost on existing homeowners in order to pay for any new development. Instead, these fees put the monetary cost on the developers and in turn, these fees will hopefully be enough to pay for the new infrastructure. The implication then is that this supply side regulation will decrease the amount of new construction and increase the price of housing, which a large body of literature has shown.[1] This seems to be a suitable outcome because in a Fort Mill town hall meeting about the proposed increase of the development fees, one individual responded with:

> "That's gonna just raise the property value of all those houses as they get built, and so it's actually gonna raise my property value, which is a good thing for me," Knol said. "If it slows the growth a little, so we can catch up, like with the roads, that's a win too, I think."

The individual's comments are in-line with previous evidence on supply side regulations and development fees. Due to the fact that developers must pay more to build a single unit of housing, the final price of the *new* home must also be higher in order to break even. In addition, developers may also build fewer units due to the increased cost of construction, which will slow the growth in the area. However, it is not clear whether *existing* property values will also increase. That is, it is not clear whether the decrease in new housing supply will increase the property value of existing homes. In fact, since growth in the area will slow down as a result of the development fees, there is uncertainty whether the increased development fees will cover the costs of the new amenities. If the development fee is set too low or too high and the amenities have not been completed, then existing housing must be taxed. This means that the expected tax burden on existing homeowners increases, which might lower prices of existing homes today.

This view, however, is inconsistent with the previous literature. There is a small, but consistent view in the literature that suggests that prices of *all* homes will increase.[2] Laid out in Ihlanfeldt and Shaughnessy (2004), there are currently two views

[1]See Mayer and Somerville (2000), Quigley and Raphael (2005), and Ihlanfeldt (2007) to name a few.

[2]Delaney and Smith (1989a), Delaney and Smith (1989b), Singell and Lillydahl (1990), and Ihlanfeldt and Shaughnessy (2004) all empirically examine the effect development fees have on existing and new homes. Each of those studies find that prices on both existing and new homes increased after the fees were put in place.

on development fee's effect on home prices: the old and new view. The old view suggests that development fees increase the cost of construction, reduce the quantity of new construction, reduce the value of undeveloped land, and increase the price of existing housing. The new view suggests that development fees reduce the price of new housing because of lower quality but this is offset by new amenities (thus, no net impact on new housing prices) and increase the price of existing homes because of amenity effects. Ultimately, the new view incorporates the new amenities into home values and posits that existing home prices should increase. However, we argue that if the amenities are important and the development fee is set too low or possible growth is exhausted, then the only way to fund the completion of the amenities is to tax existing homeowners. This lowers existing home prices today and is exactly what we find empirically.

Unlike previous literature that largely looks at different cities and whether they have development fees or not, we have a unique opportunity to exploit how a dramatic *change* in the development fee can impact the housing market. Specifically, we examine the *change* in an already existing development fee within a single school district within a single county, which provides us with a nice framework for a quasi-natural experiment. In 1996, the Fort Mill school district in York county, South Carolina implemented a $2,500 per-unit development fee to construct new homes. A study for the district was conducted and proposed that "the district would need 1.68 new elementary schools, 0.86 new middle schools, and 0.66 new high schools by the year 2020." Between 1996 and 2019, the Fort Mill school district constructed 7 elementary schools, 5 middle schools, and 2 high schools. Fees went untouched in Fort Mill until 2015 when the Fort Mill town council implemented additional development fees of $1,822.39 on August 24, 2015. These fees were collected beginning on October 1, 2015.

As the area saw extensive growth in 2017, the Fort Mill school district conducted a study that suggested that development fees for single family and multi-family housing increase to $18,158 and $12,020 respectively. On June 27, 2018, the York county council passed the proposed development fee increase, and on July 16, 2018, the York county council instituted the newly proposed development fee. Any building permit obtained before July 16, 2018 still adhered to the previous development fee of $2,500, but all new permits required payment of the increased development fee. Although housing developers sued the state of South Carolina on August 14, 2018 to stop the development fee increase, the development fee was still collected, and the proceeds were saved in an escrow pending the outcome of the lawsuit.

Using the dates from above, we are able to set up a difference-in-difference framework to test how *changes* in the development fees impact the single-family housing market. We find that after the first fee was passed and collected in 2015, home prices on existing homes decreased by 5.27%. However, home prices on new homes increased by 3.97%. This dynamic occurred again in 2018 after the passage of the new development fee. We find that home prices on existing homes decreased by 13.26%, while home prices on new homes increased by 7.18%.

To our knowledge, we are the first paper to find that a *change* in development fees decrease prices on existing homes and decrease the demand for existing homes. We believe a main driver of our results is the interaction between the level of the development fee and the amount of growth in the area. Fort Mill has grown a tremendous amount since 2010 relative to the other school districts within York county. According to the U.S. Census, Fort Mill has seen an average year-to-year increase in population of 6.91% between 2010-2018, whereas the other school districts within York county have seen an average of about 0.72%-4.22%. In addition, Fort Mill constructed 14 new schools between 1996-2019 when they were only expected to build three. Thus, when the 2015 development fee increase of $1,822.39 went into effect, the fee may have been viewed as too small to adequately cover the costs of the new amenities. Thus, current and future residents may have expected an increase in future tax burden to cover these amenity costs. However, when the new development fee went into place in 2018, the same price effects occurred. The most likely reason is that the significant increase in the development fee to $18,158 may have led consumers to think growth will significantly stall as home developers exit the Fort Mill housing market. This would reduce the incoming revenue for the local government so that in response, taxes must increase to pay for the amenities.

The rest of the paper is as follows: Section 4.2 discusses the data, Section 4.3 discusses the main methodology used, Section 4.4 discusses the results, and Section 4.5 includes the closing remarks.

4.2 Data

We use two main data sources to study the effects of development fees on the housing market: the MLS and Metrostudy. Both cover the single-family housing market in York county, South Carolina between 2010-2018. Both data sets contain the sale price, sale date, address, and other housing characteristics. Most importantly, the MLS contains information not only on sold homes but also on homes that listed and did not sell. Thus, we can observe the total number of homes for sale in the market at a given time. Furthermore, each observation in the MLS also contains the list price as well as the listing contract date, which we use to find the days on the market for each sold home.

One downside of the MLS data is that it only covers homes that were sold by realtors, so it may miss a considerable amount of homes that were sold directly by the homeowner as well as information on land sales. For these reasons, we complement our analysis with the Metrostudy data. Metrostudy is a leading provider in residential construction activity across the nation, but they also have data with transaction level history for single-family homes and land sales. This transaction level housing data comes from CoreLogic, a major provider of housing data for the U.S, as well as from deed and tax assessor data from individual counties. Using the combination of data between the MLS and Metrostudy, we can capture a fairly complete picture of the housing market.

Finally, given that our primary focus is on the price effects in the Fort Mill school district within York county, South Carolina, we need to identify which homes were sold in this district. There are four school districts within York county: York school district 1, Clover school district 2, Rock Hill school district 3, and Fort Mill school district 4. Identifying homes within each school district was an easy task since each observation in both the MLS and Metrostudy data provided us with the physical location of the home. Therefore, we classified homes as either within the Fort Mill school district or outside the Fort Mill school district. Any observation missing a physical address was dropped from our analysis. In addition, we removed any observations with outliers such as 20+ bathrooms and 10+ bedrooms, as well as any observations with missing attributes.

Table 4.1 shows the summary statistics for the MLS property sale data, the Metrostudy property sale data, and the Metrostudy land sale data. As expected, we can see that the biggest discrepancy between the MLS and Metrostudy data is the total number of observations. The MLS data set has 15,999 total observations, while the Metrostudy data set has 26,842 total observations. As a result, the percentage of new homes sold within the MLS data set is 17.43%, whereas the percentage of new homes sold within the Metrostudy data set is 13.80%. Similarly, the percentage of homes sold within the Fort Mill school district in the MLS data set is 55.92%, whereas the percentage of homes sold within the Fort Mill school district in the Metrostudy data set is 34.62%. However, these differences do not alter our main results.

Although the MLS data set has a significantly smaller number of observations, it does provide interesting detail on the listing activity within the housing market. From Table 4.1, the average list price of a sold home in the sample was $297,533.26, and the average number of days until a home sold was 123.74. Figure 4.1 shows these dynamics between 2010-2018: the total number of homes sold, the average sale price, and the average list price all increased slowly over time, while the average days on market decreased. This positive relation between total homes sold and prices and the inverse relation between prices and days on market has been well documented in the literature such as Stein (1995) and Genesove and Mayer (2001). Overall, Figure 4.1 suggests that the York county housing market as a whole has seen significant growth. These patterns are also similar when York county is split into areas within the Fort Mill school district and areas outside the Fort Mill school district, which can be seen in Figure 4.2. This is important because our empirical framework relies on the assumption that trends in the treatment group (within the Fort Mill school district) and control group (outside the Fort Mill school district) are similar before the development fee increase is introduced.

As mentioned previously, one of the main reasons for using the Metrostudy data is that Metrostudy has an adequate amount of data on both land sales and construction activity at the local level. The last column in Table 4.1 shows the summary statistics for the land data. There are 4,785 total land sales between 2010-2018 with 1,420 (29.68%) of the sales occurring within the Fort Mill school district. The average price of a land sale is $194,072.01. Figure 4.3 plots the total quarterly land sales and average quarterly land price by school district. During the sample period, land sales have steadily increased in areas outside of the Fort Mill school district, while land sales within the school district have been relatively flat. However, prices have shown the opposite: the average quarterly land sale price within the Fort Mill school district has increased since 2014, while the average quarterly land sale price outside the Fort Mill school district has remained relatively flat.

Figure 4.3A and Figure 4.3B show that more, relatively cheaper land is being purchased outside the Fort Mill school district. This seems to have slowed development, which is evident in Figure 4.3C, which shows the total quarterly number of lots submitted for development approval/zoning to the municipality.[3] Since 2014, all areas within York county have seen a significant rise in planned construction of future housing. However, as areas outside the Fort Mill school district continue to see an increase in planned construction of future housing, the Fort Mill school district has seen a relatively flat amount in planned construction of future housing. In fact, since the second half of 2016, the Fort Mill school district has seen the number of lots submitted for consideration decrease and saw a further decrease in 2018.

4.3 Methodology

The summary statistics laid out in Section 4.2 suggest that from 2010-2018 York county has seen significant growth in the housing market. However, future development in the housing market may be a concern, especially within the Fort Mill school district, as land sales and future housing inventory has either remained flat or decreased between 2014-2018. Since growth may have stalled and individuals are forward looking, it is possible that individuals have priced this slow down in growth into current home prices. Thus, the main empirical question of this study: how did the dramatic increase in development fees impact the housing market?

Unlike previous literature, we have the exact dates of when the development fee increases were passed and executed. Furthermore, the increased development fees only pertained to a single school district within a single county, giving us reason to consider the increase as a quasi-natural experiment. Using the two motives above, we use a difference-in-difference regression approach to analyze the effect of the increased development fee on the housing market.

As mentioned in the introduction, the Fort Mill school district passed two separate development fee increases: one on August 24, 2015 and one on June 27, 2018. Collection of the first fee began on October 1, 2015. Collection of the second fee began

[3]These data are collected by Metrostudy.

on July 16, 2018 and collection continued even after developers sued the state on August 14, 2018. Thus, with this design we have five separate events to analyze. To utilize our difference-in-difference approach, we set the treatment group as the Fort Mill school district and the control group as all other areas within York county, South Carolina but outside the Fort Mill school district.

Our main specification can be written as:

$$Dep_{i,t} = \beta_0 + \alpha_1 FMSD_{i,t} + \sum_{j=1}^{5} \delta_j IncreasedPeriod_{i,j,t} + \\ \sum_{j=1}^{5} \gamma_j FMSD_{i,t} * IncreasedPeriod_{i,j,t} + \beta \mathbf{X}_{i,t} + L_i + T_t + \epsilon_{i,t} \qquad (4.1)$$

where $Dep_{i,t}$ is the outcome of interest (such as the log sale price), $FMSD_{i,t}$ is a dummy variable if the observation is from the Fort Mill school district at time t, $IncreasedPeriod_{i,j,t}$ are five separate dummy variables if the observation is from one of the five periods of increased development fees, $\mathbf{X}_{i,t}$ are other relevant characteristics (such as number of bedrooms, number of bathrooms, size of lot (acres), finished square footage, and age of the home), and L_i and T_t are location and time fixed effects, respectively. For the majority of our analysis we use city, year, month, day, and day of the week fixed effects. In addition, $IncreasedPeriod_{i,j,t}$ are five separate dummy variables that cover the periods between 8-24-15 and 9-30-15 (the passing of the first fee), 10-1-15 and 6-27-18 (the collection of the first fee), 6-28-18 and 7-15-18 (the passing of the second fee), 7-16-18 and 8-12-18 (the collection of the second fee), and 8-13-18 and beyond (the lawsuit by the developers).

We reemphasize that the increase in development fees may have different effects on new and existing homes. While we agree with previous literature that prices of *new* homes will increase after the development fee increase, the price effects on *existing* homes are ambiguous. For this reason, in addition to the specification in Equation (4.1), we also run a difference-in-difference-in-difference test which includes whether the home was a new or existing home. This can be formally written as:

$$\begin{aligned} y_{i,t} = \beta_0 + \alpha_1 FMSD_{i,t} + \alpha_2 New_{i,t} + \sum_{j=1}^{5} \delta_j IncreasedPeriod_{i,j,t} + \\ \sum_{j=1}^{5} \theta_j New_{i,t} * IncreasedPeriod_{i,j,t} + \sum_{j=1}^{5} \gamma_j FMSD_{i,t} * IncreasedPeriod_{i,j,t} + \\ \sum_{j=1}^{5} \lambda_j FMSD_{i,t} * IncreasedPeriod_{i,j,t} * New_{i,t} + \\ \beta \mathbf{X}_{i,t} + L_i + T_t + \epsilon_{i,t} \end{aligned} \tag{4.2}$$

where $y_{i,t}$ is the log price of the home and $New_{i,t}$ is a dummy variable equal to one if the observation is a new home and zero otherwise. All other variables are the same as in Equation (4.1). The baseline case in Equation (4.2) is an existing home sale outside the Fort Mill school district before the introduction of the increased development fee.

Since developers must pay more to build a single unit of housing after a development fee increase, the final price of a *new* home must be higher in order for the developer to break even, so we expect positive values for the λ_j's. However, unlike previous literature, we expect negative values for the γ_j's because growth in the area may have decreased as a result of the development fees. This decrease in growth may increase the uncertainty on whether the new development fees will cover the costs of the proposed amenities. For this reason, the expected tax burden on existing homeowners might increase, which will lower prices of existing homes today.

4.4 Empirical Results

4.4.1 Impact on Home Sale and List Prices

Table 4.2 shows the main results for our difference-in-difference-in-difference test on home prices. In the first two columns, the dependent variable is the final sale price, and in the final column, the dependent variable is the list price. Furthermore, when the dependent variable is the final sale price, the first two columns of Table 4.2

uses the MLS and Metrostudy data, respectively.

Across both data sets, after the first fee was passed, existing home sale prices within the Fort Mill school district decreased. Specifically, existing home sale prices within the Fort Mill school district decreased by 7.19%. This decrease in the final sale price of existing homes may come from the fact that list prices of existing sold homes also decreased by 5.63%. This dynamic continues even after collection of the first fee began. That is, after collection of the first fee began, existing home sales decreased by 5.27%, while list prices of existing sold homes decreased by 4.35%. On the other hand, prices on new homes reacted in the opposite direction. After the first fee was passed, final sales prices on new homes increased by 11.6%, while list prices on new homes increased by 10.32%. However, these results are not statistically significant. Nonetheless, once collection of the first fee began, final sales prices on new homes increased by 3.97%, while list prices on new homes increased by 5.07%. Both of these results are statistically significant.

Consistent with the previous literature, the results herein suggest that after the passage and collection of the development fee increase, prices on new homes increased. However, after this initial passage, development fees only increased by $1,822.39, which is a relatively low amount. As the Fort Mill school district was experiencing extreme growth, this small increase in the development fee may not have been enough to fund the new infrastructure. Current homeowners might have realized future tax burdens may increase and priced this expectation into existing homes. Thus, we see the negative impact on prices after the passage and collection of the initial development fee increase.

With that rationale in mind, one may suggest that a higher increase in development fees should then help fund the infrastructure and reduce any future increases in the tax burden of current homeowners. However, once the second development fee increase of $18,158 went into effect, the previous dynamics continue. That is, after collection of the second fee began, the final sale price of existing homes decreased by 13.26%, while the list price of existing homes decreased by 11%. New home prices showed similar behavior in which the final sale price increased by 7.18% and the list price increased by 6.92%. These results are both statistically and economically significant. To put the results in monetary value, after collection of the second fee increase began, home prices on existing homes decreased by $41,512, while home prices on new homes increased by $27,647.

While the first fee may have been viewed as too low by current homeowners, the second fee may have been viewed as too high. As a result, growth in the area may slow down, which would reduce the revenue to the Fort Mill school district and further reduce funding for future infrastructure. This would increase the future tax burden of current homeowners and lead to a strong negative impact of development fees on existing home prices.

4.4.2 Impact on Land Sales and Construction

The previous section presented the results of the development fee impact on home prices. The results suggest that future growth within the Fort Mill school district may be negatively impacted by the development fees. In this section, we analyze this idea. Specifically, we examine the impact of the development fee increase on land prices, total land sales, and future development inventory. We quantify future development inventory as the total quarterly number of lots submitted to the municipality for development approval/zoning. If our conjecture is accurate, we expect total land sales and total future development inventory to decrease after the *collection* of the increased development fees. The obvious reason being that it is more costly to develop new housing after the development fee increase is in effect.

However, the effect during the *passing* of the development fee is ambiguous. On one hand, the monetary cost has not yet increased for developers, which means that developers can purchase building permits at the pre-increase price level. This means that developer behavior may not change, and they will continue to purchase the same amount of land as before the passing of the fees. On the other hand, it is also possible that immediately after the passing of the development fee increase, developers may want to quickly purchase land so they can purchase the building permits before the increase goes into effect. One way to accomplish a quick sale is to pay a higher price.

Table 4.3 shows the results of this difference-in-difference test. After collection of both development fee increases began, the total number of weekly land sales decreased. After collection of the first development fee increase began, total land sales decreased by 2.27 lots per week, and after the collection on the second development fee increase, total land sales decreased by 4.64 lots per week. These results are both statistically significant and suggest that once the monetary effect was in place, developers chose not to purchase more land.

This reduction in land purchases directly influences the future inventory of new houses because as fewer land sales occur, fewer housing units can be built. This impact can be viewed in column 3 of Table 4.3: after collection of the first fee began, the total number of future inventory decreased by 93 units. The impact was even larger after the collection of the second fee began: the total number of future inventory decreased by 1,527 units. However, neither of these results are statistically significant. The main reason for the weak statistical significance is that the future inventory data is aggregated at a quarterly frequency and runs until the end of 2018. This means that we only have one full quarter of data immediately following the collection of the second development fee increase. As a result, we can not determine the long term impact of the development fee on construction growth.

Finally, Table 4.3 also shows the impact of development fee increases on land prices. As discussed above, developers may want to quickly purchase land so that they can purchase the building permits before the new development fees are put into effect. One way to facilitate a quick sale is to pay a higher price. Column 1 of

Table 4.3 shows exactly this. After the passing of the first development fee increase, land prices within the Fort Mill school district increased by 23.94%. Land prices within the Fort Mill school district further increased by 69.71% after the passing of the second development fee increase. Since there were essentially no changes in the total weekly land sales during these two periods, it is entirely possible that some developers purchased at higher prices in order to outbid other developers and speed up the development process. This increase in land prices during the *passing* stage may also be a main driver for the increase in final sale prices of new homes during the *collection* stage that we see in Table 4.2. Additionally, once collection on these fees began, there were no changes in land prices. These results suggest that the increase in development fees slowed future inventory while also making newly built homes more expensive through increased land prices and development cost.

4.4.3 Impact on Demand

The land price dynamics in the previous section shed light on the way prices of *new* homes may have been effected. That is, in order to start new development before collection of the fees began, developers needed to purchase land quickly, and they did so by purchasing land at higher prices. This increased the cost of development. However, the land dynamics do little to explain why home prices of *existing* homes decreased. Our main hypothesis for this is if the development fee is not optimally chosen and the new amenities are not fully funded by the development fees, then existing housing must be taxed to fulfill the shortfall. If this were the case, the expected tax burden on existing homeowners would increase, which would then lower prices of existing homes today.

To test this hypothesis, we examine the demand for housing. We believe that if current and future homeowners expect taxes to increase in the Fort Mill school district, then current homeowners may choose to put their home on the market while future homeowners may choose not to purchase in the Fort Mill school district. If this is the case, we should expect more homes to gradually be put up for sale but not be sold and homes should take longer to sell. In order to accomplish this, we utilize our general framework in Equation (4.1), which now uses new weekly sales, new weekly listings, cumulative weekly listings, and the days on the market as the dependent variables. If our hypothesis is correct, we should expect a positive impact on both the cumulative weekly listings and the days on the market after the passing and collecting of the new development fees.

Columns one and two of Table 4.4 show the impact of the development fees on new weekly sales and new weekly listings, respectively. There is no significant effect on new weekly sales after the new development fees were introduced. Furthermore, the introduction of the new development fees had a mixed impact on the total number of weekly listings. After the first development fee went into effect, the total number of weekly listings increased by 1.75. However, after the second development fee went into effect, the total number of weekly listings decreased by 4.01, although this effect is not statistically significant at conventional levels. As a result, it seems as though the development fee increase did not have an immediate impact on the sale patterns within the Fort Mill school district.

A further look into the homes for sale within York county suggests a different story. In a majority of cases, homes do not sell immediately after being listed for sale. In fact, within York county during the sample period, an average home was on the market for 123.74 days until it sold. For this reason, at any given time, there may be homes on the market that have been listed for many days but have yet to sell. This typically suggests a slow market (i.e. more sellers than buyers), which results in lower home prices on average. Over time, some homes will sell and be taken off the market and different homes will be placed on the market to be sold. If more homes are being put on the market and are not bought, then there are more sellers than buyers, which would explain the negative impact on existing home prices. Column three of Table

4.4 shows exactly this. After the increased development fees were passed and put into effect, the cumulative number of homes that were listed on the market but did not sell increased. Specifically, after the first development fee increase was passed, the total number of homes listed for sale increased by 1,820. After the second development fee increase was passed, the total number of homes listed for sale increased by 2,382. This suggests that the introduction of the new development fees increased the ratio of sellers to buyers. Regardless of whether the increase was due to more listings or due to fewer sales, the results suggest that fewer people wanted to be homeowners within the Fort Mill school district.

The final column of Table 4.4 further suggests a decrease in housing demand within the Fort Mill school district. As mentioned earlier, the market is considered slow if there are more sellers in the market than buyers. This will not only result in lower prices, but it will also result in longer times for a home to sell. The final column of Table 4.4 reports results that show once the development fee increases were passed and put into effect, the time-on-market increased for the average home within the Fort Mill school district. Specifically, after the first and second development fee increases were put into effect, the average time-on-market increased by 16.67% and 15.47%, respectively. Both of these results are statistically significant and suggest that homes within the Fort Mill school district took longer to sell after the increased development fees were introduced and put into effect.

4.5 Conclusion

Using a difference-in-difference approach, we find that the increased development fees in the Fort Mill school district correspond with a decrease in asking and closing prices and an increase in available inventory in the period immediately after the fee increase. Specifically, after the first fee was passed and collected in 2015, we find that home prices on existing homes decreased by 5.27%, while home prices on new homes increased by 3.97%. This dynamic occurred again in 2018 after the passage of a new development fee, which was considerably higher than the previous fee. We find that home prices on existing homes decreased by 13.26% and home prices on new homes increased by 7.18%. Furthermore, we find that after the second development fee increase was passed, the total number of homes listed for sale increased by 2,382. In addition, our results suggest that future housing development slowed down in response to the increased development fee.

Our results shed new light on how development fees may impact a local housing market and how these development fees might be selected. Unlike previous literature, we focus on how a dramatic *change* in development fees impact a local housing market rather than focusing on whether an area has a development fee or not. Given the amount of growth within the Fort Mill school district between 1996-2019, there would seem to be an optimal level for development fees such that they are neither too low nor too high. One indirect source of evidence is the somewhat arbitrary manner by which the new fees of $18,158 was selected by the Fort Mill county commission. The second source of evidence is that in the aftermath of passing and collecting the new dramatically higher development fees in 2018, there was lower future inventory, lower existing home prices, and lower demand for single-family housing within the Fort Mill school district. The main argument for this was that if possible growth is exhausted, then the only way to fund the completion of the amenities is to tax the existing homeowners. We believe that our framework gave us a unique opportunity to extract this feature and leaves an area for future work to examine how to optimally set development fees given an area's current and future growth.

However, it is also important to note that even though existing home prices decreased after the fees were passed, this may only be a short term effect. In fact, since future development also seemed to have decreased after the new development fees were introduced, then future home prices may increase in the long term if demand was to increase. Given that housing affordability is a growing concern across the U.S., increasing development costs may not be the most effective tool to raising new funds dedicated to new infrastructure construction as it may reduce the number of lower valued homes in the area.

REFERENCES

Delaney, C. J. and Smith, M. T. (1989a). Impact Fees and the Price of New Housing: An Empirical Study. *Real Estate Economics*, 17(1):41–54.

Delaney, C. J. and Smith, M. T. (1989b). Pricing Implications of Development Exactions on Existing Housing Stock. *Growth and Change*, 20(4):1–12.

Genesove, D. and Mayer, C. J. (2001). Loss Aversion and Seller Behavior: Evidence from the Housing Market. *Quarterly Journal of Economics*, 116(4):1233–1260.

Ihlanfeldt, K. R. (2007). The Effect of Land Use Regulation on Housing and Land Prices. *Journal of Urban Economics*, 61(3):420–435.

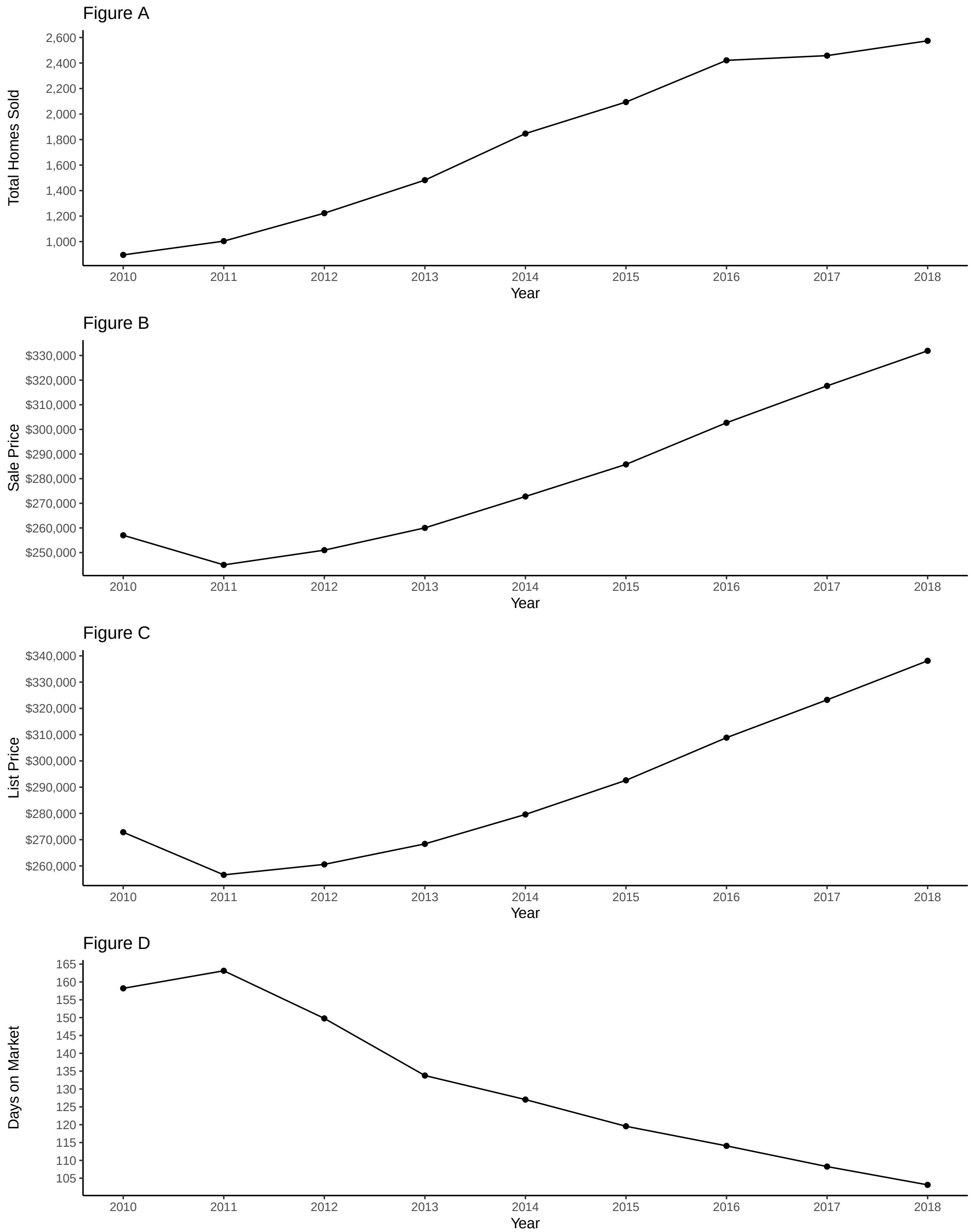
Ihlanfeldt, K. R. and Shaughnessy, T. M. (2004). An Empirical Investigation of the Effects of Impact Fees on Housing and Land Markets. *Regional Science and Urban Economics*, 34(6):639–661.

Mayer, C. J. and Somerville, C. T. (2000). Land Use Regulation and New Construction. *Regional Science and Urban Economics*, 30(6):639–662.

Quigley, J. M. and Raphael, S. (2005). Regulation and the High Cost of Housing in California. *American Economic Review*, 95(2):323–328.

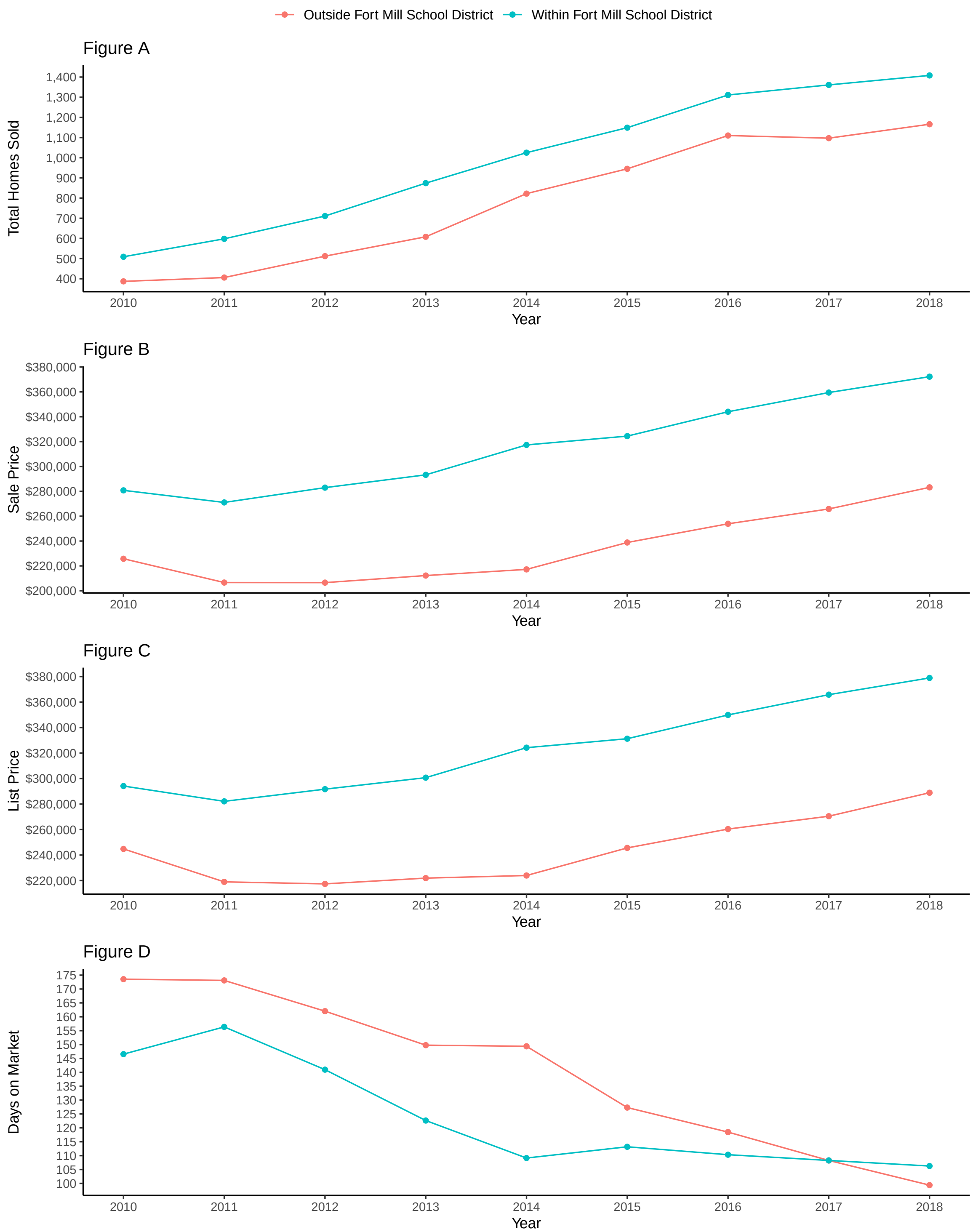
Singell, L. D. and Lillydahl, J. H. (1990). An Empirical Examination of the Effect of Impact Fees on the Housing Market. *Land Economics*, 66(1):82–92.

Stein, J. C. (1995). Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects. *Quarterly Journal of Economics*, 110(2):379–406.



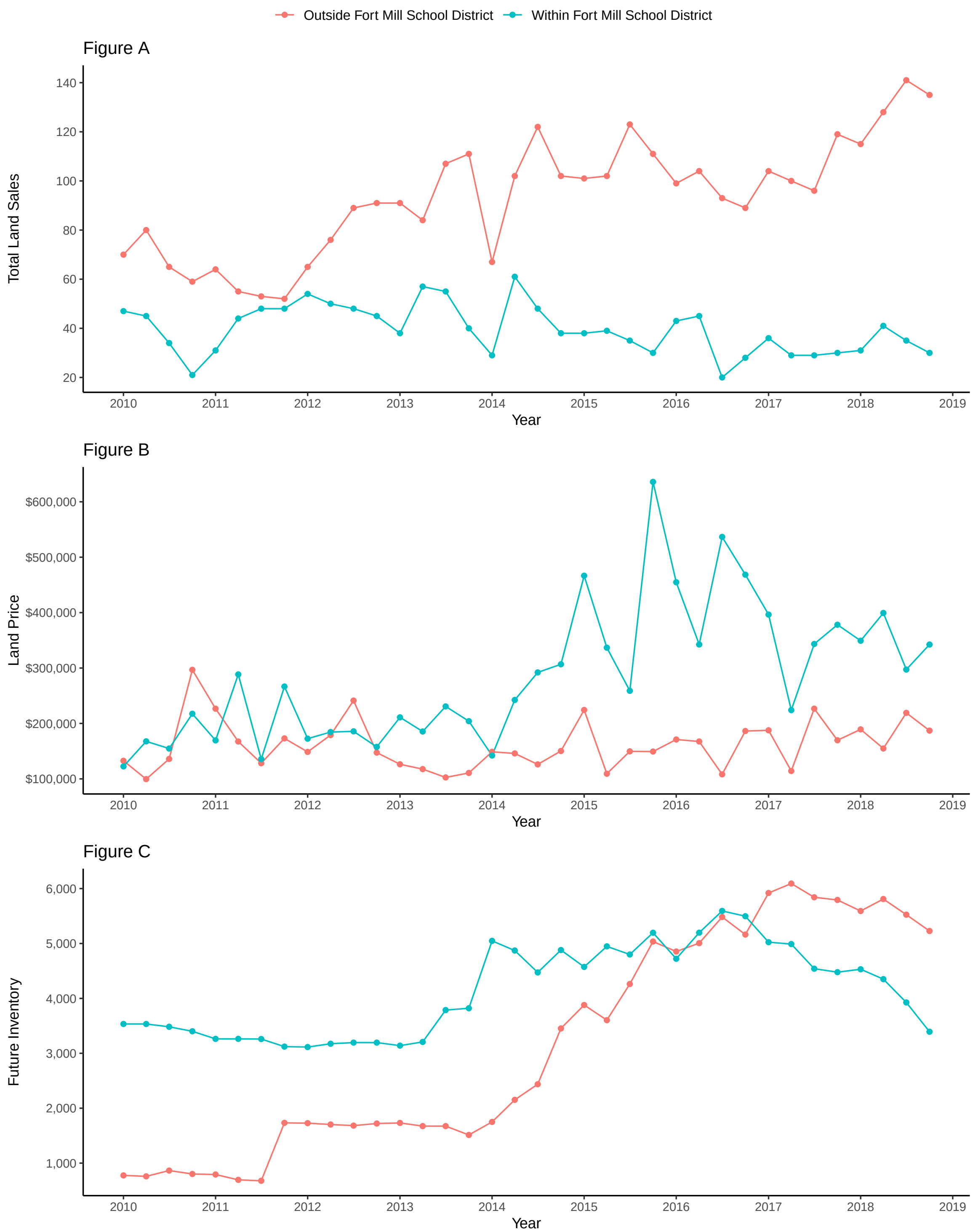
Data set is the MLS property data, which covers York County, South Carolina from 2010−2018. DOM is the average days on market until a home sold.

Figure 4.1: Total Homes Sold, Average Sale Price, Average List Price, and Average Days on Market



Data set is the MLS property data, which covers York County, South Carolina from 2010−2018. DOM is the average days on market until a home sold.

Figure 4.2: Total Homes Sold, Average Sale Price, Average List Price, and Average Days on Market by School District



Data set is the Metrostudy total quarterly land sales, average quarterly land price, and average construction data, which covers York County, South Carolina from 2010−2018. Future inventory is the total number of lots submitted for consideration to the municipality for development approval/zoning. These lots are not ready for homes to be constructed upon them.

Figure 4.3: Land Sales, Land Prices, and Future Construction

Table 4.1: Summary Statistics

Variables	MLS Property	Metrostudy Property	Land Sales
Total Observations	15,999	26,842	4,785
New Home Sales	2,789	3,703	
Within Fort Mill School District	8,946	9,294	1,420
Sale Characteristics			
Sale Price	$289,939.57	$249,877.55	$194,072.01
Age	14.54	18.14	
Bedrooms	3.65	3.57	
Sqft Finished	2,555.79	2,384.22	
Sqft Lot	0.92	0.80	
Baths Full	2.45	2.35	
Baths Half	0.56	0.51	
DOM	123.74		
List Price	$297,533.26		

Note:
Data set covers York County, South Carolina from 2010-2018. Sale characteristics are average values. DOM is the days on the market until a home sold.

Table 4.2: Difference-in-Difference-in-Difference Property Estimates

	Sale Price		List Price
Variable	MLS	Metrostudy	MLS
firstfeepass x fmschools	-0.072**	-0.081***	-0.056*
	(-2.369)	(-5.838)	(-1.780)
firstfeepass x fmschools x new	0.116	0.120***	0.103
	(1.631)	(5.706)	(1.366)
firstfeecoll x fmschools	-0.053**	-0.085***	-0.044*
	(-2.110)	(-4.940)	(-1.754)
firstfeecoll x fmschools x new	0.040***	0.118***	0.051***
	(3.972)	(3.381)	(6.775)
secondfeepass x fmschools	-0.049	-0.092***	-0.025
	(-1.216)	(-3.203)	(-0.599)
secondfeepass x fmschools x new	0.091**	-0.060	0.057
	(2.032)	(-1.113)	(1.291)
secondfeepasscounty x fmschools	-0.091**	-0.135***	-0.078*
	(-2.269)	(-3.476)	(-1.852)
secondfeepasscounty x fmschools x new	0.020	0.004	0.065**
	(1.078)	(0.521)	(2.167)
secondfeecoll x fmschools	-0.133***	-0.170***	-0.110***
	(-4.316)	(-5.801)	(-3.683)
secondfeecoll x fmschools x new	0.072***	0.032**	0.069***
	(3.298)	(1.996)	(4.342)
Intercept	5.459***	5.250***	5.787***
	(17.461)	(26.276)	(18.560)
Housing Characteristics	Yes	Yes	Yes
R^2	0.765	0.742	0.767
N	15,999	26,842	15,999

Note:
*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] The log sale price is the independent variable in the first two columns and the log list price is the independent variable in the last column. fmschools is equal to 1 if the home sale was in the Fort Mill school district. firstfeepass is equal to 1 if the home was sold between 8-24-2015 and 10-1-2015. firstfeecoll is equal to 1 if the home was sold between 10-1-2015 and 6-28-2018. secondfeepass is equal to 1 if the home was sold between 6-28-2018 and 7-16-2018. secondfeepasscounty is equal to 1 if the home was sold between 7-16-2018 and 8-13-2018. secondfeecoll is equal to 1 if the home was sold after 8-13-2018.

[b] Each regression includes city identifiers and time-based fixed effects (year, month, day, and day of week sold). Standard errors are clustered at the city level. t-statistics are in parenthesis.

Table 4.3: Difference-in-Difference Land and Construction Estimates

Variable	Log Sale Price	Weekly Lot Sales	Quarterly Future Inventory
firstfeepass x fmschools	0.239***	-2.054	
	(2.603)	(-1.255)	
firstfeecoll x fmschools	0.124	-2.267***	-92.931
	(1.497)	(-5.703)	(-0.266)
secondfeepass x fmschools	0.697***	-1.569	-1092.507
	(5.564)	(-0.586)	(-1.123)
secondfeepasscounty x fmschools	0.029	-7.319***	
	(0.245)	(-3.850)	
secondfeecoll x fmschools	0.002	-4.644***	-1526.507
	(0.011)	(-5.288)	(-1.568)
Intercept	10.493***	3.649***	224.655
	(84.818)	(5.299)	(0.917)
R^2	0.437	0.463	0.789
N	4,785	901	144

Note:
*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] fmschools is equal to 1 if the sale was in the Fort Mill school district. firstfeepass is equal to 1 if the sale took place between 8-24-2015 and 10-1-2015. firstfeecoll is equal to 1 if the sale took place between 10-1-2015 and 6-28-2018. secondfeepass is equal to 1 if the sale took place between 6-28-2018 and 7-16-2018. secondfeepasscounty is equal to 1 if the sale took place between 7-16-2018 and 8-13-2018. secondfeecoll is equal to 1 if the sale took place after 8-13-2018.

[b] Land price regression includes city, year, month, day, and day of week fixed effects. Weekly lot sale regression includes year and week fixed effects. Quarterly future inventory includes year and quarter fixed effects. Standard errors are clustered at the city level for the land price regression in column one. t-statistics are in parenthesis.

Table 4.4: Difference-in-Difference Demand Estimates

Variable	New Weekly Sales	New Weekly Listings	Cumulative Weekly Listings	DOM
firstfeepass x fmschools	-2.803	-3.784	736.392	0.218***
	(-0.964)	(-1.435)	(1.601)	(4.080)
firstfeepass x fmschools x new				0.205
				(1.289)
firstfeecoll x fmschools	0.654	1.755**	1820.271***	0.167***
	(0.790)	(2.390)	(16.002)	(7.620)
firstfeecoll x fmschools x new				0.040
				(0.487)
secondfeepass x fmschools	4.900	2.950	2334.475***	0.183***
	(0.898)	(0.595)	(2.950)	(3.061)
secondfeepass x fmschools x new				0.278*
				(1.797)
secondfeepasscounty x fmschools	2.649	-4.157	1928.475***	0.153***
	(0.647)	(-1.119)	(3.435)	(3.530)
secondfeepasscounty x fmschools x new				0.550**
				(2.386)
secondfeecoll x fmschools	-1.910	-4.005*	2382.075***	0.155***
	(-0.941)	(-1.907)	(9.243)	(3.223)
secondfeecoll x fmschools x new				0.190
				(1.414)
Intercept	-5.161***	-5.395	7823.758***	1.938***
	(-3.492)	(-0.982)	(66.622)	(3.558)
Housing Characteristics	No	No	No	Yes
R^2	0.681	0.706	0.900	0.095
N	938	1,034	1,404	15,999

Note:
*** indicates significance at the 1% level, ** indicates significance at the 5% level, * indicates significance at the 10% level

[a] The independent variables in the first two columns are the new weekly sales and new weekly listings respectively. The independent variable in the third column is the cumulative weekly listings. The independent variable in the last column is the log of the days on the market. fmschools is equal to 1 if the sale was in the Fort Mill school district. firstfeepass is equal to 1 if the sale took place between 8-24-2015 and 10-1-2015. firstfeecoll is equal to 1 if the sale took place between 10-1-2015 and 6-28-2018. secondfeepass is equal to 1 if the sale took place between 6-28-2018 and 7-16-2018. secondfeepasscounty is equal to 1 if the sale took place between 7-16-2018 and 8-13-2018. secondfeecoll is equal to 1 if the sale took place after 8-13-2018.

[b] The new sales, new listings, and cumulative listings regressions include year and week fixed effects. The DOM regression includes city, year, month, day, and day of week fixed effects. Standard errors are clustered at the city level for the DOM regression in the last column. t-statistics are in parenthesis.

CHAPTER 5: CONCLUSION

This dissertation took a deeper dive into the housing market both during and after the Great Recession. Through three articles, this dissertation addressed the potential mismeasurement of home price indices and its implications on mortgage default research during the recession and examines the potential drivers of the housing affordability issue since the recession.

The first essay provided a rational reason why the distribution of home sellers is drastically different during economic downturns than during normal economic times. I showed that when the economy enters into a recession, a very small subset of current homeowners will voluntarily choose to put their house on the market to sell. This subset of homeowners is not representative of the total homeowner population, which implies that selection bias may affect observed home transaction prices. Since home price indices use these observed transaction prices, attempts to impute the value of every home in the housing stock is incorrect and will lead to inference issues.

The second essay examined the entrance of institutional investors in the single-family housing market before and after the 2007-2009 recession. I empirically tested whether the entrance of institutional investors contributed to the subsequent increase in home prices after the recession. Using just under 337,000 home sale transactions for the Charlotte region between the years 2005-2017, I found that institutional investors paid a discount of about 8.13%-11.19% per transaction. Additionally, I found that an increase in institutional investor home purchases in the single-family housing market had a positive statistical impact on individual home prices but only a moderate economic impact.

The third essay investigated the impact of increased development fees in the residential housing market of York county, South Carolina. In July 2018, the Fort Mill school district, located within York county, increased new residential impact fees from $2,500 to $18,158. As the increased development fees only pertained to a single school district in the county, I considered the increase as a quasi-natural experiment and tested for difference-in-differences in listing prices, closing price, and inventory after two separate increases in development fees. The evidence suggested that increased development fees correspond with a decrease in asking and closing prices and an increase in available inventory in the period immediately after the increase. This result is consistent with the capitalization of expected future tax liabilities in the prices of existing residential properties.

In conclusion, this dissertation showed that that there is still much to learn about the housing market. Through theoretical evidence, I showed that home price indices must take into account the selection bias to accurately impute the value of every home in the housing market. In addition, I showed that home prices have been increasing significantly since 2012, but institutional investors are not the main contributor of this increase. In order to address housing affordability issues, economic policies need to address both demand from the overall population as well as the lack of supply in the housing market. These are areas that desperately need additional research.

REFERENCES

Bhutta, N., Dokko, J., and Shan, H. (2017). Consumer Ruthlessness and Mortgage Default During the 2007 to 2009 Housing Bust. *Journal of Finance*, 72(6):2433–2466.

Case, B. and Quigley, J. M. (1991). The Dynamics of Real Estate Prices. *Review of Economics and Statistics*, 73(1):50–58.

Haurin, D. R. and Hendershott, P. H. (1991). House Price Indexes: Issues and Results. *Real Estate Economics*, 19(3):259–269.

Laufer, S. (2018). Equity Extraction and Mortgage Default. *Review of Economic Dynamics*, 28:1–33.