Title
Essays in Housing Markets and the Real Economy

Author
Lako, Christopher M

Publication Date
2020

Essays in Housing Markets and the Real Economy

by

Christopher M Lako

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Nancy Wallace, Chair
Professor Martin Lettau
Professor Emi Nakamura
Professor David Sraer

Spring 2020

Essays in Housing Markets and the Real Economy

Abstract

Essays in Housing Markets and the Real Economy

by

Christopher M Lako

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Nancy Wallace, Chair

This dissertation consists of three chapters on the effects that housing markets have on the real economy. In the United States, personal real estate had an aggregate market value of $30 trillion in 2019.[1] Additionally, data on real estate transactions and the underlying loans and properties are extremely high quality, thereby making real estate an ideal empirical laboratory.

The first chapter shows that mortgage credit access is vital for small business financing and alleviating credit constraints. To do this, I rely on micro-data from a merge of the personal home equity extraction activity of business owners to the confidential IRS tax records of their businesses. With direct measurement, I find that one out of four small businesses created during the mid-2000s were funded by personal home equity, double the rate previously thought based on evidence from survey data. Entrepreneurs use their personal home equity to alleviate credit constraints, which this chapter finds has a long-run effect on both the survival and employment levels of small businesses. Not only are new businesses credit constrained, but existing small businesses also face credit constraints that have a persistent effect. Following the Great Recession, restrictions to mortgage credit access have caused one-third of the decline in firm entry rates in the post-crisis period.

In the second chapter, I show how variations in personal housing wealth feeds into professional behavior through studying mutual fund managers. The literature on overconfidence in financial markets has primarily focused on retail investors. Using novel data that identifies the personal real estate holdings of fund managers, this chapter studies the degree to which overconfidence affects the returns and investment behavior of institutional investors. Positive shocks to the personal real estate of mutual fund managers should not affect their professional behavior. However, this chapter finds that a one standard deviation positive home price shock leads to a decline in 4-factor alpha of 37bps per year. This is due to fund managers becoming overconfident in their underperforming trading positions, making worse selling choices, and trading more frequently. Fund managers who are more likely to be affected by overconfidence, such as less educated and less experienced fund managers,

[1] Source: Federal Reserve Flow of Funds

show a much stronger response. This chapter provides evidence that overconfidence is time varying and shows how institutional investors respond to behavioral shocks that should be orthogonal to their professional duties.

In the third chapter, I, along with my co-authors, Aya Bellicha, Richard Stanton, and Nancy Wallace, study how the vast outstanding stock of mortgage debt affects U.S. Treasury bond yields. We propose an empirical duration measure for the stock of U.S. Agency MBS that appears to be less prone to model risk than measures such as the Barclays Effective Duration measure. We find that this measure does not appear to have a strong effect on the 12-month excess returns of ten-year Treasuries as would be expected if shocks to MBS duration lead to commensurate shocks to the quantity of interest rate risk borne by professional bond investors (Hanson, 2014; Malkhozov, Mueller, Vedolin, and Venter, 2016). Given this negative reduced form result, we then explore the mortgage and treasury hedging activities of the primary MBS investors such as commercial banks, insurance companies, the agencies, the Federal Reserve Bank, mutual funds, and foreign investors. We find that the only investors that may follow the models of Hanson (2014) and Malkhozov et al. (2016) are life insurance firms. We also find a relation with banks however we cannot rule out that this is merely correlation. Life insurance firm market share has declined over the period, dropping below 10% since 1996 and reaching 4% in 2016. Of the investors we are not able to study, hedge funds and pensions/retirement funds are the two investor groups that may trade along the Hanson (2014) and Malkhozov et al. (2016) models. However, although these two investor groups held almost 25% of the Agency MBS market (including households and non profit organizations) in the late 1990s, post crisis their share has fallen below 10%.

To my father

Contents

List of Figures

List of Tables

Acknowledgments

I am deeply grateful to the many people who have assisted and guided me in my doctoral studies. First, I would like to thank Nancy Wallace, my dissertation chair, for her guidance, support, and immense wealth of knowledge. I cannot imagine a more supportive, helpful, and encouraging advisor. I am also grateful to my other committee members, Martin Lettau, Emi Nakamura, and David Sraer, whose input and guidance helped shaped my research. I also owe a special thank you to Amir Kermani and Christopher Palmer who provided early mentoring and training while I was their Research Assistant. I could not have completed my third chapter without my co-authors, Aya Bellicha, Richard Stanton, and Nancy Wallace, who spent long hours discussing our research, which helped train me as an economist. In addition to the faculty mentioned, I am grateful for all of the faculty in both the Haas School of Business and the Economics department who taught me, attended my seminars, and were always willing to meet one on one to discuss my research.

I thank my parents who raised me and always pushed me to focus on my education. I am especially grateful to my late father who encouraged and supported my desire to pursue a PhD. I also thank my husband, who moved to California after his post-doc, for helping provide a source of support through my PhD. I apologize for the long hours and for often being distracted. My friends outside the PhD program were immensely helpful in providing an outlet outside academia and distraction from research. Prior to my doctoral studies, I worked in industry at Freddie Mac. This period helped me gain the technical expertise and institutional knowledge that assisted me in my doctoral work. I would like to thank the economists that I worked with, Zheng Liu, Jan Luytjes, Frank Nothaft, and Hanqing Zhou, as well as my director Charles McKinney, who encouraged me to pursue my PhD.

I am also very thankful to my fellow classmates. I am lucky to have had my cohort as a study group during the coursework phase of the PhD and for the friendships that we forged. From feedback in pre-seminar, to talking casually in the office or over coffee, to dinner conversations, and everything else in between, my fellow classmates have been invaluable.

Lastly, I am grateful for having access to both the PhD office at Haas and the Fisher Center for Real Estate and Urban Economics. Linda Algazalli, Thomas Chappelear, Melissa Hacker, Charles Montague, and Lisa Villalba provided exceptional support navigating the administrative issues during my PhD. Additionally, Thomas Chappelear provided exceptional editing services for my thesis, as well as being a great resource overall. Paulo Issler has built an amazing research lab and data infrastructure, without which I would not have been able to complete my research, and was always willing to discuss my research. Lastly, I would like to thank Angela Andrus who provided guidance with the approval process for Census data and assistance working with the data.

Chapter 1

The Long-Run Effects of Mortgage Credit Access on Entrepreneurship

1.1 Introduction

For most households, personal housing equity is the largest source of savings (Campbell, 2006). In corporate finance, pledging durable assets such as real estate is a common way to alleviate credit constraints (Kiyotaki and Moore, 1997), making the link between mortgage credit access and small business financing a natural one. This paper shows that the personal home equity of entrepreneurs was used to fund one out of four small businesses created during the mid-2000s, highlighting the vital role of mortgage credit as a funding source for small businesses. This role was previously greatly underestimated in the literature. Since the Great Recession, the share of new (entrant) small businesses funded with personal home equity has fallen to one out of twenty. Variation in the supply of mortgage credit affects the ability of small businesses to fund projects. On the intensive margin, this paper finds that both entrant small businesses (those in their first year of operation) and continuing small businesses (those that have survived at least three or four years) face severe credit constraints that lead to a permanent disadvantage.[1] Furthermore, tighter mortgage credit standards explain one-third of the decline in firm entry rates since 2006.

Explicitly showing the link between mortgage credit and entrepreneurship is difficult due to a lack of individual level data that shows both entrepreneurs extracting personal housing equity and a longitudinal panel of business outcomes. This has led past research to proxy for entrepreneurs extracting housing equity with changes in housing wealth, and to largely focus on the extensive margin of entry into entrepreneurship (Corradin and Popov, 2015; Kerr,

[1] The intensive margin effect for entrant firms shows how initial size, long-run size, and survival are affected by exogenous relaxations in credit constraints in the first year of operation (conditional on entry). As opposed to the extensive margin effect, which would show how the choice to start a business is affected by exogenous relaxations in credit constraints. The intensive margin effect for continuing firms shows how size, long-run size, and survival are affected by exogenous relaxations in credit constraints in year three or four (for firms that have survived for at least three of four years, respectively).

Kerr, and Nanda, 2015; Adelino, Schoar, and Severino, 2015). Consequently, the findings of these papers have been mixed and as a result, the degree to which small businesses are credit constrained remains an open question in the academic literature (Hurst and Lusardi, 2004). This paper contributes to this literature by constructing a novel household-firm linked panel dataset from multiple administrative datasets, including the restricted use Longitudinal Business Database (LBD). With these data, this paper more precisely estimates the response of small businesses to relaxations in their credit constraints, estimates long-run effects, and separately shows the effect for both entrant (conditional on entry) and continuing small businesses.

This paper shows that a small exogenous increase in available credit at entry causes small businesses to start larger, grow faster, permanently remain larger, and have higher higher survival rates (measured out to five years). Initial funding has an irreversible effect on small businesses. For every $88,000 increase in a small business's credit access at entry, one additional job is created. Similarly, continuing small businesses strongly and immediately respond to credit shocks. When small continuing businesses receive a positive credit shock, they immediately expand and permanently remain larger. For continuing firms, a $153,000 credit shock creates one additional job (implying that they are less constrained than entrant firms). Schmalz, Sraer, and Thesmar (2017) establish that there is a persistent impact of credit constraints at entry for businesses registered in 1998 in France. This paper complements theirs by showing that the persistent effect at entry holds for the more recent period in the United States, which has a very different banking market, and is true for continuing firms as well.[2]

Identification is a challenge for studying the severity of credit constraints given the high-dimensional heterogeneity across both the associated households and their businesses. To establish a causal impact of credit shocks on the long-run outcomes of businesses, this paper utilizes variation in the availability of home equity via a matched pair framework. At a high level, two similar businesses that are located in the same zip code and that are owned by similar entrepreneurs who live in different zip codes from the businesses and from each other are compared. This isolates exogenous variation in the availability of credit stemming from differential zip code level home price growth within a narrow geographical area. A similar approach has been used by Bernstein, McQuade, and Townsend (2018) to study innovation and Stoffman, Pool, Yonker, and Zhang (2018) to study asset pricing.

As a third contribution, this paper finds causal evidence that the tightening of mortgage credit standards post-crisis has been an important channel for the lack of recovery in business formation rates. Small businesses rely on housing wealth to alleviate their credit constraints. Consequently, when mortgages become harder to obtain, business entry is negatively impacted. Despite home prices recovering between 2009 and 2016, cash-out refinancing volume remains depressed due to a tightening of mortgage credit standards. From its peak in

[2]In France, homeowners with a mortgage cannot extract housing equity, while in the United States they can. In addition, mortgages in France generally cannot be prepaid and housing wealth is not viewed as something that would be borrowed against.

2005, the amount of home equity extracted by households has declined more than 60%, with volume currently at 2001 levels (Figure 1.1). To identify a causal link, this paper constructs a Bartik shift share instrument that exploits an institutional detail about independent mortgage banks (IMBs). IMBs, such as Quicken Loans (one of the largest mortgage lenders), aggressively target their existing customers to persuade them to refinance. This provides a source of exogenous variation in local refinancing activity. Robustness tests rule out that the predicted IMB share directly correlates with measures that drive business formation, other than refinancing. To further show that the effect is driven by a lack of mortgage credit access and not changes in local demand, the effect is shown to hold for industries that are not reliant on local demand.

Past research largely focused on the impacts of housing wealth on business entry during the housing bubble of the mid-2000s. However, the impact of a lack of access to housing wealth on business entry in the decade since the 2008 financial crisis has not been studied. If mortgage debt was a substitute for other forms of credit then reductions in access to liquid housing wealth should not have a noticeable impact on business entry. Recent work by Davis and Haltiwanger (2019) looks at the impact that home prices have had on the share of firms that are less than five years old between 1999 and 2014. This paper differs in that it shows the impacts that a tightening in mortgage credit supply has had specifically on firm entry in the period since 2009—a period during which home prices rose dramatically and mortgage refinancing volume stagnated.

The benefits of home equity extraction found in this paper are in in contrast to the results for the average household that extracted home equity during the mid-2000s. Generally, households that extracted home equity had lower credit scores and used the extracted home equity to increase their consumption (Bhutta and Keys, 2016). However, home equity extraction was also used by a large subset of the population with prime credit scores to start and grow businesses, which led to job and GDP growth. This would be irresponsible if entrepreneurs who use home equity start weaker businesses, but this is not the case. This paper finds that home equity funded businesses are similar to and face similar survival rates as their non-home equity funded counterparts. Alleviating credit constraints via housing collateral allows entrepreneurs to start productive businesses at a more optimal size, which in recent years has not been possible.

Recent work by Bahaj, Foulis, and Pinter (2017) finds that increases in personal home values of firm directors in the United Kingdom leads to a contemporaneous increase in corporate investment (via a personal guarantee for business debt). In contrast, this paper estimates the long-run response to credit supply shocks by age for small firms in the United States (for survival, employment, and payroll). Additionally, this paper shows more broadly the impacts that restrictions on mortgage credit have had on firm entry and directly shows the actual rates of home equity use by small businesses.

This paper builds on the impacts of credit allocation in the corporate finance literature. Early research showed that credit may be rationed due to informational asymmetries (Stiglitz and Weiss, 1981). One way to overcome asymmetric information is through lending on soft information, which is acquired via relationships between business owners and bankers (Petersen and Rajan, 1994; Berger and Udell, 1995). However, acquiring a banking relationship takes time and as a consequence often is not an option for entrant small businesses. To overcome this, entrepreneurs generally rely on personal assets to fund their entrant businesses (Robb and Robinson, 2014). Early research on credit constraints studied wealth and found that entry into entrepreneurship increases with wealth (Evans and Jovanovic, 1989). The literature progressed to studying transitions into entrepreneurship after receiving an inheritance (Blanchflower and Oswald, 1998; Hurst and Lusardi, 2004). Lately, the focus of this literature has shifted to studying housing wealth shocks. More broadly, in recent years the corporate finance and household finance literatures have focused on the impacts that credit supply shocks have on businesses and households.[3] This paper fits into these literatures by showing the macro implications that disruptions to mortgage credit access have on entrepreneurship and the long-run intensive margin effects of credit constraints by firm age. Lastly, this paper builds on the literature of business dynamism by identifying one of the channels behind the decline in firm entry rates.[4]

The paper proceeds as follows. Section 2 describes the data and provides new stylized facts on the use of home equity by small business owners. Section 3 provides the empirical methodology and results for the impact of credit constraints on the intensive margin. Section 4 shows causal evidence that the lack of recovery in home cash-out refinancing activity since the Great Recession has negatively impacted business formation rates. Lastly, Section 5 concludes.

1.2 Data and the Funding of Small Businesses By Housing Equity

1.2.1 Micro-data Overview

Empirical research on small businesses has been hindered by a lack of time-series data at the business level linking business outcomes and the household balance sheet of the entrepreneur. This is true even for administrative datasets. In the study of personal home equity as a funding source for small businesses, and more broadly, the effect of credit constraints on small businesses, there have been a wide range of findings due to this lack of data. This paper relies on novel merges to create a "big" dataset of small business administrative data in the United States linked to the mortgage and refinancing activity of the business owner.

[3]See Chodorow-Reich (2013), Greenstone, Mas, and Nguyen (2014), Krishnan, Nandy, and Puri (2014), Laufer and Paciorek (2018), Benmelech and Ramcharan (2017), DeFusco, Johnson, and Mondragon (2017), Goodman (2017), Bord, Ivashina, and Taliaferro (2018), Gete and Reher (2018), Mondragon (2018), Nguyen (2019).

[4]Past work has largely identified the issue of the long-term decline in firm entry rates and the related economic consequences. See Decker, Haltiwanger, Jarmin, and Miranda (2014), Hathaway and Litan (2014), Gourio, Messer, and Siemer (2016), Siemer (2016). Recent work has identified import competition and changing demographics as two channels causing the decline in firm entry rates (Pugsley and Şahin, 2018; Karahan, Pugsley, and Şahin, 2019).

The dataset constructed in this paper is a unique merge between administrative data on business outcomes from the restricted use Longitudinal Business Database (LBD) and administrative data for real estate transactions from ATTOM. The LBD is a confidential dataset housed within Federal Statistical Research Data Centers of the US Census Bureau. The LBD provides annual establishment level data on survival, employment, and payroll.[5] The data are available for all businesses with at least one paid employee in the United States and are constructed from IRS tax return data. A key feature of the LBD is the LBDNUM variable, which allows an establishment to be longitudinally tracked over time. In this paper, small businesses are defined as businesses with initial employment of ten or fewer employees and that started as single-unit firms.

Despite the richness of the LBD data, it does not contain information on credit access. For this, data from ATTOM are utilized. ATTOM is a comprehensive property and transaction level dataset on residential (and commercial) real estate purchases and refinances, the successor to Dataquick. In the United States, county recorder offices track every real estate transaction, including residential refinances. ATTOM consolidated these records and created a dataset of this information, from which the amount of home equity extracted from a property can be constructed (described in the Online Appendix). At a high level, this paper longitudinally links real estate transactions through time at the property by homeowner level. Using the purchase transaction and past refinancing records (if any), the outstanding mortgage balance is estimated at the time of refinance and the amount of home equity extracted is constructed from this based on the refinance loan amount.

The LBD does not contain personally identifiable information and does not identify the business owner. Consequently, ATTOM is first merged to NETS. The NETS data are based on Dun and Bradstreet (D&B) data, which contain the universe of all businesses (including non-employer businesses) in the United States. The advantage of NETS is that it lists the business name, a longitudinal panel of addresses, and the name of the business owner (for 60% of businesses).[6] ATTOM is linked to NETS by merging the name of the homeowner in ATTOM to the name of the business owner in NETS (described in the Online Appendix). With the link of ATTOM to NETS, an entrant business is classified as being funded with home equity if the business owner extracts personal housing equity in the year that the business is created or the prior year.

Lastly, NETS is merged to the LBD (described in the Online Appendix) for small businesses founded between 2001 and 2011. This provides a direct link of longitudinal home equity extraction activity of business owners to longitudinal outcomes of their businesses. From which this paper shows the persistent impact that credit constraints have on both entrant and continuing businesses. For notation, entrant firms are firms in their first year of

[5]Jarmin and Miranda (2002) provide a detailed overview of the LBD (and the companion dataset, the SSEL). Revenue measures are not currently available, but should be available once the revenue enhanced LBD is completed (Haltiwanger, Jarmin, Kulick, and Miranda, 2016).

[6]The dataset also contains annual employment information. However, analysis on the reliability of the data has led to a view that it should not be relied on for high frequency time series information on businesses (Neumark, Zhang, and Wall, 2007; Barnatchez, Crane, and Decker, 2017).

operation. Whereas continuing firms are mature firms that have survived for at least three or four years.

Table 1.1 compares the overall LBD small business population to the final population utilized in this paper. 461,000 small entrant firms are successfully merged to all datasets. The small businesses in the merged sample are biased towards larger businesses with higher survival rates. As such, this should bias the results for the effect of credit constraints towards zero (since stronger firms should be less impacted by credit constraints). Unfortunately, the micro-data end with businesses formed in 2011. To study the effect of a contraction in mortgage credit supply on business entry rates since the 2008 financial crisis, county-level data between 2009 and 2016 are used. This is described in the Decline in Refinancing Activity and Business Formation section.

1.2.2 The Role of Housing Equity in Small Business Financing

Past research that looked at the funding structure of entrant small businesses relied on survey data. Robb and Robinson (2014) provided the first insight into the funding of entrant small businesses by using confidential survey data from the Kauffman Family Survey (KFS).[7] They find that 16% of entrant businesses in 2004 were founded with housing equity. Similarly, Kerr et al. (2015) used the public use 2007 Survey of Business Owners (SBO) micro-data, which is based on survey results from businesses alive as of 2007. They find that 12 to 14% of small businesses were funded with home equity at entry between 2003 and 2007.

Survey data has issues with accuracy (Hurst, Li, and Pugsley, 2014) and non-response (Campbell, 2006). As a solution to these issues, this paper directly observes if an entrepreneur extracts housing equity in a narrow period around when their business is formed.[8] Table 1.2 compares the degree to which small business owners fund their business with housing equity to the rates found in prior research. The actual use of housing equity is almost double what has been reported on surveys. Despite housing wealth effects not being particularly large during the mid-2000s (Guren, McKay, Nakamura, and Steinsson, 2018), this paper shows that housing equity was used by roughly one out of every four small business startups during this period.

Figure 1.2 shows the use of home equity across regions and time for entrant small business by the year they are founded.[9] Nationally, one out of four small entrant firms were funded with the personal home equity of the firm's owner in the mid-2000s, but since 2008 this

[7] Ballou, Barton, DesRoches, Potter, Zhao, Santos, and Sebastian (2007) provide a detailed overview of the KFS.

[8] While it is not observed if the extracted home equity is invested in the business, it is unlikely that a noticeable amount of the home equity extracted was used for purposes other than funding businesses. In later sections, strong results from the extracted housing equity on the business are found.

[9] The underlying data (ATTOM merged to NETS) is restricted to business owners who also own a home, which would bias upwards the statistics on the percent of business owners who use personal home equity. To correct for this, the statistics are adjusted by the percentage of the total population of business owners who own a home within the same cohort that the statistic is calculated for. Home ownership rates are calculated from the American Community Survey (ACS) micro-data.

has fallen to one out of twenty.[10] Between 2001 and 2003 the use of home equity by small business owners rose sharply in all regions. After 2003, there were strong regional differences in the patterns of home equity use. On the West Coast, home equity use was fairly uniform between 2003 to 2006, while in the Midwest the use of home equity peaks in 2003. The pattern for the Midwest is similar to the overall national pattern of home equity extractions, which also peaked in 2003. Bhutta and Keys (2016) attribute the decline in home equity extraction rates after 2003 to the rise in interest rates, highlighting the role that funding costs have in the decision to extract home equity.

Following the 2008 crisis, home equity sharply declined as a source of credit for entrant small businesses. While the data in this paper end in 2011, it is unlikely that the use of housing equity has recovered given the aggregate lack of recovery in cash-out refinancing volume (Figure 1.1). Additionally, comparing within survey data from the Census, the use of home equity has fallen by almost two-thirds to 5% in 2016 from 15% in 2005 (Table 1.2). This new finding highlights the importance of understanding how mortgage credit access propagates through the economy.

1.3 Intensive Margin Impact of Credit Constraints

To understand why small business owners rely so heavily on their personal housing collateral, it is necessary to understand the degree to which they are credit constrained. This section estimates the intensive margin long-run impact of credit constraints on small businesses. Section 3.1 explains the general empirical methodology to isolate exogenous variation in credit access. Sections 3.2 and 3.4 explain the specific approach to studying entrant and continuing businesses, respectively. Sections 3.3 and 3.5 present the results.

1.3.1 General Empirical Methodology

A key challenge in studying how small businesses respond to credits shocks with micro-data is that both business owners and their businesses are highly idiosyncratic. To isolate the long run response of a business to exogenous credit shocks, variations in both the business owner and the business have to be controlled for. Simple controls and fixed effects are unlikely to be sufficient in controlling for the high-dimensional idiosyncratic differences. This paper uses variation in the amount of home equity extracted from idiosyncratic home price growth as a credit shock. To isolate idiosyncratic home price growth, business owners who live in similar zip codes are studied. Simply comparing the amount of home equity extracted would lead to estimating a regression across heterogeneous populations, as the amount of home equity extracted by a homeowner is strongly correlated to their home value. This in turn is associated with their wealth, among other covariates. A homeowner who extracts $250,000

[10] The NETS data included non-employer firms. To show that this finding holds for employer firms as well, the time series are shown for small firms of various initial sizes. In all cases, the ratios of firms funded with home equity at entry are the same across time (Figure 1 in the Online Appendix).

of home equity will vary on observable qualities between a homeowner who extracts \$100,000 of home equity.

For example, linear or non-parametric controls for home prices would not control for differences in business owners with the same home value across different commuting zones. A business owner who owns a \$200,000 home in San Francisco is distinct from a business owner with a \$200,000 home in Cleveland. Even within San Francisco, a business owner with a \$200,000 home in 2001 would be very different from a business owner with a \$200,000 home in 2007. Unless home values are interacted with controls for location and time, the estimated response to credit shocks would capture differences in ex-ante wealth across business owners.

Additionally, business characteristics exhibit strong heterogeneity as well. A business owner who starts a restaurant in 2002 in the 02120 zip code of Boston is very different from a business owner who starts a restaurant in that same zip code in 2006. Without controlling for the interaction of zip code, year of creation, and industry, the estimated effect of credit shocks would pick up differences in endogenous wealth, risk aversion, skill, etc. Based on these dimensions alone, controls for home values at the same point in time within the same region among business owners who start a business in the same industry in the same year and within the same zip code would need to be included in the regression model. This strong heterogeneity across many dimensions lends itself naturally to a coarsened exact matching approach.

Matching non-parametrically isolates the causal effect of credit constraints from idiosyncratic variation in the amount of home equity extracted among a heterogeneous population in observational data. Assume business Y_i had credit access of $\$X$ at time t. Ideally, business outcomes for Y_i could also be observed if Y_i instead had credit access of $\$X + \Delta$ at time t. However, it is not possible to observe Y_i in both cases. The causal effect of credit constraints can be directly estimated if the amount of housing equity extracted was randomly assigned. In reality, between two business owners, the one who extracted greater home equity will be correlated with various covariates, such as firm industry, firm location, year of firm entry, home value, housing leverage (combined loan to value ratio), etc. Differences in these covariates will directly affect business outcomes.

Controlling for these differences with matching will allow for a causal interpretation to be uncovered (Card and Sullivan, 1988; Angrist, 1998).[11] However, exact matching will not be feasible since many of the covariates are continuous and the covariate vector is high-dimensional. Instead, a coarsened-exact matching approach is utilized (see recent work by Sarsons, 2017; Iacus, King, and Porro, 2019). At a high level, the matching procedure utilized in this paper selects pairs of businesses where both business owners and their corresponding businesses are similar across a set of characteristics. The matched pairs are restricted to businesses that are located in the same zip code (z_1). However, the business owners live in different zip codes from each other and from their firms (z_2 and z_3), where $z_1 \neq z_2 \neq z_3$ and the three zip codes are all located within the same commuting zone. This allows for the two

[11] For early work on matching see: Rubin (1973), Rubin (1974), Rubin (1977), and Rosenbaum and Rubin (1983).

business owners in the pair to experience the same local shocks to their businesses in zip code z_1. Figure 1.3 visualizes this approach. The pair member that experiences greater (less) lagged 3-year zip code level home price growth is labeled as the treated (control) member.[12] For notation, the treated business owner (A) is the one who lives in z_2 and the control business owner (B) is the one who lives in z_3.

Business owners A and B are similar and live in similar zip codes, but business owner A receives additional home price growth growth because of the zip code in which they live. Zip code level home price growth exhibits strong cross-sectional and time-series variation within a commuting zone. Across similar zip codes, variation in home price growth is not large, but even 5% differential 3-year home price growth would translate into a $15,000 credit shock for a home worth $300,000. It is testable if zip code level home price growth does not exhibit strong variation, as this will lead to a weak first stage. It may be possible to predict that future home price growth will be higher for some zip codes, for example a gentrifying zip code that attracts younger people. However, after constructing a sample of similar treated and control business owners, this is unlikely to be the case. As a result, differential home price growth between the treated and control members of the pair can be used as an exogenous source of credit that is orthogonal to business outcomes. A similar identification approach has been used by Bernstein et al. (2018) and Stoffman et al. (2018).

This paper studies how both entrant and continuing small businesses respond to credit shocks. For both analyses, the pairs are exact matched on: year of business creation, industry (SIC division), zip code of the business, and the zip codes of the homes being different from one another and the businesses (though the homes are located within the same commuting zone).[13] Given the strong and persistent effects that will be shown for the initial conditions of entrant firms, continuing firms are also exact matched on initial employment. The exact matching criteria creates a sample of similar firms, while the coarsely matched variables will create a sample where the business owners are similar as well. The coarsened section of the matching procedure is adapted for the entrant and continuing analyses and is explained below.

1.3.2 Empirical Methodology for Entrant Businesses

This section studies exogenous variation in the amount of home equity extracted for entrant small firms. The sample is restricted to businesses funded with home equity at entry to control for the fact that a firm not funded with home equity likely has differential access to other forms of funding, such as family or personal savings. Figure 1.4a graphically illustrates the identification set-up. As an example, A and B are two similar business owners who both start restaurants in zip code 94610 in year t and used personal home equity to fund the

[12] Zip code level home price growth data is provided by Zillow.

[13] SIC divisions are broad industry categories of: agriculture/forestry/fishing, mining, construction, manufacturing, transportation/communications/electric/gas/sanitary services, wholesale trade, retail trade, finance/insurance/real estate, and services. 2-digit SIC industry fixed effects are included in the regressions to control for differences across the industries within each SIC division.

business. Three years earlier, A and B both owned homes of similar value in neighboring zip codes 94611 and 94612. Over the subsequent three years, A received 10% additional home price growth, which allowed A to extract $30,000 of additional home equity. If credit constraints have a long-term effect on entrant firms, A should start their business at a larger size (three instead of two employees) and subsequently grow at a faster rate compared to B.

For this analysis, an entrant business is defined as being funded with home equity if the entrepreneur extracts >$10,000 (or >5% of home value for less valuable homes) in the year that the business is created or the prior year. This restriction removes business owners who are not extracting a sizable amount of home equity and as such are both unlikely to respond to small variations in home price growth and are more likely to have other larger sources of funding secured. While the use of the extracted home equity is not observed, it can be assumed that the majority of the funds are put towards the business given the strength of the results and the sizable cost of starting a business. This introduces measurement error into the result, which will downwards bias the coefficients towards zero in OLS. Businesses funded with home equity likely use other sources of funds as well (i.e., savings, credit cards, business loans, etc.), which will introduce a second source of measurement error that will also downwards bias the results in OLS.

To form matched pairs, business owners and their businesses are exact matched on the criteria listed in the General Empirical Methodology section. In addition, the pairs are coarsened matched to cases where the home values (measured three years prior to business creation year) are within 20% or $100k (for less valuable homes) of each other and the combined loan to value ratios (CLTV) at purchase are within 20bps of each other.[14] The coarsened match criteria restricts to similar firm owners, after having already been restricted to similar firms with the exact match criteria. The treated firm owner is the owner within each pair who experienced greater home price growth in the 3-year period prior to the business being formed. To make sure that the entrepreneurs personally experience the home price growth, the entrepreneurs must have bought their homes at least three years prior to starting the businesses. Lastly, for cases where more than one control firm is matched to a treated firm, the control firm that has the most similar home value as the treated firm is selected. A business can be a control firm multiple times (and can be both a control and treated firm in different pairs) but can only be a treated firm once (the majority of firms are only in one pair).

The matching algorithm forms a sample of 5,100 firms (based on 4,600 distinct firms). Matched firms tend to be slightly larger and their owners have higher CLTV ratios and lower home values compared to the overall population of home equity funded firms (Tables 1.4 and Tables 1.9). The starting sample is 124,000 home equity funded firms, approximately 60% of which are owned by entrepreneurs who live in different zip codes from their firms, and 45% of which are owned by entrepreneurs who bought their home at least three years prior to

[14]CLTV at purchase is the ratio of the sum of all liens at time of home purchase relative to purchase price. LTV only accounts for the 1st lien, however during the housing bubble many households included secondary liens (so called "piggyback" second liens, Lee and Tracy, 2012).

starting their businesses.[15] A large number of observations are not included in the analysis in order to create a sample where both the treated and control firm owners are similar on observables, except for the differential home price shock.

For home price growth to be exogenous within a pair, the treated and control owners must on average be similar on observable characteristics and live in similar zip codes as each other. This is the goal of matching. Table 1.3 provides summary statistics for the difference in values between the treated and control owners. On average, differences across the variables are close to zero. Covariates for home zip codes are not matched, but are also similar for both the treated and control firm owners. For robustness, a matched pair sample is constructed from the ATTOM to NETS merged sample with the same criteria that is used to form the entrant matched pairs with the LBD data. With this data it is shown that the treated and control firm owners both live the same distance from their firms on average (Figure 1.5). This is not shown using the sample formed from the LBD data due to disclosure restrictions.

One concern is that if the home zip codes of either the treated or control owners are consistently more correlated to shocks to the firm zip code, then business outcomes could be correlated to the idiosyncratic home price growth. Since the home zip codes are similar and the firm owners live similar distances from their firms, this is unlikely to be the case. Despite the treated and control owners being from similar zip codes, the treated owners receive, on average, 6.5% greater home price growth. The average home value for the population is \$206,000, which leads to an average increase of \$13,000 in available credit for the treated owners relative to the control owners.

With this matched pair sample, the effect of credit shocks on firm outcomes is estimated using two-stage least squares:

$$Y_{i,j,t+k} = \alpha + \beta ln(\widehat{\$\text{ Amount Extracted}}_{i,t}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,j,t+k} \tag{1.1}$$

firms are denoted by i, firm creation year by t, and the pair that firms belongs to by j. $\mathbb{X}_j$ are fixed effects for each pair. $Y_{i,j,t+k}$ are the outcome variables: survival in years 1, 3, and 5 (estimated as a linear probability model), employment and employment growth in years 1 through 5, and payroll in years 1 through 5. Employment growth is calculated using the standard approach in the entrepreneurship literature[16]:

$$\frac{E_{i,t} - E_{i,t-1}}{.5 * (E_{it} + E_{i,t-1})} \tag{1.2}$$

$ln(\widehat{\$\text{ Amount Extracted}}_{i,t})$ is the instrumented amount of home equity that an entrepreneur extracts in the year that the business is created or the prior year. Even with pair fixed effects, the amount of home equity extracted is endogenous. To isolate exogenous variation in the amount of home equity extracted, the following first stage is estimated:

[15] These statistics are compiled from the ATTOM to NETS merged sample due to disclosure restrictions.

[16] See Törnqvist, Vartia, and Vartia (1985), Davis, Haltiwanger, and Schuh (1996), and Haltiwanger, Jarmin, and Miranda (2013).

$$ln(\text{\$ Amount Extracted}_{i,t}) = \\ c + \tau\Delta_3 ln(\text{Zip Code Home Price}_{i,t}) + \gamma\mathbb{X}_j + \omega\text{Controls}_i + \epsilon_{i,t} \quad (1.3)$$

The first stage instruments the amount of personal home equity extracted with the log of zip code level home price growth for the 3-year period prior to business entry year t. $\mathbb{X}_j$ partials out the common trend in home price growth within each pair, leaving idiosyncratic home price growth between two similar and geographically close zip codes. If greater lagged home price growth leads to more home equity being extracted, then $\tau > 0$.

Standard errors are clustered at the SIC division by firm zip code level, which is the primary exact match criteria of the pairs. The clustering level does not include firm creation year to allow for arbitrary correlation across firm creation years among firms located within the same zip code and broad industry category. Controls$_i$ is a vector of zip code level and firm level controls that includes: log of CLTV at purchase, log of home value at year $t-3$, 2-digit SIC industry fixed effects, the log of the number of months between home purchase and year t, legal form of organization (LFO) fixed effects, and non-parametric controls for home zip code characteristics of percent white, percent renter, percent below poverty line, and median income.[17]

The control variables attempt to correct for any heterogeneity within the pair that is not accounted for by the coarsened exact matching algorithm. The first two controls, home value and CLTV, are coarsely matched, however they are included as controls in case any heterogeneity remains. Likewise, 2-digit SIC industry fixed effects are included in case heterogeneity within SIC division remains. The number of months between home purchase and firm creation year t is included to control for differences between how long a business owner waits to start a business after buying a home. LFO fixed effects control for differences among incorporation types of businesses. Lastly, the home zip code characteristic controls remove variation from the choice of home zip code within a commuting zone, if any remain after matching.

1.3.3 Entrant Business Empirical Results

If a new business is credit constrained and receives an exogenous relaxation in their credit limit, how are the business's survival, size, and growth affected? If the business is ex-ante credit constrained and the additional credit is spent productively, then loosening the credit constraints should result in the business starting larger, growing faster, and surviving longer. Second, it is important to understand if businesses that have less access to credit at entry are able to catch up from their initial disadvantage.

The effect on business survival is shown first. Table 1.5 column 1 shows the raw OLS effect on the 5-year survival rate when only the pair fixed effects are included. Having additional

[17] Zip code level control variables are from the 2000 Decennial Census.

access to credit leads to an increased survival rate. The result survives and remains stable when the vector of additional controls are included, implying that these controls are not a source of heterogeneity between the treated and control pair members. Column 3 tests the first stage and shows that within the pair, the business that receives additional personal home price growth in the 3-year period before the business is founded extracts additional housing equity. The instrument is strongly significant and the F-statistic is above 10, indicating that the instrument is not weak. The coefficient of 0.88 on lagged home price growth means that the treated member extracts 88% of the additional home equity that they have access to from variation in home price growth. The mean difference in home price growth between the treated and control firms is 6.5%, implying that the average treated firm has 5.7% of additional realized funding at creation relative to the control firm. Since the treated firm owner extracts most of the additional home equity that they have access to, this is preliminary evidence that firms are credit constrained. Otherwise, there would not be as strong of a response to exogenous variation in home price growth.

In column 4, the regression model is estimated with the amount of home equity extracted instrumented with 3-year lagged home price growth. The coefficient shows a strong positive causal effect.[18] A 10% increase in credit increases a firm's 5-year survival rate by 5.1%. For the sample used in these regressions, the average amount of home equity extracted is $100,000, which means a 10% credit shock translates into $10,000 of additional credit. The 5.1% increase in the 5-year survival rate from this small credit shock would lead to the survival rate increasing from 67% to 71%, on average. This result shows that there is a long term effect on survival from a positive exogenous credit shock. Therefore, the initial credit access of a business is strongly important and small businesses are credit constrained at entry. Schmalz et al. (2017) also found a persistent impact of initial housing wealth on small businesses for small businesses founded in 1998 in France. However, this is the first time it has been conclusively shown for small entrant businesses in the United States. Due to strong differences in the banking and loan markets and entrepreneurial demand/preferences between the two countries, it is not obvious that the results should apply to the United States.

A natural question is whether the effect on survival is apparent immediately or if it slowly accumulates. Columns 7 and 8 show the effect in years 1 and 3, respectively. The effect is present even in the firm's initial year, with the survival rate increasing from having access to additional credit. The effect in year 3 is larger than in year 5, implying a concave relationship. After year 3, the control firm starts to catch up but remains at a disadvantage in year 5. Unfortunately, survival beyond the 5th year cannot be tested due to restrictions on Census RDC disclosure. However, a 5-year survival rate effect implies that the effect is permanent.

Effects on other business outcomes are estimated to understand why business survival increases with additional credit, and to show robustness that the effect is not confined to

[18] As a robustness check, if the standard errors are clustered at the SIC division by firm zip code by creation year level, the result remains (column 6).

survival. In Table 1.6, the regression is estimated for the effect of credit constraints on employment. Columns 1-5 show that employment increases in years 1 through 5 from having access to additional credit at entry. Similar to the effect on survival, the effect for employment is persistent and remains through year 5. The sample size declines for the results on employment after year 1, due to businesses being omitted for disclosure reasons if they have missing data after their initial year.[19]

A 10% increase in credit leads to a 4.9% increase in employment at creation. In the LBD, 5.6 million small entrant firms were founded between 2001 and 2011 and created 13.5 million jobs at entry. Approximately 22% of these firms were funded with home equity at entry, leading to 2.9 million jobs created at entry by home equity funded firms.[20] For every 10% increase in housing collateral, a back of the envelope calculation shows that 140,000 additional jobs would be created. Additionally, the 0.49 coefficient for the effect on initial employment translates to one additional job being created from every $88,000 positive credit shock a small entrant business receives.

In addition to starting larger in year 1, the treated businesses grow faster between years 1 and 2 (Table 1.7). A 10% increase in initial credit increases initial firm growth rates by 18% between years 1 and 2. After year 2, growth rates between the treated and control businesses stabilize and this allows the treated businesses to remain larger through year 5. The effects on employment show that small increases in initial credit have an important effect on a firm's ability to start at a more optimal size. In addition, greater initial credit access allows businesses to operate more productively, which leads to larger initial growth rates. Businesses with less access to credit at entry are never able to catch up from their initial disadvantage.

As a last test, the effect on payroll is estimated (Table 1.8). Initial payroll is not statistically significantly larger in year 1 from a credit shock at entry. In later years, payroll increases from the credit shock at entry. A possible reason for this is that having additional initial credit allows businesses to attract more customers through increased advertising or a more attractive interior. This in turn increases revenue and demand, which necessitates a need to hire more workers. However, businesses may have to decrease the wages of workers in order to hire more employees. Studying how wages are affected by a firm's access to credit is left for future research.

A concern with the empirical approach is that differential home price growth might affect the decision to start a business. If this is the case, differential home price growth between the two pair members might directly affect business outcomes. The findings are based on small differences in home price growth between the treated and control entrepreneurs (6.5% on average), which makes this less of a concern. However, if the treated entrepreneurs required

[19]Due to restrictions on disclosure, after year 1 pairs are dropped if one of the businesses in a pair ever has missing employment/payroll data in years 2 through 5. In addition, if a firm is in multiple pairs and one of the pairs is excluded then all of their other pairs are excluded as well.

[20]This statistic is based on the merged sample of ATTOM to LBD. It is adjusted by the percent of small business owners who are homeowners to account for the fact that the merged sample only includes homeowners. The percent of small business owners who are homeowners is calculated from ACS micro-data.

additional credit in order to be convinced to start a business they would likely be of a lower skill type compared to the control entrepreneurs. This would make the treated entrepreneurs less successful, which would downwards bias the effect. Second, if the control entrepreneurs were severely constrained, because they had less home price growth, and still decided to start a business it could reveal that they are of a higher skill type. This would again downwards bias the effect. In both cases the bias from this potential concern would decrease the effect that is found.

The population of small businesses is restricted to businesses that started as a single-establishment and with ten or fewer employees. The threshold of ten was chosen because larger businesses rely less on home equity, due to the greater costs involved in starting larger businesses. Adelino et al. (2015) find that firms with more than ten employees do not respond to home price growth shocks. In addition, Patnaik (2017) shows that firms with ten or fewer employees rely on housing wealth, while firms with more than ten employees rely on bank loans. Additionally, fewer than 10% of entrant businesses between 2001 and 2011 started with more than ten employees (Figure 2 in the Online Appendix). Including this small population of larger firms would add noise to the results since they are likely not relying on housing wealth and are inherently different from smaller firms.

An additional concern is if treated firm owners are consistently more or less likely to originate their cash-out refinance mortgages with traditional banks, such as Bank of America. Traditional banks cross-sell customers, as a result bankers might encourage firm owners obtaining business loans to also obtain cash-out refinance mortgages in order to have additional credit. Non-traditional banks, such as Quicken Loans, do not originate business loans. If treated firm owners are biased towards or away from traditional banks, there could be a bias in their total amount of funding relative to control firm owners. Using the matched pair sample that is constructed from the ATTOM to NETS merged sample, it is shown that the treated and control firm owners have the same likelihood of having originated their cash-out refinance mortgages from traditional banks on average (Figure 1.6). This is not shown using the sample formed from the LBD data due to disclosure restrictions.

1.3.4 Home Equity Funded Businesses

A key assumption to generalizing the findings to all small businesses is that home equity funded businesses are similar to non-home equity funded businesses. Table 1.9 compares the populations of home equity to non-home equity funded businesses. Overall, home equity funded businesses have similar survival rates and tend to start slightly larger compared to non-home equity funded businesses. The entrepreneurs of home equity funded businesses have greater housing leverage (at time of home purchase), larger home values 3-years prior to when the business is formed, and bought their homes more recently. The average business owner who uses home equity as a funding source extracts \$100,000 of home equity. A regression model is utilized to rigorously test the differences between these two populations:

$$Y_i = \alpha + \beta \mathbb{I}(\text{Home Equity Funded}_i) + \omega \text{Controls}_i + \epsilon_i \tag{1.4}$$

where i indexes a business.

As a first test, this paper checks if housing characteristics vary across the two funding types, by including the interaction of home zip code by home purchase year fixed effects in the regression model. Additional controls include legal form of organization (LFO) and 2-digit SIC industry fixed effects. Standard errors are clustered at the home zip code level. Table 1.10 columns 1 and 2 rigorously show that home equity funded firms tend to be started by entrepreneurs who own slightly more expensive homes (5% greater home values) and who have slightly greater housing leverage (4 point higher CLTV ratios) within a zip code at a given point in time.

Next, this paper tests if home equity funded businesses differ on initial size and survival, by including the triple interaction of 2-digit SIC industry by creation year by firm zip code fixed effects in the regression model. Additional controls include log purchase value, CLTV, and fixed effects for LFO. Standard errors are clustered at the 2-digit SIC industry by firm zip code level. Table 1.10 columns 3 and 4 show that home equity funded firms start slightly larger in terms of employment, but not payroll. Interestingly, there is no differential outcome for survival.

To generalize the results for the existence of credit constraints, survival is the most important variable to test, since an insignificant difference in survival between home equity and non-home equity funded firms implies that firms funded with home equity at entry are not weaker firms. Adelino et al. (2015) presents suggestive evidence that home equity funded firms were not more likely to exit during the 2008 financial crisis compared to firms not funded with home equity. The results of this section show further proof that this is indeed the case. In results available upon request, it is found that during the 2007-2009 period, small businesses funded with home equity at entry were less likely to close compared to small businesses not funded with home equity at entry.[21]

Lastly, the mortgage default rates of small business owners who extract personal home equity to fund their businesses are shown. Overall, it is well established that homeowners who extracted home equity during the 2000s realized larger mortgage delinquent rates compared to homeowners who did not. Bhutta and Keys (2016) find that homeowners with a cash-out refinance between 2001 and 2003 experienced 20% greater 4-year default rates. By 2006, the 4-year default rate of homeowners who extracted home equity was double the default rate of homeowners who did not extract home equity.

To show mortgage default rates for business owners, Equation 4 is estimated as a linear probability model with a dependent variable equal to one if the business owner experiences a foreclosure on their personal home within four years of starting the business.[22] Controls include risk characteristics for combined loan to value (CLTV) at purchase, FICO at purchase, and initial mortgage rate.[23] The interaction of home purchase year, firm creation

[21]Due to disclosure reasons, this result is shown using survival information from NETS and not the restricted-use LBD.

[22]This analysis uses the data from the merged ATTOM-NETS sample that is further merged to McDash. Foreclosure is estimated from ATTOM. FICO and initial mortgage rate are from McDash.

[23]Home purchase value and initial firm employment are included as additional controls.

year, home zip code, firm zip code, and 2-digit SIC industry is included in order to estimate the mortgage default probability differential among similar business owners. It is important to compare among borrowers who bought their home in the same year as mortgage default rates vary substantially across home purchase years (Palmer, 2015).

Overall, there is no differential probability of foreclosure for business owners who fund their businesses with home equity (column 1, Table 1.11). Even though in the overall population, homeowners who extract home equity default on their mortgages at greater rates, business owners who rely on home equity do not default at higher rates compared to similar business owners who do not extract home equity to fund their business. For business owners who started a business between 2005 and 2007, there is an elevated probability of foreclosure if home equity is extracted to fund the business (column 3, Table 1.11). The default rate rises by 25.6%, from the population mean default rate of 24.8% to 31.2%. While the default rate is elevated during this period, the increase in default is much smaller compared to the increase in default rates for the overall population of homeowners who extracted home equity during this period. Bhutta and Keys (2016) note that default rates for the overall population of homeowners who extracted home equity during this period rose by 80-100% of the population mean. Lastly, business owners who funded their businesses with home equity since 2008 have had lower mortgage default rates compared to all business owners who started a business during this period (column 4, Table 1.11). This is consistent with the notion of tighter mortgage credit standards deterring firm entry, which is explored in Section 4.

In general, mortgage default rates are elevated for business owners, with mean default rates on mortgage debt of 19% during the 2001 to 2011 period (column 1, Table 1.11). This is likely a result of the riskiness of small business ownership. If the business fails, the business owner is at increased risk of losing their home. However, the use of home equity to fund businesses does not noticeably affect the already high mortgage default rate of business owners. This is additional evidence, beyond low survival rates, to the riskiness of small business ownership.

1.3.5 Empirical Methodology for Continuing Businesses

Continuing businesses are often ignored in the literature on credit constraints. Entrant businesses, in theory, have a more difficult time obtaining credit due to a lack of both soft and hard information and business collateral. However, continuing businesses have hard information (financial statements), the ability to build a relationship with a banker for soft information, and potentially have business collateral. Additionally, data on the long-run response of continuing businesses is difficult to obtain. This paper has the ideal dataset to test if continuing businesses are also credit constrained.

To obtain a measure of credit constraints for continuing businesses, exogenous variation in the amount of personal home equity extracted (if any) in years 3 or 4 of a firm's life is studied. Figure 1.4b graphically illustrates the identification set-up and shows the outcome if continuing businesses are credit constrained. A and B are two similar business owners who both start restaurants with the same number of employees in zip code 94610 in year t. In year t, A and B also owned homes in neighboring zip codes 94611 and 94612, respectively, with the same home value. Over the next two years, their firms grew at a similar rate in terms of employment. In year $t+3$, A has accumulated 10% additional home price growth and because of this extracts $100,000 of home equity, while B does not. (It could also be that B extracts home equity as well, but extracts less due to experiencing less home price growth.) With this exogenous variation in credit access in year $t+3$, A expands their firm and permanently remains larger. The example shows the matching exercise for extraction in year 3 (referred to as the event year)—however, matching is also performed for firms in year 4.

For a firm to be matched, it must have survived through the event year. Firms can be matched in both the age three and four event year cohorts and the response to credit shocks is jointly estimated with data from both event years, which assumes that businesses do not deferentially respond to credit shocks at age three versus age four. It is more difficult to isolate exogenous variation for continuing businesses compared to entrant businesses due to the additional dimensions of the business's past record that also must be matched on. In the section for entrant businesses, it is shown that initial conditions for a business play a persistent role in the business's success and size. Therefore, a business's initial conditions must be controlled for in the coarsened-exact match algorithm for continuing businesses by also exact matching on initial employment. Continuing firms are coarsely matched on two covariates. First, home values (as measured three years prior to the event year) must be within 20% or $100k (for less valuable homes) of each other. Second, the number of employees one year prior to the event year must be within three employees of each other. This second constraint restricts to firms that are on similar growth paths (after starting at the same size).

The treated firms are the ones that experience greater home price growth within the 3-year period prior to the event year. To confirm that the entrepreneurs personally experience the home price growth, they must have bought their homes at least three years prior to the event year. Lastly, in cases where more than one control firm is matched to a treated firm for a given event year, the control firm that has the most similar home value as the treated firm is selected. A business can be a control firm multiple times, but can only be a treated firm once within each of the two event years. Roughly half of the firms in the sample are only in one matched pair.

From the matching algorithm, 17,500 firms are matched (11,500 unique firms). The starting population for this sample are firms that have survived through the matched event year, which restricts the sample to at most 345,000 firms. Of these, approximately 60% of the firm owners live in different zip codes from their firm and 45% bought their home at least three years before the event year.[24] A large number of observations are removed in order to create a sample in which both the treated and control firm owners are similar on

[24]Due to disclosure restrictions, exact numbers are not provided and these statistics are compiled from the ATTOM to NETS merged sample.

observables, except for the differential home price shock.

On average, \$7,909 of home equity is extracted in the event year (Table 1.12). The average includes zeros for the businesses that did not extract home equity in the event year. To ensure that the matching does not bias the treated or control members towards certain characteristics, average differences between key variables are calculated. There are no noticeable differences between the treated and control businesses/owners based on a set of observable characteristics (Table 1.13). The treated owners receive an average 10% of additional home price growth, so the average home value of \$228,000 translates into a credit shock of roughly \$23,000.

To estimate the effect of credit constraints on continuing businesses, a regression model similar to equation 1.1 is utilized:

$$Y_{i,j,t+p+k} = \alpha + \beta ln(\widehat{\$ \text{ Amount Extracted}}_{i,t+p}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,j,t+p+k} \quad (1.5)$$

where p is the number of years since founding t that home equity is potentially extracted for expansion (two or three, which corresponds to firms of either age three or four). The difference between equation 1.5 and equation 1.1 is the vector of control variables, since continuing businesses have additional dimensions that need to be controlled. For continuing businesses, Controls_i also includes a fixed effect for if the business was initially funded with home equity, and depending on the regression also includes non-parametric fixed effects for initial payroll and employment and payroll growth between years 1 and 2. These last two controls attempt to control for variation in ex-ante business success. The initial payroll control is in addition to matching on initial employment.

To isolate exogenous variation in the amount of home equity extracted, the following first stage is estimated:

$$\begin{aligned} ln(\$ \text{ Amount Extracted}_{i,t+p}) = \\ c + \tau \Delta_3 ln(\text{Zip Code Home Price}_{i,t+p}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,t+p} \quad (1.6) \end{aligned}$$

The first stage instruments the amount of personal home equity extracted with the log of zip code level home price growth for the 3-year period prior to the event year $t+p$. $\mathbb{X}_j$ partials out the common trend in home price growth within each pair, leaving idiosyncratic home price growth between two similar and geographically close zip codes. If greater lagged home price growth leads to more home equity being extracted then $\tau > 0$.

β measures the causal impact of how continuing businesses respond to exogenous availability of home equity at age three or four $(t + p)$. The response is then studied for one year before to three years after year $t + p$. If the treated and control firms are similar before year $t + p$, there should be no difference in their outcomes in year $t + p - 1$. If home equity extracted in year $t + p$, due to exogenous variation in past home price growth, relaxes the credit constraints of continuing firms and they productively use the credit, a positive β should be found in year $t + p$ onward.

1.3.6 Continuing Business Empirical Results

This section starts with testing if exogenous credit shocks to continuing businesses affect firm survival. Due to restrictions on RDC disclosure, only one survival period is studied, which is three years after the event year (year 6 or 7 of the business's life depending on which age the event year corresponds to). Firms had to survive to the event year (age three or four) to be included, which restricts the population to a set of stronger businesses given low initial survival rates. Table 1.14 columns 1 and 2 estimate the regression model using OLS. No effect is found when using OLS and the coefficient is very close to zero. This implies that there is no raw effect.

The first stage shows a very strong positive effect from past home price growth on home equity extraction in the event year and the F-statistic is above 10 (column 3), indicating that the instrument is not weak. In the second stage, no significant effect is found for survival and the coefficient is still roughly zero (column 5). To test for robustness around the clustering level, columns 6 and 7 cluster at firm zip code by SIC division by firm creation year and firm commuting zone by SIC division levels, respectively. The coefficient remains insignificant. Relaxing credit constraints on continuing businesses does not affect their continued survival likelihood. This is likely the result of survival rates being convex, in that many firms close within their first few years. Following which the rate of survival levels off. Given that these firms have already survived through their most volatile years, having additional credit will not affect their future survival.

This does not imply that continuing firms are not credit constrained. To test if firm size is affected, the regression is estimated with employment as the dependent variable. First, column 1 of Table 1.15 shows that in year prior to the event year, year $t + p - 1$, there is no significant difference in employment between the treated and control firms. In the event year, employment for treated firms significantly exceed the employment of control firms (column 2). This implies that between years $t+p-1$ and $t+p$, treated firms expanded more rapidly than control firms. Column 6 confirms this by showing that the employment growth rate is larger for treated firms. In addition, treated firms remain larger for the following three years, which is when the sample ends (columns 3 through 5). Similarly, Table 1.16 shows that payroll between the treated and control firms is similar before the event year and immediately increases for the treated firms in the event year. Overall, when continuing businesses receive a credit shock, they immediately expand and permanently remain larger as a result.

The findings in Table 1.15 show that when continuing firms extract additional home equity because of exogenous availability of home equity, they immediately expand and subsequently permanently remain larger. These results provide evidence that continuing businesses are credit constrained. Having access to 10% of additional credit leads to an increase in employment of 1.1%, with the effect increasing to 1.6% after three years. The coefficient of 0.114 translates into a credit shock of $153,000 immediately creating one additional job at a small continuing firm.

Since continuing firms require a larger positive credit shock to hire one additional worker, continuing firms are less credit constrained than entrant firms. To see why this implies that continuing firms are less credit constrained, it is helpful to think in terms of a "marginal propensity to hire (MPH)", similar to the concept of marginal propensity to consume or borrow. One worker hired from an $88,000 credit shock implies that 0.011 jobs are created for every $1,000 of credit that is extended, this is the MPH for entrant firms. While continuing firms have an MPH of 0.0065 for every $1,000 of credit that is extended. Since continuing firms have a lower MPH, this implies that they have less of a need to hire from a credit shock.

These are the first results to test if continuing businesses are credit constrained. Continuing businesses have survived their early years, when they are most prone to failure. This segments the population to a stronger group of businesses. As such they are often ignored in the study of credit constraints on small businesses since it is assumed they can more easily obtain external financing. However, the results in this section show that businesses remain credit constrained as they age. When continuing businesses receive an exogenous increase in credit, they immediately expand and permanently remain at a larger size.

1.4 Decline in Refinancing Activity and Business Formation

As shown earlier, entrant small businesses relied heavily on home equity for financing during the mid-2000s. However, starting in 2008, this credit channel was virtually eliminated (Figure 1.2 and Table 1.2). Mortgage lending has become very tight and as a result, borrowers are having a difficult time obtaining mortgage credit. Goodman (2017) finds that among all home buyers, credit standards have become twice as restrictive as they were in 2001, which was prior to the looser lending standards of the subprime bubble of the mid-2000s.

Recent research has shown that this tightening of mortgage credit supply has caused fewer mortgage originations—particularly among younger, middle income, and black borrowers—and higher rental prices (Laufer and Paciorek, 2018; Gete and Reher, 2018). Given the reliance of small business owners on home equity and the strong credit constraints that they face, another possible effect of a tightening of credit standards is a reduction in business formation rates. During the Great Recession, household refinancing activity and business formation rates simultaneously fell, and both have remained permanently lower since the Great Recession (Figure 1.1). This is despite strong home price growth over the post-crisis period—between 2009 and 2018, home prices have generally reached their pre-crisis levels and in many cases exceeded them (Figure 1.7). As a result, households are likely to have significant home equity. Correlations between home equity extraction activity and firm entry rates imply that a lack of refinancing activity of small business owners is a channel for the steep decline in firm entry rates since the Great Recession. This section will show this causally.

In the United States, increased housing wealth is realized by entrepreneurs through extracting home equity and then using the cash to fund their businesses. Normally, housing wealth variations and home equity extraction activity are highly correlated. This has allowed past research to proxy home equity extraction activity with either home price growth or available housing wealth in studying the effects on business formation. However, the period since 2009 has seen a recovery in home price growth *without* a concurrent recovery in cash-out refinancing rates, thereby breaking this link. Instead, this section tests the effect of home equity extraction activity on business entry. The micro-data that is utilized earlier in this paper does not have data on firms that are created after 2011. Therefore, this section utilizes county-level measures of cash-out refinancing activity to study the impacts on business formation during the 2009 to 2016 period. Unlike data for home price growth, local measures of cash-out refinancing activity are not readily available. Using mortgage transaction level data from ATTOM, this paper constructs 1-year growth rates in the number of households extracting home equity at the county-level between 2009 and 2016.

To test the hypothesis that increased local refinancing activity leads to an increase in business formation, 1-year growth rates in county-level home equity refinancing rates are regressed onto 1-year growth rates in county-level business formation rates. However, this will suffer from omitted variable bias due to confounding demand and banking shocks that affect both refinancing activity and business formation activity. In addition, the explanatory variable of county-level growth in refinancing activity suffers from measurement error. The measure of refinancing activity is for all households in the county, not only entrepreneurs. Since growth in refinancing activity and growth in business formation rates are assumed to be positively related, the coefficient will be biased downwards in OLS.

To isolate a causal relationship that is driven by the collateral channel, two steps are undertaken. First, an instrument that isolates exogenous variation in refinancing activity that does not directly affect business formation is utilized. The instrument will also correct for measurement error. Second, to further isolate the collateral channel effect from the demand channel effect, industries that are reliant on local demand are removed. For robustness, it is shown that the result generalizes to all industries.

To construct an instrument, this paper exploits the institutional detail that independent mortgage banks (IMBs, which includes mortgage originators such as Quicken Loans) are prone to churning their originations. This means that IMBs aggressively target their prior customers to encourage them to refinance their existing mortgage. Consequently, greater IMB market share will lead to greater refinancing activity.[25] Figure 1.8 shows this effect

[25] In a SEC filing from 2015, Ellington Financial noted the aggressive targeting of customers by IMBs to refinance as a risk to MBS holdings:
"*As improved technology spreads throughout the lending industry, we believe that the lending industry will change in a number of important ways... borrowers should start to prepay more efficiently, as demonstrated by the higher prepay speeds of mortgage loans serviced by Quicken, which reportedly uses proprietary algorithms to target borrowers who are more likely to refinance, and is particularly quick at contacting borrowers about refinancing incentives... A technology-driven, broad-based increase in prepayment efficiency may put pressure on MBS prices and/or reduce the excess spread enjoyed by MBS investors.*"

graphically. Purchase mortgages that were originated by Quicken Loans are much more likely to terminate within one year compared to purchase mortgages originated by the top three bank lenders. This relationship has been true since the start of the data in 2001 and is not a new phenomenon. Buchak, Matvos, Piskorski, and Seru (2018) also noted the increased prepayments of loans originated by these IMBs.

Based on this institutional detail, this paper constructs a shift-share instrument in the spirit of (Bartik, 1991; Blanchard, Katz, Hall, and Eichengreen, 1992). For each county c, the predicted annual (t) IMB market share for purchase mortgages is calculated using the pre-existing IMB market share for the county interacted with national growth rates in IMB share (excluding county c):

$$\text{Avg }\%\text{ IMB}_{c,t-1} * \Delta\%\text{ IMB}_{-c,t} \tag{1.7}$$

The pre-existing shares (Avg % $\text{IMB}_{c,t-1}$) exhibit strong cross-sectional variation.[26] These shares are interacted with 1-year growth rates in national IMB share (excluding county c) to construct a Bartik style instrument, which is based on the premise that national shocks to IMB market share will affect a given county more when their pre-existing IMB market share is larger. If a greater predicted share of recent purchase mortgages in a county are originated by IMBs, there will be greater refinancing activity within the county due to those households receiving aggressive marketing to refinance. Therefore, this will provide exogenous variation in refinancing activity. For robustness, a second version of this instrument based on many exogenous shocks from each large IMB is also implemented. Additional robustness tests are shown at the end of this section to rule out that within-county changes in the instrument stem from reasons related to local business activity.

As a second concern to estimating the causal impact of refinancing activity on business formation rates, increases in refinancing activity may lead to an increase in business formation because of a local demand channel. Households in aggregate extract home equity, not only entrepreneurs, which may lead to a confounding local consumption shock. An increase in local consumption might in turn lead to increased business formation in industries reliant on local demand. To further isolate a causal link driven by the collateral channel, industries that rely on local demand are removed from the calculation of business formation growth rates. First, industries that are non-tradeable based on Mian and Sufi (2014) are removed—retail trade and accommodation/food services (NAICS 44, 45, and 72). Second, firms in the construction, finance, insurance, and real estate industries (NAICS 52 and 53) are also removed (Adelino et al., 2015).

Publicly available data from the Census Statistics of US Businesses (SUSB) are used to calculate measures of local establishment creation. The SUSB is an annual dataset that measures the number of new establishments created at the county by 2-digit NAICS industry

[26]Figure 3 in the Online Appendix shows Avg % $\text{IMB}_{c,t}$ for Los Angeles, Broward, and Suffolk counties. Each county experiences differing growth rates in IMB share over time.

level. To estimate the causal effect, weighted 2SLS are estimated as follows:

$$\Delta ln(\#\text{ New Establishments}_{c,t+1}) = \alpha + \beta\Delta ln(\widehat{\#\text{ Extracted}}_{c,t+1}) + \omega_c + \theta_{t+1} + \zeta\mathbb{Z}_{c,t} + \epsilon_{c,t+1} \tag{1.8}$$

where ω_c and θ_{t+1} are county and annual time fixed effects, respectively, which are included to isolate the effect within-county while controlling for common annual shocks. $\mathbb{Z}_{c,t}$ is a vector of time-varying county-level non-parametric growth rate controls (between $t-1$ and t), including: home prices, number of home purchases, unemployment rate, small business loan volume, and the total number of establishments.[27] The sample covers 1,493 counties (c) for the 8-year period between 2009 and 2016. t starts in 2009 in order to focus on the post-crisis period during which refinancing and business formation rates have stagnated. Standard errors are clustered at the county-level and the regressions are weighted by the county's population from the 2010 Census.

Results in columns 1 through 7 of Table 1.17 are based on growth rates of business formation that exclude industries reliant on local demand. Columns 1 and 2 show that greater refinancing activity is associated with greater business entry (column 2 adds the vector of time varying county-level controls $\mathbb{Z}_{c,t}$) using WLS. To test the strength of the instrument, column 3 estimates the first stage:

$$\Delta ln(\#\text{ Extracted}_{c,t+1}) = \alpha + \tau\text{Avg }\%\text{ IMB}_{c,t-1} * \Delta\%\text{ IMB}_{-c,t} + \omega_c + \theta_{t+1} + \zeta\mathbb{Z}_{c,t} + \epsilon_{c,t+1} \tag{1.9}$$

The first stage isolates variation within county for cash-out refinancing activity based on predicted IMB market share, while controlling for national shocks. If IMB market share positively affects cash-out refinancing activity within county, τ will be positive and significant. Column 3 finds that a 1 point increase in IMB share increases refinancing activity by 0.25 points. The instrument is significant at the 1% level and produces an F-statistic over 10, indicating that it is not a weak instrument.

Column 4 estimates the second stage. The coefficient from weighted 2SLS is larger than the coefficient estimated with WLS, likely a result of the measurement error discussed earlier. A 10% increase in refinancing activity causes an increase in business entry growth rates of 2.5%. For robustness, column 5 shows that the coefficient is similar when the explanatory variable of growth in the number of households extracting home equity is replaced with growth in the dollar amount of home equity extracted.[28] If the dependent variable is replaced with growth in establishment entry rates (the ratio of the number of entering establishments to the population of establishments) the result is unaffected (column 6). To show that the

[27] County-level home price data are from FHFA. Volume of home purchases data are calculated from HMDA. Unemployment rate data are from BLS LAU. Small business loan volume data are from the CRA and restricted to loan amounts <$100,000. Data on the total number of establishments is from the Census SUSB.

[28] Growth in the dollar amount of home equity extracted mixes effects from changes in the number of households extracting home equity and the amount households extract.

result generalizes, the coefficient is shown to be largely unaffected when only businesses in the non-tradeable sectors are excluded (column 8) and when business from every industry are included (column 9). Since the coefficient is stable when the regression model is estimated for all industries and also for industries that are not reliant on local demand, the effect is from the collateral channel (as opposed to the demand channel from an increase in demand for businesses).

The coefficient's magnitude is large. A one standard deviation shock to cash-out refinancing activity (0.53) would cause an increase in business entry growth rates of 11%. This 53% shock would increase annual cash-out refinancing activity by $65 billion, which would lead to a recovery in cash-out refinancing volume to levels seen in 2002, a period of normal lending standards prior to the housing bubble. Although the recovery in refinancing activity is modest, the 11% increase in business entry growth rates from this shock would recover one-third of the decline in business entry rates experienced since 2006.

A possible concern with the instrument is that a shift towards non-banks for mortgage lending might increase small business lending by banks. If banks hold their originated mortgages on balance sheet and do not securitize the mortgages then bank mortgage originations can potentially reduce non-mortgage lending. As IMBs increase their market share of mortgages, banks may as a result have increased lending capacity for non-mortgage loans and subsequently originate more business loans. This in turn may stimulate business formation rates. Given the extent to which banks securitize their mortgage originations, this is likely not a concern. To test that small business lending is not affected by IMB share, the effect of the instrument on 1-year growth rates in small business lending is estimated using WLS (county and time fixed effects and $\mathbb{Z}_{c,t}$ [excluding lagged small business lending volume] are included). The regression shows that the instrument is not correlated with changes in small business lending volume (column 1, Table 1.18).

Another concern is if within-county changes in IMB market share are correlated with changes in the unemployment rate, share of residents who are white, or share of residents with less than a high school education. Buchak et al. (2018) find that IMB share is higher in the cross-section for counties with larger values for these variables. Their finding is cross-sectional, although it is possible that IMB share within a county evolves with changes in these variables, which would subsequently affect business entry growth rates. To show that this is not a concern, a WLS regression is estimated to test if 1-year county-level growth rates in these variables forecast predicted IMB share. Columns 2 to 4, Table 1.18 show that there is no relationship between predicted IMB share and changes in these variables, thereby alleviating this concern.

The robustness tests attempted to rule out potential issues with the instrument. As an additional test, the instrument is transformed to exploit exogenous variation from the shocks in case endogeneity remains a concern. The instrument uses national shocks to overall IMB market share, excluding county c, so one exogenous shock is utilized for each county c. With many exogenous shocks, the estimated effect for β is consistent even if the shares are endogenous (Borusyak, Hull, and Jaravel, 2018; Goldsmith-Pinkham, Sorkin, and Swift, 2018). To transform the instrument to many exogenous shocks, for each county c, national growth rates in market share for each IMB (excluding county c) are calculated. Only large IMBs, those that operate in at least 100 counties, are included to minimize the chance that growth rates in IMB market share are correlated with local economic activity.[29]

Each IMB market share growth rate is weighted by the IMB's ex-ante share of total purchase mortgage origination volume in county c in year $t-1$:

$$\sum_{j\epsilon\{IMB_{c,t-1}\}} \omega_{c,t-1} * \Delta\% \ j_{-c,t}, \text{ where } \omega_{c,t-1} = \frac{\$ \text{ Purchase Origination } j_{c,t-1}}{\$ \text{ Purchase Origination}_{c,t-1}} \tag{1.10}$$

where the sum is indexed over the set of all large IMBs that operate in county c in year $t-1$ and $\Delta\% \ j_{-c,t}$ is the national growth rate of IMB j's market share excluding county c between years $t-1$ and t. Even in the presence of endogenous lagged shares, the many exogenous shocks will allow for a consistent estimation of β. With this instrument, the estimated coefficient does not noticeably change (column 7, Table 1.17). In results available upon request, the standard error approach from Adao, Kolesár, and Morales (2019), which accounts for correlation across counties with similar IMB shares, is shown to not reduce the significance of the result.

1.5 Conclusion

Using novel data, this paper provides evidence that both entrant and continuing small businesses are negatively impacted by credit constraints. At entry, positive credit shocks cause small businesses to start larger, grow faster, permanently remain larger, and have a greater chance of survival. Similarly, continuing businesses immediately respond when their credit constraints are loosened and permanently remain larger. Credit constraints have a permanent effect on small businesses, highlighting the economic benefits of alleviating credit constraints earlier in a firm's life. In terms of direct and immediate employment effects, one additional job is created from a positive credit shock of \$88,000 to entrant firms and \$153,000 to continuing firms.

To alleviate these credit constraints, businesses owners have historically heavily relied on their personal housing equity. In recent years, this channel of credit has almost entirely disappeared, with only the most creditworthy entrepreneurs able to tap into their personal home equity. A tightening of credit standards since 2008 has been an important reason for the lack of recovery in business formation rates since the 2008 financial crisis. While home prices have recovered, personal home equity extraction has not recovered due to more stringent lending standards, which in turn has led to fewer small businesses being started.

The results in this paper shed light on another avenue through which disruptions in lending since the 2008 financial crisis have affected the economy. While restrictions on lending

[29] In the simplest of cases, if an IMB operates in one county then growth rates in that IMB's share will be related to local economic conditions. As the number of counties that the IMB operates in increases, the smaller the relation that growth rates will have to local economic conditions.

to businesses and households have been studied separately, they have not been studied jointly. The strong and previously underappreciated role that mortgage credit had on small business formation prior to the 2008 crisis highlights the need to understand the implications that mortgage credit access has on business formation. While there has been a decline in firm entry rates for decades, the decline since 2006 was perhaps the sharpest in recent history. This paper shows that a tightening of mortgage credit availability to potential entrepreneurs is an important reason for this decline. A recovery of cash-out refinancing activity to its level in 2002 would recover one-third of the decline in firm entry rates since 2006. While tighter mortgage standards in the aftermath of the Great Recession are often viewed as net positive, the negative externality that this can have on the creation, size, and strength of small businesses is overlooked.

1.6 Figures and Tables

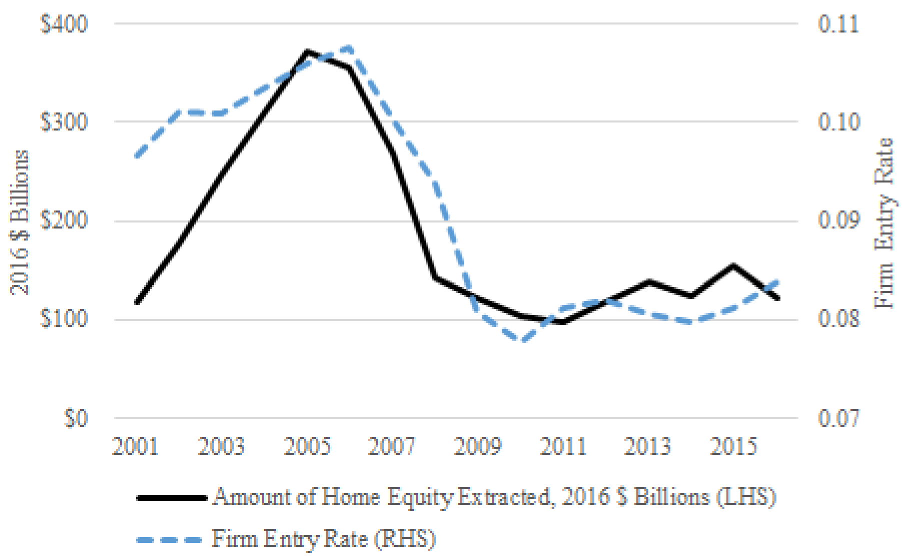


Figure 1.1: Home Equity Extraction and Firm Entry Rates

The solid line shows volume of home equity extracted by year between 2001 and 2016, constructed from ATTOM (in billions of dollars adjusted to 2016 prices). The dashed line reports firm entry rates from the Census Business Dynamic Statistics data. Firm entry rate is calculated as the ratio of the number of private sector firms created in a given year to the number of all active private sector firms in the respective year (Pugsley and Șahin, 2018).

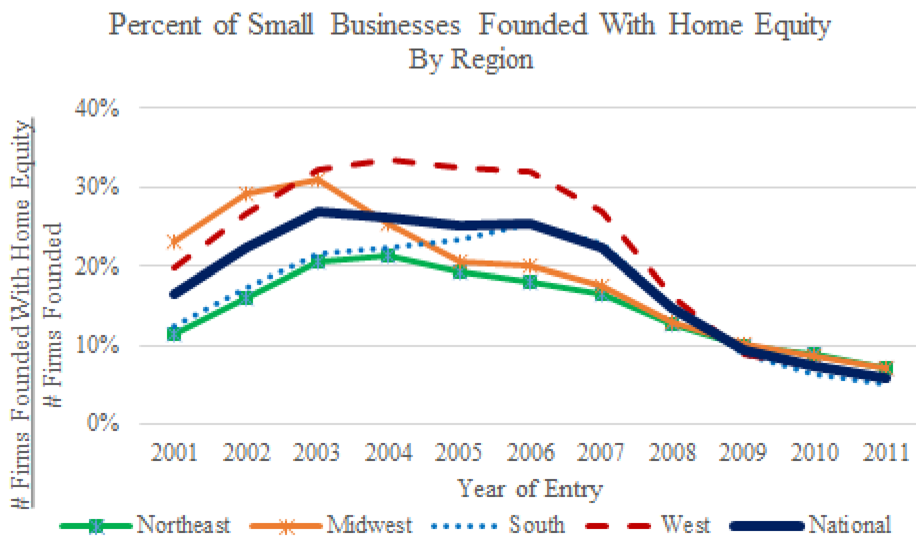


Figure 1.2: Share of Small Businesses Founded with Home Equity Funding By Region

The share of entrant small business funded by personal home equity by year of formation. The data are constructed from a merge of ATTOM to NETS. Small businesses are defined as having ten or fewer employees and being single-unit firms at entry. A business is classified as being funded by home equity if the owner extracts over $5,000 of home equity in the year that the business is created or the prior year. Regions follow the Census region classification. The raw underlying data only includes business owners who own a home. To correct for this, the time series are adjusted by the home ownership rate of business owners based on the population of business owners by region and year from the American Community Survey micro-data.

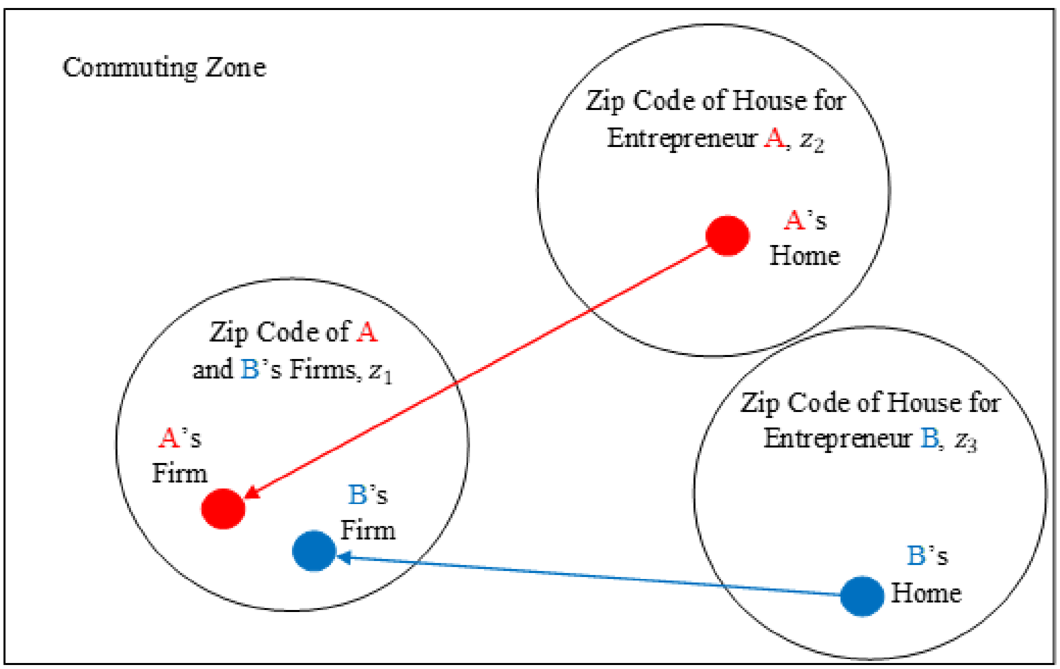
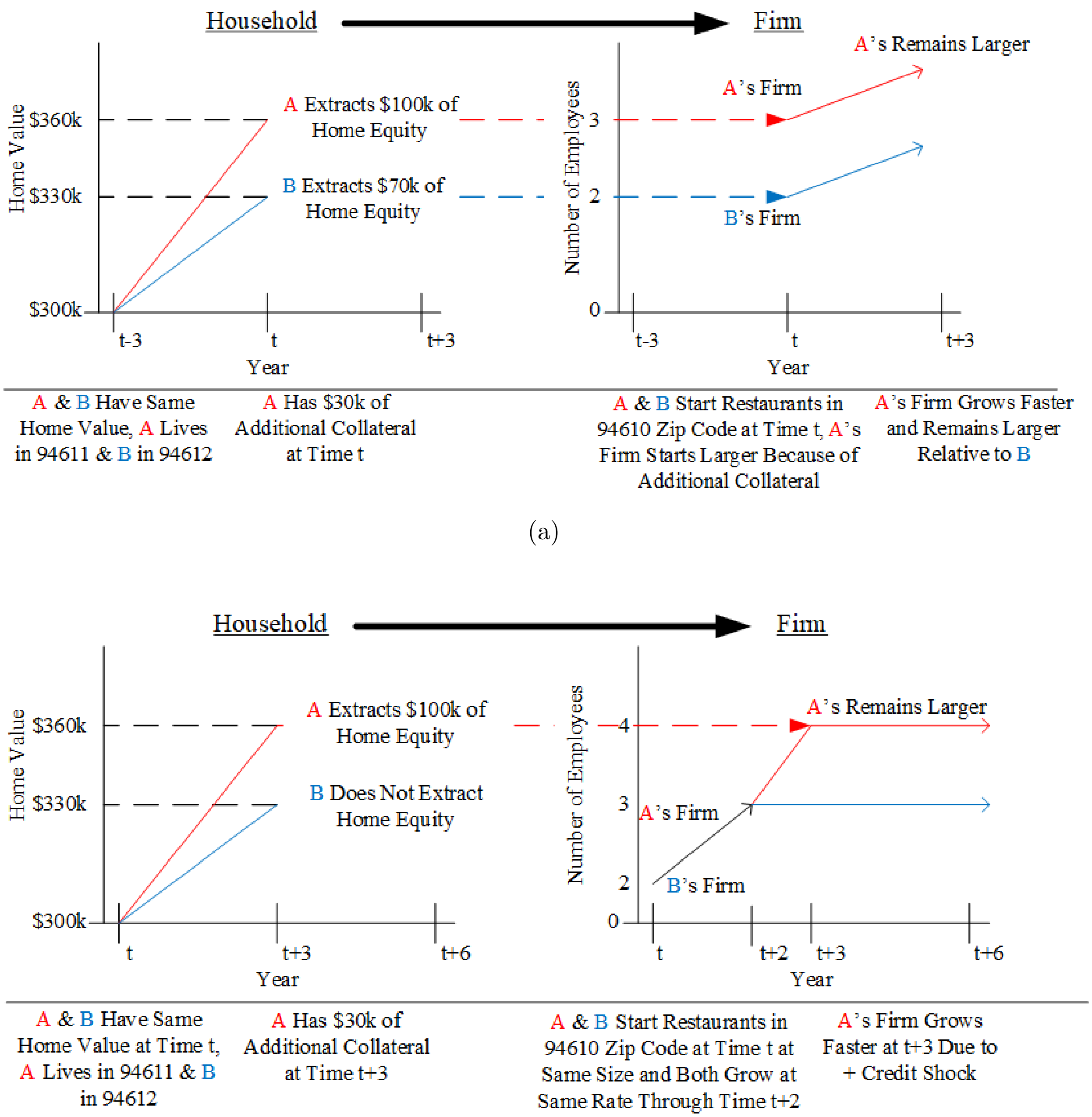


Figure 1.3: Identification Setup

Illustration of identification setup based on location matching.



(a)

(b)

Figure 1.4: Identification Illustration

Figures A and B show the ideal set-up for measuring exogenous credit shocks to entrant and continuing businesses, respectively.

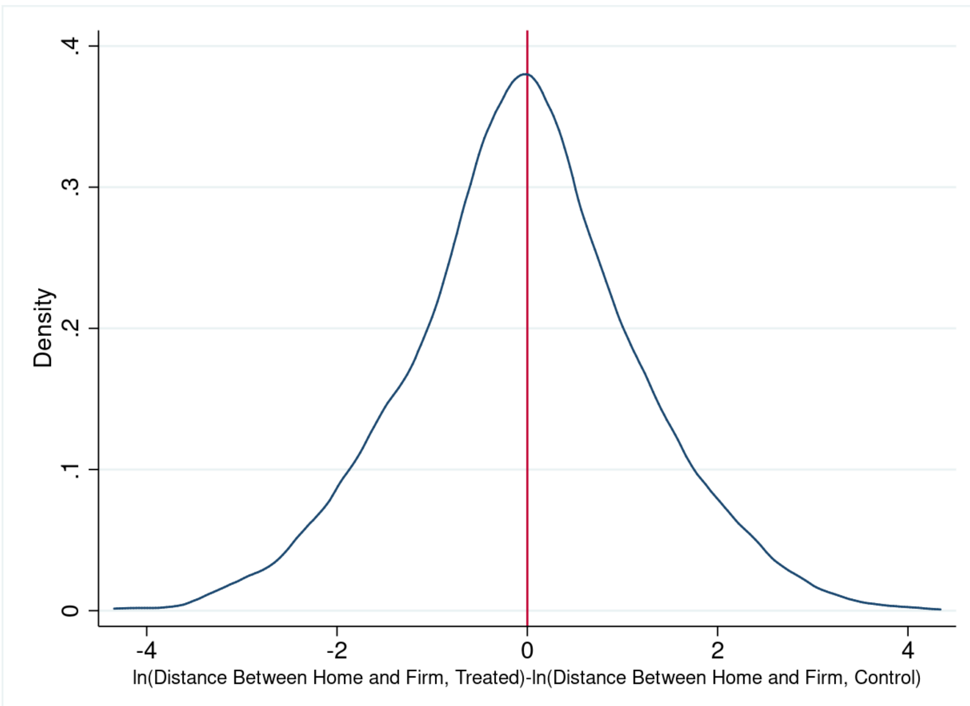


Figure 1.5: Difference in Distance from Home to Firm Between Treated and Control

Density plot of the ln difference between entrant treated and control firms for the distance between a firm owner's home and firm. The entrant matched pairs for this plot are constructed from the ATTOM to NETS merged sample with the same criteria to form matched pairs from the sample merged to LBD for entrant firms. The mean difference is -0.0338, with a standard deviation of 1.191.

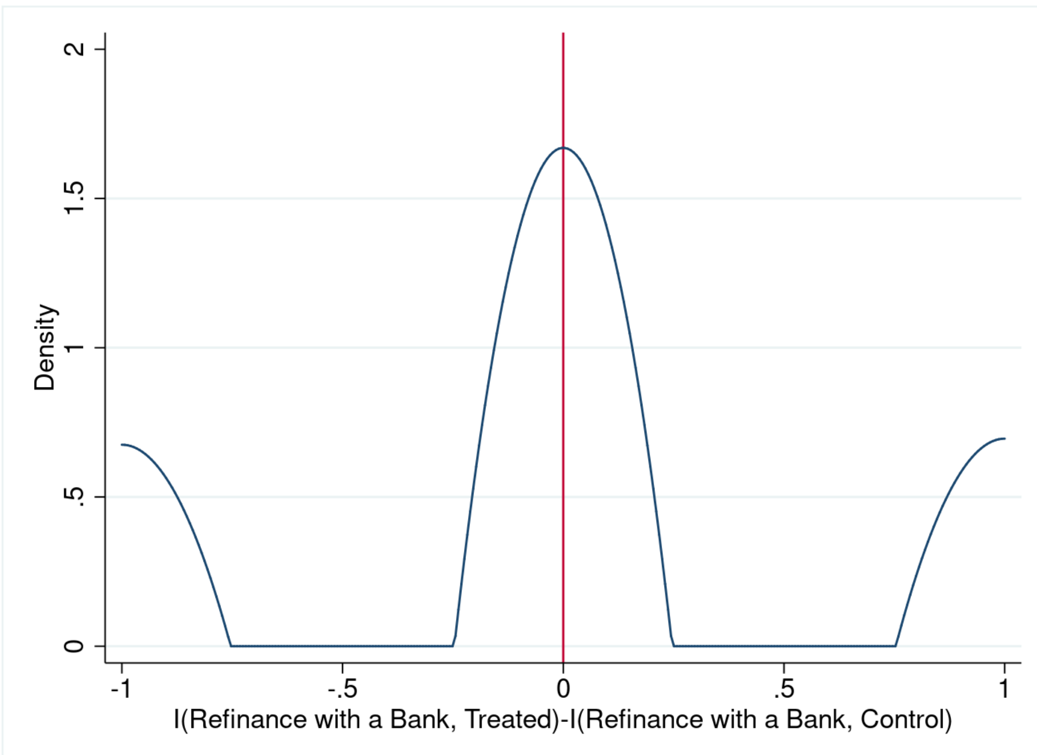


Figure 1.6: Difference in Propensity to Refinance With a Traditional Bank Between Treated and Control

Density plot of the difference between the entrant treated and control firm owners' propensity to originate their cash-out refinance mortgage through a traditional bank (=1 if traditional bank). The entrant matched pairs for this plot are constructed from the ATTOM to NETS merged sample with the same criteria to form matched pairs from the sample merged to LBD for entrant firms. The mean difference is 0.007, with a standard deviation of 0.671.

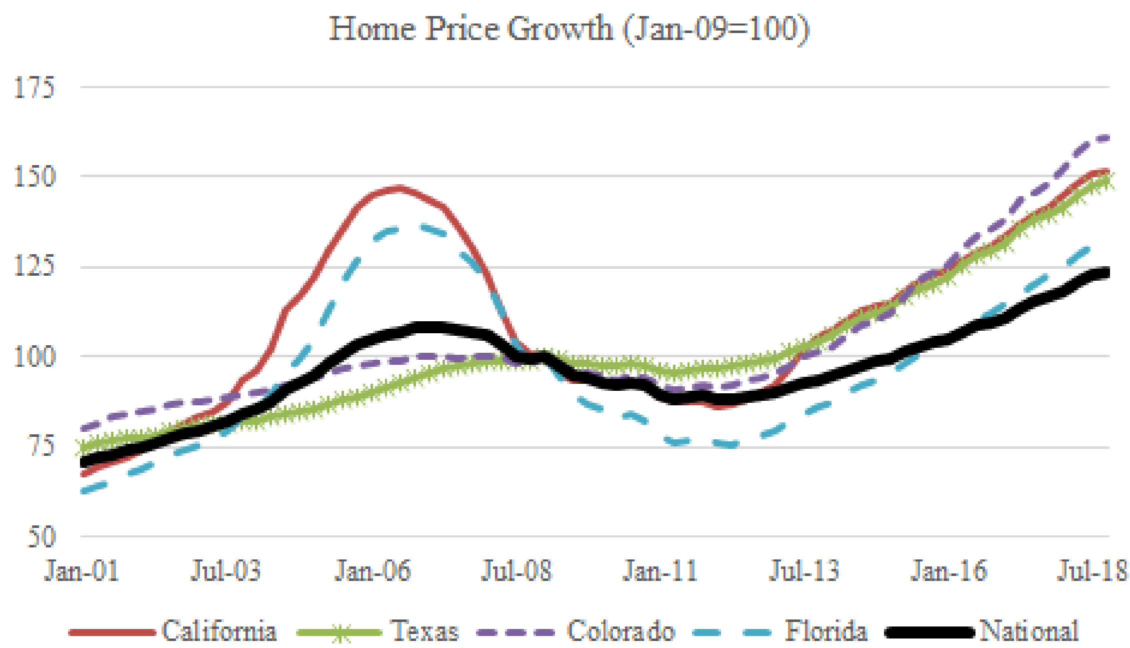


Figure 1.7: Home Price Growth Example

Quarterly all transaction (purchase and refinance) home price growth between January 2000 and December 2018 for California, Florida, Texas, Colorado, and nationally. Data are from FRED. Values are normalized to 100 in January 2009.

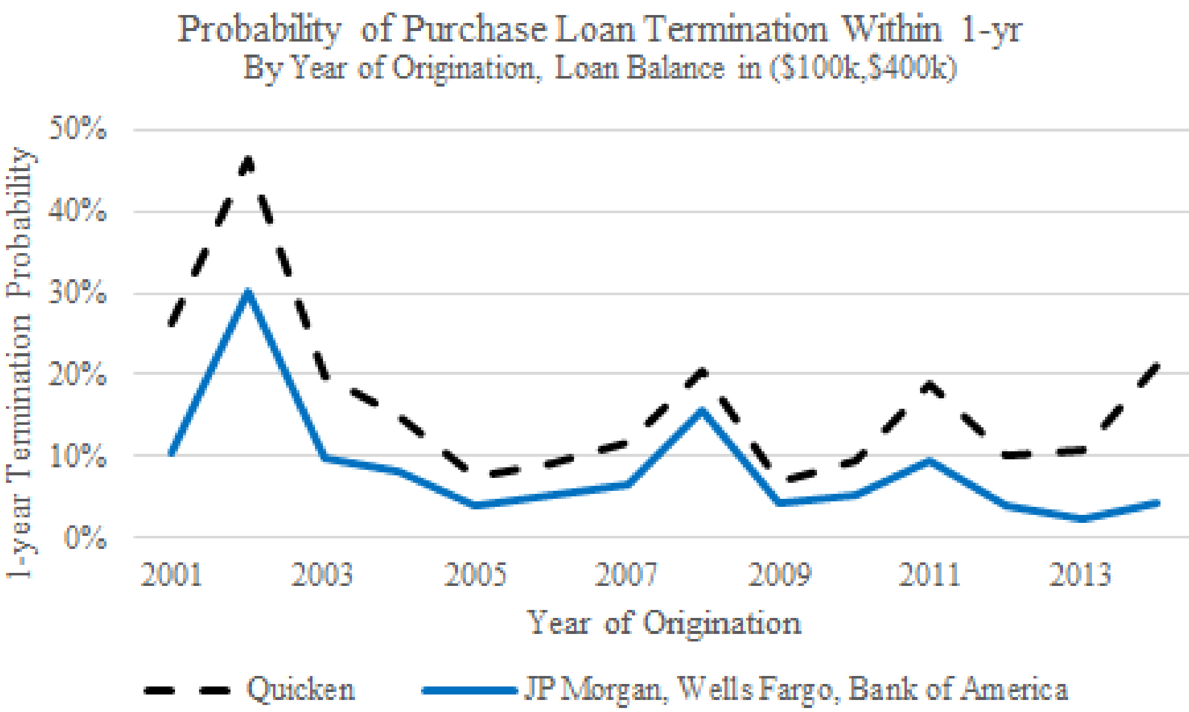


Figure 1.8: Loan Termination Rates

Probability of loan termination within one year of origination. Data are from a merge of loan level public use Freddie Mac and Fannie Mae data to ATTOM. The data are restricted to purchase mortgages with an origination balance between $100k and $400k.

Table 1.1: Summary Statistics for LBD

	Small Businesses in LBD								
	All			Merged to All Datasets			Did Not Merge		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Initial Employment	5,656,000	2.37	2.247	461,000	2.573	2.228	5,195,000	2.352	2.248
Initial Payroll ($000s)	5,656,000	57.75	87.07	461,000	78.51	102.4	5,195,000	55.91	85.34
Survived to Year 5	5,656,000	0.4888	-	461,000	0.6386	-	5,195,000	0.4755	-

Summary statistics for small businesses in the restricted use Longitudinal Business Database (LBD) that were founded between 2001 and 2011. Note, this only includes employer businesses. Due to restrictions on disclosure, only rounded sample counts, mean, and standard deviation are shown for select variables. Small businesses are defined as having ten or fewer employees and being single-unit firms at entry. A business is counted as merged to all datasets if it was successfully merged to: the SSEL, NETS, ATTOM, and zip code level data from Zillow and the 2000 decennial Census.

Table 1.2: Summary Statistics of Entrant Business Financing Sources

	Census Surveys						Robb and Robinson (2014), Kauffman Firm Survey	This Paper
Year Founded	Survey	Home Equity	Bank Loan	Savings	Credit Card	Family	Personal Bank Loan (Includes Home Equity)	Extracted Home Equity in Year of or Year Prior to Business Creation
2003	2007 Survey of	13%	16%	76%	20%	4%	-	27%
2004	Business	14%	15%	75%	18%	4%	16%	26%
2005	Owners (SBO,	15%	16%	74%	20%	5%	-	25%
2006	Survived Until	14%	16%	73%	19%	5%	-	25%
2007	2007)	12%	15%	73%	18%	5%	-	22%
2013-2014	Annual Survey	6%	12%	69%	-	5%	-	-
2014-2015	of	6%	13%	72%	-	6%	-	-
2015-2016	Entrepreneurs	5%	11%	70%	-	5%	-	-

Summary statistics on sources of start-up financing from: 2007 Survey of Business Owners (SBO), the 2013-2014, 2014-2015, and 2015-2016 waves of the Annual Survey of Entrepreneurs (ASE), the 2004 Kauffman Family Survey based on results in Robb and Robinson (2014), and the NETS-ATTOM merged dataset utilized in this paper. The 2007 SBO includes businesses that survived through 2007. The ASE is based on surveys in 2014, 2015, and 2016 and provides statistics for businesses founded within the two year period before the survey. The SBO data are restricted to firms with between 1 and 25 employees, the amount of startup capital reported, and with owners who founded the business. Tabulation weights are used. For the ASE, these restrictions are not possible as a micro-data set is not available. The Kauffman Firm Survey is based on firms of all sizes, but is primarily comprised of firms with ten or fewer employees. Robb and Robinson (2014) report 16% utilize personal bank loans (which includes home equity) as initial funding capital. This statistic restricts to firms that do not have missing survey results over the 2004 to 2007 period (or were known to have gone out of business). The NETS-ATTOM merged dataset utilized in this paper is restricted to businesses with ten or fewer employees initially and to business owners who own a home. To correct for this last restriction, the ratios from the NETS-ATTOM population are adjusted by the home ownership rate for the population of business owners each year between 2003 and 2007 using micro-data from the American Community Survey (ACS).

Table 1.3: Summary Statistics Between Treated and Control Members for Entrant Business Matching

	(1) N	(2) Mean	(3) SD
LN Difference in Home Value at First Year-3	2,500	-0.01565	0.261
LN Difference in CLTV at Purchase	2,500	0.001952	0.05182
LN Difference in Number of Months from Purchase Until First Year of Firm	2,500	0.003815	0.5278
LN Difference in Median Home Zip Code Income	2,500	-0.07032	0.3436
Difference in Home Zip Code Percent of Residents Who Are White	2,500	-0.02046	0.205
Difference in Home Zip Code Percent of Households Below Poverty Line	2,500	0.01142	0.0714
Difference in Home Zip Code Percent of Households Who Rent	2,500	0.02711	0.2038
Ratio of Lagged 3-Year Home Zip Code HPI Growth	2,500	0.06525	0.06252

Summary statistics at the pair level for entrant small business pairs created based on the entrant small business matching algorithm. Pairs are formed from exact matches on the zip code of the business, the zip codes of the homes being different from each other (and the businesses), SIC division, and year of business creation. The pairs are further restricted to cases where the home values (as measured three years prior to business creation year) are within 20% or $100k (for less valuable homes) of each other and the CLTV ratios at purchase are within 20bps of each other. The treated firm is the one that experienced greater home price growth. For each treated firm, a control firm that has the most similar home value as the treated firm is selected after the matching exercise. Small businesses are defined as having ten or fewer employees and being single-unit firms at entry. Firms in the sample were founded between 2001 and 2011. Rounded sample size, mean, and standard deviation are reported for the differences in variables between the treated and control member of each pair.

Table 1.4: Summary Statistics of Entrant Business Matched Sample

	(1) N	(2) Mean	(3) SD
Initial Employment	4,600	2.778	2.32
Initial Payroll ($000s)	4,600	$94.72	$125
Survived to Year 1	4,600	0.9494	-
Survived to Year 3	4,600	0.7899	-
Survived to Year 5	4,600	0.6732	-
Amount of Home Equity Extracted	4,600	$100,800	$102,200
Home Zip Code 2000 Median Family Income	4,600	$64,330	21330
Home Zip Code Percent White	4,600	0.7474	0.1833
Home Zip Code Percent Below Poverty Line	4,600	0.08752	0.06056
Home Zip Code Percent Renter	4,600	0.3093	0.1564
CLTV at Purchase	4,600	0.8716	0.1648
Home Value at Purchase ($000s)	4,600	$206.8	122.1
3-year HPI Growth Before Firm Creation	4,600	1.507	0.2728
# Months Between Home Purchase and Firm Creation	4,600	76.87	34.18

Summary statistics at the business level for entrant small business that were matched to another business. The matched entrant firm sample is constructed as small businesses that were funded with home equity at entry, where the entrepreneur lives in a different zip code from the business, and where the home was purchased at least three years prior to the firm entry year. Pairs are formed from exact matches on the zip code of the business, the zip codes of the homes being different from each other (and the businesses), SIC division, and year of business creation. The pairs are further restricted to cases where the home values (as measured three years prior to business creation year) are within 20% or $100k (for less valuable homes) of each other and the CLTV ratios at purchase are within 20bps of each other. The treated firm is the one that experienced greater home price growth. For each treated firm, a control firm that has the most similar home value as the treated firm is selected after the matching exercise. Small businesses are defined as having ten or fewer employees and being single-unit firms at entry. Firms in the sample were founded between 2001 and 2011 and each firm is only included once even if matched multiple times. Rounded sample size, mean, and standard deviation are reported.

Table 1.5: Effect of Credit Access on Entrant Business Survival

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Survival							
	5-year OLS, Only Pair FE	5-year OLS	5-year 1st Stage	5-year 2SLS	5-year 2SLS	5-year 2SLS	1-year 2SLS	3-year 2SLS
ln($ Amount Extracted$_{i,t}$)	0.0551***	0.0532***		0.513**	0.513*	0.513**	0.202*	0.712***
	(3.02)	(2.89)		(2.19)	(1.88)	(2.24)	(1.73)	(2.69)
$\Delta_3 ln$(Zip Code Home Price$_{i,t}$)			0.882***					
			(3.21)					
# Obs	5100	5100	5100	5100	5100	5100	5100	5100
R-squared	0.525	0.553	0.655	0.3	0.3	0.3	0.332	-0.187
F-statistic	-	-	10.32	-	-	-	-	-
Cluster firm zip*SIC division?	Y	Y	Y	Y	N	N	Y	Y
Cluster commuting zone*SIC division?	N	N	N	N	Y	N	N	N
Cluster firm zip*SIC division*firm entry year?	N	N	N	N	N	Y	N	N

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if entrant small businesses have differential survival outcomes based on initial credit access:

$$\mathbb{I}(\text{Survived}_{i,j,t+k}) = \alpha + \beta ln(\widehat{\text{\$ Amount Extracted}_{i,t}}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,j,t+k}$$

Firms are denoted by i, firm creation year by t, and the pair that the firm belongs to by j. $\mathbb{X}_j$ are fixed effects for each pair. The regressions are a linear probability model with $\mathbb{I}(\text{Survived}_{i,j,t+k})$ equal to one if the business survived at least k years. ln($ Amount Extracted$_{i,t}$) is the amount of home equity that the entrepreneur extracted in the year that the business is created or the prior year. Controls$_i$ is a vector of zip code level and firm level controls that includes: log of CLTV at purchase, log of home value at year $t-3$, 2-digit SIC industry fixed effects, the log of the number of months between home purchase and year t, legal form of organization (LFO) fixed effects, and non-parametric controls for home zip code characteristics of percent white, percent renter, percent below poverty line, and median income. Standard errors are clustered at the firm zip code by SIC division level. The regressions are estimated on the entrant matched pair sample and includes small businesses that were initially funded with home equity and were founded between 2001 and 2011. Columns 1 and 2 estimate the effect on 5-year survival rate using OLS (only controlling for pair fixed effects in column 1). Column 3 estimates the first stage with the instrument being lagged log 3-year home zip code level home price growth from Zillow. Columns 4 through 6 estimate the causal impact on 5-year survival. For robustness, column 5 clusters at the commuting zone by SIC division level and column 6 clusters at the firm zip code by SIC division by firm creation year level. Columns 7 and 8 estimate the effect on 1-year and 3-year survival, respectively.

Table 1.6: Effect of Credit Access on Entrant Business Employment

	(1)	(2)	(3)	(4)	(5)	(6)
			ln(Employment)			
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 5
	2SLS	2SLS	2SLS	2SLS	2SLS	OLS
ln(\$ Amount Extracted$_{i,t}$)	0.488*	1.380**	1.458**	1.421**	1.241**	0.111***
	(1.84)	(2.39)	(2.24)	(2.21)	(2.00)	(2.75)
# Obs	5100	4300	4300	4300	4300	4300
R-squared	0.4	-0.141	-0.079	0.0561	0.223	0.57

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if entrant small businesses have differential employment based on initial credit access:

$$ln(\text{Employment}_{i,j,t+k}) = \alpha + \beta ln(\widehat{\text{\$ Amount Extracted}_{i,t}}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,j,t+k}$$

Firms are denoted by i, firm creation year by t, and the pair that the firm belongs to by j. $\mathbb{X}_j$ are fixed effects for each pair. The outcome variables are the log of employment in years 1 through 5. ln(\$ Amount Extracted$_{i,t}$) is the amount of home equity that the entrepreneur extracted in the year that the business is created or the prior year. Controls$_i$ is a vector of zip code level and firm level controls that includes: log of CLTV at purchase, log of home value at year $t-3$, 2-digit SIC industry fixed effects, the log of the number of months between home purchase and year t, legal form of organization (LFO) fixed effects, and non-parametric controls for home zip code characteristics of percent white, percent renter, percent below poverty line, and median income. Standard errors are clustered at the firm zip code by SIC division level. The regressions are estimated on the entrant matched pair sample and includes small businesses that were initially funded with home equity and were founded between 2001 and 2011. For employment in years 2 through 5 (columns 2 through 6) the sample is restricted to pairs where both members never had missing employment/payroll data in the LBD within their first five years (if a firm is in multiple pairs and one of the pairs is excluded then all of their other pairs are excluded as well). Columns 1 through 5 estimate the effect on employment in years 1 through 5, respectively, using 2SLS with the second stage reported. Column 6 estimates the effect on year 5 employment using OLS.

Table 1.7: Effect of Credit Access on Entrant Business Employment Growth Rates

	(1)	(2)	(3)	(4)
	Employment Growth			
	Years 1 to 2	Years 2 to 3	Years 3 to 4	Years 4 to 5
ln($ Amount Extracted$_{i,t}$)	1.809**	0.970	0.207	0.231
	(2.18)	(1.45)	(0.50)	(0.52)
# Obs	4300	4300	4300	4300
R-squared	-0.299	0.398	0.702	0.750

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if entrant small businesses have differential employment growth rates based on initial credit access:

$$\frac{Employment_{i,j,t+k+1} - Employment_{i,j,t+k}}{.5 * (Employment_{i,j,t+k+1} + Employment_{i,j,t+k})} = \alpha + \beta ln(\widehat{\text{\$ Amount Extracted}}_{i,t}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,j,t+k}$$

Firms are denoted by i, firm creation year by t, and the pair that the firm belongs to by j. $\mathbb{X}_j$ are fixed effects for each pair. The outcome variables are 1-year employment growth rates for the business's first four years. ln($ Amount Extracted$_{i,t}$) is the amount of home equity that the entrepreneur extracted in the year that the business is created or the prior year. Controls$_i$ is a vector of zip code level and firm level controls that includes: log of CLTV at purchase, log of home value at year $t-3$, 2-digit SIC industry fixed effects, the log of the number of months between home purchase and year t, legal form of organization (LFO) fixed effects, and non-parametric controls for home zip code characteristics of percent white, percent renter, percent below poverty line, and median income. Standard errors are clustered at the firm zip code by SIC division level. The regressions are estimated on the entrant matched pair sample and includes small businesses that were initially funded with home equity and were founded between 2001 and 2011. The sample is restricted to pairs where both members never had missing employment/payroll data in the LBD within their first five years (if a firm is in multiple pairs and one of the pairs is excluded then all of their other pairs are excluded as well). Columns 1 through 4 estimate the effect on 1-year employment growth in for the businesses first four years, respectively, using 2SLS with the second stage reported.

Table 1.8: Effect of Credit Access on Entrant Business Payroll

	(1)	(2)	(3)	(4)	(5)	(6)
	ln(Payroll)					
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 5
	2SLS	2SLS	2SLS	2SLS	2SLS	OLS
ln(\$ Amount Extracted$_{i,t}$)	0.557	2.685**	3.664**	3.306**	2.972**	0.296***
	(0.92)	(2.13)	(2.25)	(2.13)	(1.96)	(2.98)
# Obs	5100	4300	4300	4300	4300	4300
R-squared	0.566	0.128	-0.0752	0.150	0.278	0.586

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if entrant small businesses have differential payroll based on initial credit access:

$$ln(\text{Payroll}_{i,j,t+k}) = \alpha + \beta ln(\$ \widehat{\text{Amount Extracted}}_{i,t}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,j,t+k}$$

Firms are denoted by i, firm creation year by t, and the pair that the firm belongs to by j. $\mathbb{X}_j$ are fixed effects for each pair. The outcome variables are the log of payroll in years 1 through 5. ln(\$ Amount Extracted$_{i,t}$) is the amount of home equity that the entrepreneur extracted in the year that the business is created or the prior year. Controls$_i$ is a vector of zip code level and firm level controls that includes: log of CLTV at purchase, log of home value at year $t-3$, 2-digit SIC industry fixed effects, the log of the number of months between home purchase and year t, legal form of organization (LFO) fixed effects, and non-parametric controls for home zip code characteristics of percent white, percent renter, percent below poverty line, and median income. Standard errors are clustered at the firm zip code by SIC division level. The regressions are estimated on the entrant matched pair sample and includes small businesses that were initially funded with home equity and were founded between 2001 and 2011. For payroll in years 2 through 5 (columns 2 through 6) the sample is restricted to pairs where both members never had missing employment/payroll data in the LBD within their first five years (if a firm is in multiple pairs and one of the pairs is excluded then all of their other pairs are excluded as well). Columns 1 through 5 estimate the effect on payroll in years 1 through 5, respectively, using 2SLS with the second stage reported. Column 6 estimates the effect on year 5 payroll using OLS.

Table 1.9: Summary Statistics Comparing Entrant Businesses by Initial Funding Source

	(1)	(2)	(3)	(4)	(5)	(6)
	Home Equity Funded			Not Home Equity Funded		
	N	Mean	SD	N	Mean	SD
Initial Employment	124,000	2.661	2.264	337,000	2.541	2.213
Initial Payroll ($000s)	124,000	$81.04	$102.9	337,000	$77.59	$102.2
Survived to Year 1	124,000	0.9453	-	337,000	0.9378	-
Survived to Year 3	124,000	0.7703	-	337,000	0.764	-
Survived to Year 5	124,000	0.6382	-	337,000	0.6387	-
Amount of Home Equity Extracted for Initial Capital	124,000	$100,600	$123,700	33,7000	-	-
Home Zip Code 2000 Median Family Income	124,000	$65,470	$21,660	337,000	$63,480	$21,900
Home Zip Code Percent White	124,000	0.7945	0.1761	337,000	0.7893	0.1871
Home Zip Code Percent Below Poverty Line	124,000	0.08109	0.05934	337,000	0.08664	0.06415
Home Zip Code Percent Renter	124,000	0.2868	0.1561	337,000	0.2956	0.1621
CLTV at Purchase	124,000	0.7768	0.3322	337,000	0.7006	0.3852
Home Value at Purchase ($000s)	124,000	$277.5	$223.7	337,000	$264.9	$230.7
3-year HPI Growth Before Firm Creation	124,000	1.362	0.296	337,000	1.172	0.3339
# Months Between Home Purchase and Firm Creation	124,000	52.21	41.5	337,000	59.58	47.39

Summary statistics comparing home equity funded businesses to non-home equity funded businesses within the sample of small businesses founded between 2001 and 2011 in the LBD that were merged to the other datasets. Small businesses are defined as having ten or fewer employees and being single-unit firms at entry. A business is labeled as home equity funded if the entrepreneur extracted at least $5,000 of personal home equity in the year that the business is created or the prior year. Due to restrictions on disclosure, only rounded sample counts, mean, and standard deviation are shown for select variables.

Table 1.10: Comparing Differences in Businesses and Entrepreneurs By Initial Funding Source

	(1) ln(Purchase Value)	(2) CLTV	(3) $ln(\text{Employment}_t)$	(4) $ln(\text{Payroll}_t)$	(5) 1-year Survival	(6) 3-year Survival	(7) 5-year Survival
$\mathbb{I}(\text{Home Equity Funded}_i)$	0.0535***	0.0397***	0.0161**	0.00954	0.0022	0.00114	-0.000473
	(22.23)	(28.1)	(2.1)	(0.48)	(0.64)	(0.19)	(-0.07)
$ln(\text{Purchase Value}_i)$			0.0295***	0.158***	0.00467**	0.00907**	0.0120***
			(5.69)	(11.66)	(2.01)	(2.3)	(2.73)
CLTV_i			-0.0141	-0.0582**	-0.00559	-0.0171**	-0.0184**
			(-1.37)	(-2.17)	(-1.2)	(-2.16)	(-2.06)
# Obs	461000	461000	461000	461000	461000	461000	461000
R-squared	0.67	0.45	0.789	0.784	0.753	0.77	0.776

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if home equity funded businesses differ from non-home equity funded businesses:

$$Y_i = \alpha + \beta\mathbb{I}(\text{Home Equity Funded}_i) + \omega\text{Controls}_i + \epsilon_i$$

where i indexes a business. $\mathbb{I}(\text{Home Equity Funded}_i)$ is an indicator variable = 1 if a business is initially funded with home equity (entrepreneur extracted >$5,000 in the year that the business is created or the prior year). The population is the merged LBD to ATTOM dataset and includes small businesses (started with ten or fewer employees and initially single-unit firms) founded between 2001 and 2011. Outcome variables are from ATTOM and the LBD. Columns 1 and 2 use housing information from ATTOM as outcome variables and Controls_i includes fixed effects for the interaction of home zip code by home purchase year, LFO, and 2-digit SIC industry. Standard errors are clustered at the home zip code level. Columns 3 through 7 use data on the businesses from LBD as outcome variables and Controls_i includes log purchase value, CLTV, and fixed effects for the triple interaction of 2-digit SIC by creation year by firm zip code and LFO. Standard errors are clustered at the 2-digit SIC by firm zip code level.

Table 1.11: Comparing Differences in Personal Foreclosure Outcome By Initial Funding Source

	(1) 2001-2011	(2) 2001-2004	(3) 2005-2007	(4) 2008-2011
$\mathbb{I}$(Home Equity Funded$_i$)	0.0156	-0.0101	0.0636***	-0.0239*
	(1.59)	(-0.22)	(4.14)	(-1.82)
# Obs	551887	57501	158062	334300
R-squared	0.880	0.924	0.865	0.881
Overall Default Rate	0.193	0.143	0.248	0.177

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if home equity funded businesses differ from non-home equity funded businesses in foreclosure probability:

$$\mathbb{I}(\text{Foreclosure}_i) = \alpha + \beta\mathbb{I}(\text{Home Equity Funded}_i) + \omega\text{Controls}_i + \epsilon_i$$

where i indexes a business. $\mathbb{I}$(Home Equity Funded$_i$) is an indicator variable = 1 if a business is initially funded with home equity (entrepreneur extracted >\$5,000 in the year that the business is created or the prior year). The population is the merged NETS to ATTOM dataset that is further merged to McDash and includes small businesses (started with ten or fewer employees and initially single-unit firms) founded between 2001 and 2011. Foreclosure is estimated from ATTOM and the dependent variable equals 1 if a business owner has a foreclosure on their home within four years of starting the business. Controls$_i$ is vector of controls that includes: home purchase year*firm creation year*home zip code*firm zip code*2-digit SIC industry FE, purchase value, combined loan to value (CLTV) at purchase, initial firm employment FE, original mortgage interest rate, and FICO at time of home purchase. Standard errors are clustered at the home zip code level. Column 1 is based on the full 2001-2011 period, while columns 2 to 4 segment into the pre-crisis (2001-2004 firm entry years), crisis (2005-2007 firm entry years), and post-crisis (2008-2011 firm entry years) periods, respectively.

Table 1.12: Summary Statistics of Continuing Business Matched Sample

	(1) N	(2) Mean	(3) SD
Initial Employment	11,500	1.711	1.096
Initial Payroll ($000s)	11,500	$62.46	$70
Survived to Event Year+3	11,500	0.7581	-
Amount of Home Equity Extracted in Event Year	11,500	$7,909	$36,350
Home Zip Code 2000 Median Family Income	11,500	$62,800	$20,710
Home Zip Code Percent White	11,500	0.7664	0.1921
Home Zip Code Percent Below Poverty Line	11,500	0.08869	0.06358
Home Zip Code Percent Renter	11,500	0.3057	0.168
CLTV at Purchase	11,500	0.7334	0.3258
Home Value at Purchase ($000s)	11,500	$227.8	$152.9
3-year HPI Growth Before Event Year	11,500	1.088	0.3848
# Months Between Home Purchase and Firm Creation	11,500	59.77	44

Summary statistics of (unique) continuing small businesses in the matched continuing firm sample. The matched continuing firm sample is constructed as small businesses that have survived through either year 3 or 4 (the event year), where the entrepreneur lives in a different zip code from the business, and where the home was purchased at least three years prior to the event year. Pairs are formed from exact matches on the zip code of the business, the zip codes of the homes being different from each other (and the businesses), SIC division, year of business creation, and same initial employment. The pairs are restricted to cases where the home values (as measured three years prior to the event year) are within 20% or $100k (for less valuable homes) of each other and employment one year prior to the event year is within three employees of each other. Within the pair, the treated firm is the one that experienced greater home price growth within the 3-year period prior to the event year. The control firm is selected as the one that has the most similar home value as the treated firm. A business can be a control firm multiple times (and can be both a control and treated firm in different pairs) but can only be a treated firm once for each of the two event years. Small businesses are defined as having ten or fewer employees and being single-unit firms at entry. Due to restrictions on disclosure, only rounded sample counts, mean, and standard deviation are shown for select variables. The amount of home equity extracted in the event year is set to 0 for the businesses that do not extract home equity in the event year.

Table 1.13: Summary Statistics Between Treated and Control Members for Continuing Business Matching

	(1) N	(2) Mean	(3) SD
LN Difference in Home Value	8,900	-0.004887	0.2766
LN Difference in CLTV at Purchase	8,900	-0.004617	0.2975
LN Difference in Number of Months from Purchase Until First Year of Firm	8,900	-0.005421	1.082
LN Difference in Median Home Zip Code Income	8,900	0.0129	0.3745
Difference in Home Zip Code Percent of Residents Who Are White	8,900	0.0042	0.2238
Difference in Home Zip Code Percent of Households Below Poverty Line	8,900	0.001866	0.07808
Difference in Home Zip Code Percent of Households Who Rent	8,900	0.01498	0.2202
Difference in Number of Employees in Year Prior to Event Year	8,900	0.001129	1.413
LN Difference in Payroll in Year 1	8,900	0.01565	1.125
Ratio of Lagged 3-Year Home Zip Code HPI Growth	8,900	0.1016	0.115

Summary statistics comparing differences in variables between the treated and control members of the matched continuing firm sample. The matched continuing firm sample is constructed as small businesses that have survived through either year 3 or 4 (the event year), where the entrepreneur lives in a different zip code from the business, and where the home was purchased at least three years prior to the event year. Pairs are formed from exact matches on the zip code of the business, the zip codes of the homes being different from each other (and the businesses), SIC division, year of business creation, and same initial employment. The pairs are restricted to cases where the home values (as measured three years prior to the event year) are within 20% or $100k (for less valuable homes) of each other and employment one year prior to the event year is within three employees of each other. Within the pair, the treated firm is the one that experienced greater home price growth within the 3-year period prior to the event year. The control firm is selected as the one that has the most similar home value as the treated firm. A business can be a control firm multiple times (and can be both a control and treated firm in different pairs) but can only be a treated firm once for each of the two event years. Small businesses are defined as having ten or fewer employees and being single-unit firms at entry. Due to restrictions on disclosure, only rounded sample counts, mean, and standard deviation are shown for select variables.

Table 1.14: Effect of Credit Access on Continuing Business Survival

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Survival to $t+p+3$				
	OLS	OLS	1st Stage	2SLS	2SLS	2SLS	2SLS
ln($ Amount Extracted$_{i,t+p}$)	-0.000121	-0.000168		-0.00533	0.00669	0.00669	0.00669
	(-0.08)	(-0.11)		(-0.10)	(0.13)	(0.13)	(0.12)
$\Delta_3 ln$(Zip Code Home Price$_{i,t+p}$)			1.162***				
			(3.76)				
# Obs	17500	17500	17500	17500	17500	17500	17500
R-squared	0.524	0.532	0.543	0.523	0.531	0.531	0.531
F-Statistic	-	-	14.11	-	-	-	-
Include controls for initial payroll, payroll/employment growth?	N	Y	Y	N	Y	Y	Y
Cluster on Firm Zip*SIC Division?	Y	Y	Y	Y	Y	N	N
Cluster on Firm Zip*SIC Division*Creation Year	N	N	N	N	N	Y	N
Cluster on Commuting Zone*SIC Division	N	N	N	N	N	N	Y

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if continuing small businesses have differential survival outcomes based on the amount of home equity extracted:

$$\mathbb{I}(\text{Survived}_{i,j,t+p+3}) = \alpha + \beta ln(\$ \text{ } \widehat{\text{Amount Extracted}}_{i,t+p}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,j,t+p+3}$$

Firms are denoted by i, firm creation year by t, the pair that the firm belongs to by j, p is the number of years since founding t that home equity is extracted for expansion (either two or three, which corresponds to the firms of age three or four). $\mathbb{X}_j$ are fixed effects for each pair. The regressions are a linear probability model with $\mathbb{I}(\text{Survived}_{i,j,t+p+3})$ equal to one if the business survived through year $t+p+3$. ln($ Amount Extracted$_{i,t+p}$) is the amount of home equity that the entrepreneur extracted in year $t+p$ (set to 0 if home equity is not extracted). Controls$_i$ is a vector of zip code level and firm level controls that includes: log of CLTV at purchase, log of home value at year $t+p-3$, 2-digit SIC industry fixed effects, the log of the number of months between home purchase and year t, legal form of organization (LFO) fixed effects, fixed effect for if the business was initially funded with home equity, and non-parametric controls for home zip code characteristics of percent white, percent renter, percent below poverty line, and median income. Depending on the regression, additional controls include non-parametric fixed effects estimated by 2-digit SIC for initial payroll and employment and payroll growth between years 1 and 2. Standard errors are clustered at the firm zip code by SIC division level. The regressions are estimated on the continuing firm matched pair sample and includes small businesses that survived through the event year $t+p$. Columns 1 and 2 estimate the effect using OLS, column 1 does not include the non-parametric controls for initial payroll and employment/payroll growth between years 1 and 2. Column 3 estimates the first stage with the instrument being lagged log 3-year home zip code level home price growth from Zillow. Columns 4 and 5 estimate the causal impact on survival, column 4 does not include the non-parametric controls for initial payroll and employment/payroll growth between years 1 and 2. For robustness, columns 6 and 7 replicate column 5 with clustering at the firm zip code by SIC division by firm year level and commuting zone by SIC division level, respectively.

Table 1.15: Effect of Credit Access on Continuing Business Employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			ln(Employment)			Employment Growth Rates	
	$t+p-1$	$t+p$	$t+p+1$	$t+p+2$	$t+p+3$	$t+p-1$ to $t+p+1$	$t+p+1$ to $t+p+3$
ln(\$ Amount Extracted$_{i,t+p}$)	-0.0383	0.114**	0.173**	0.149*	0.159*	0.208*	0.0576
	(-0.99)	(2.10)	(2.20)	(1.70)	(1.69)	(1.73)	(0.48)
# Obs	17500	17500	17500	17500	17500	17500	17500
R-squared	0.727	0.538	0.342	0.435	0.412	0.315	0.53

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if continuing small businesses have differential employment outcomes based on the amount of home equity extracted:

$$ln(\text{Employment}_{i,j,t+p+k}) = \alpha + \beta ln(\widehat{\text{\$ Amount Extracted}}_{i,t+p}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,j,t+p+k}$$

Firms are denoted by i, firm creation year by t, the pair that the firm belongs to by j, p is the number of years since founding t that home equity is extracted for expansion (either two or three, which corresponds to the firms of age three or four). $\mathbb{X}_j$ are fixed effects for each pair. The outcome variables are log employment k years around the event year $t+p$, where k ranges from -1 to 3. ln(\$ Amount Extracted$_{i,t+p}$) is the amount of home equity that the entrepreneur extracted in year $t+p$ (set to 0 if home equity is not extracted). Controls$_i$ is a vector of zip code level and firm level controls that includes: log of CLTV at purchase, log of home value at year $t+p-3$, 2-digit SIC industry fixed effects, the log of the number of months between home purchase and year t, legal form of organization (LFO) fixed effects, fixed effect for if the business was initially funded with home equity, non-parametric fixed effects estimated by 2-digit SIC for initial payroll, and non-parametric controls for home zip code characteristics of percent white, percent renter, percent below poverty line, and median income. Except for columns 1 and 6, whose outcome variables are based on data from year $t+p-1$, controls also include non-parametric fixed effects estimated by 2-digit SIC for employment and payroll growth between years 1 and 2. Standard errors are clustered at the firm zip code by SIC division level. The regressions are estimated on the continuing firm matched pair sample and includes small businesses that survived through the event year $t+p$. Columns 1 through 5 estimate the causal impact of the amount of home equity extracted on log employment from years $t+p-1$ through $t+p+3$, respectively. Columns 6 and 7 estimate the impact on employment growth from years $t+p-1$ to $t+p+1$ and $t+p+1$ to $t+p+3$, respectively.

Table 1.16: Effect of Credit Access on Continuing Business Payroll

	(1)	(2)	(3)	(4)	(5)
			ln(Payroll)		
	$t+p-1$	$t+p$	$t+p+1$	$t+p+2$	$t+p+3$
ln($ Amount Extracted$_{i,t+p}$)	-0.111	0.207**	0.300	0.218	0.202
	(-1.42)	(2.02)	(1.63)	(1.01)	(0.81)
# Obs	17500	17500	17500	17500	17500
R-squared	0.787	0.685	0.522	0.571	0.559

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if continuing small businesses have differential payroll outcomes based on the amount of home equity extracted:

$$ln(\text{Payroll}_{i,j,t+p+k}) = \alpha + \beta ln(\$ \text{ Amount } \widehat{\text{Extracted}}_{i,t+p}) + \gamma \mathbb{X}_j + \omega \text{Controls}_i + \epsilon_{i,j,t+p+k}$$

Firms are denoted by i, firm creation year by t, the pair that the firm belongs to by j, p is the number of years since founding t that home equity is extracted for expansion (either two or three, which corresponds to the firms of age three or four). $\mathbb{X}_j$ are fixed effects for each pair. The outcome variables are log payroll k years around the event year $t+p$, where k ranges from -1 to 3. ln($ Amount Extracted$_{i,t+p}$) is the amount of home equity that the entrepreneur extracted in year $t+p$ (set to 0 if home equity is not extracted). Controls$_i$ is a vector of zip code level and firm level controls that includes: log of CLTV at purchase, log of home value at year $t+p-3$, 2-digit SIC industry fixed effects, the log of the number of months between home purchase and year t, legal form of organization (LFO) fixed effects, fixed effect for if the business was initially funded with home equity, non-parametric fixed effects estimated by 2-digit SIC for initial payroll, and non-parametric controls for home zip code characteristics of percent white, percent renter, percent below poverty line, and median income. Except for column 1, whose outcome variable is from year $t+p-1$, controls also include non-parametric fixed effects estimated by 2-digit SIC for employment and payroll growth between years 1 and 2. Standard errors are clustered at the firm zip code by SIC division level. The regressions are estimated on the continuing firm matched pair sample and includes small businesses that survived through the event year $t+p$. Columns 1 through 5 estimate the causal impact of the amount of home equity extracted on log payroll from years $t+p-1$ through $t+p+3$, respectively.

Table 1.17: Extensive Margin Effects of Refinancing Activity on Business Formation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			Excluding Non-tradeable, Construction, Finance, Insurance, Real Estate					Excluding Non-tradeable	All Industries
						$Y = \Delta ln$(Entry Rate$_{c,t+1}$)	Many Shocks Instrument		
	WLS	WLS	WLS 1st Stage	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Δln(# Extracted$_{c,t+1}$)	0.00383	0.0182***		0.247**		0.224**	0.236*	0.222**	0.200**
	(0.68)	(3.26)		(2.26)		(2.23)	(1.74)	(2.30)	(2.25)
Avg % IMB$_{c,t-1}$ $*\Delta$% IMB$_{-c,t}$			0.301***						
			(3.18)						
Δln($ Amt (Extracted$_{c,t+1}$)					0.197**				
					(2.00)				
# Obs	9719	9265	9173	9173	9173	9173	9168	9173	9174
R-squared	0.269	0.381	0.437	0.200	0.00558	0.227	0.216	0.296	0.368
F-stat	-	-	10.09	-	-	-	-	-	-
$\mathbb{Z}_{c,t}$ FE?	N	Y	Y	Y	Y	Y	Y	Y	Y

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing the effect of cash-out refinancing activity on business entry in aggregate:

$$\Delta ln(\#\text{ New Establishments}_{c,t+1}) = \alpha + \beta \Delta ln(\#\ \widehat{\text{Extracted}}_{c,t+1}) + \omega_c + \theta_{t+1} + \zeta \mathbb{Z}_{c,t} + \epsilon_{c,t+1}$$

Counties are indexed by c (1,493 counties) and years by t (2009 to 2015). # New Establishments$_{c,t+1}$ are from the Census SUSB dataset, which provides annual establishment entry counts by county and industry. # Extracted$_{c,t+1}$ is calculated from ATTOM and is a county level count of the number of households who extracted personal home equity in year $t+1$ in county c. $\mathbb{Z}_{c,t}$ is a vector of non-parametric controls for time-varying zip code level lagged growth rates ($t-1 \rightarrow t$) of: home prices, number of home purchases, unemployment rate, small business loan volume, and the total number of establishments. Regressions are weighted by the county's 2010 Census population. Columns 1-2 use WLS, column 3 shows the first stage, and columns 4-9 use weighted 2SLS. Columns 1-7 estimate the effect excluding establishments in the non-tradeable, construction, finance, insurance, and real estate industries. Column 4 shows the baseline effect for industries not reliant on local demand, column 5 estimates the effect for growth in the dollar amount extracted (instead of # of households extracting home equity), column 6 replaces the dependent variable with ln growth in establishment entry rates, and column 7 uses the instrument with many exogenous shocks. Column 8 estimates the effect excluding establishments in non-tradeable industries. Column 9 includes all industries.

Table 1.18: Robustness Tests for Extensive Margin Instrument

	(1) Business Loan Volume	(2) Unemployment Rate	(3) Share of Residents Who are White	(4) Share of Residents Without a High School Education
Avg % $\text{IMB}_{c,t-1} * \Delta\%\ \text{IMB}_{-c,t}$	0.0444	0.0364	-0.00299	0.0111
	(1.45)	(1.31)	(-0.64)	(0.34)
# Obs	9266	9266	4480	4474
R-squared	0.549	0.924	0.565	0.0739

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regression estimates testing if the instrument is correlated with county level variables across time:

$$\Delta Y_{c,t} = \alpha + \tau \text{Avg } \%\ \text{IMB}_{c,t-1} * \Delta\%\ \text{IMB}_{-c,t} + \omega_c + \theta_t + \zeta \mathbb{Z}_{c,t} + \epsilon_{c,t}$$

Counties are indexed by c (1,493 counties) and years by t (2009 to 2015). $\mathbb{Z}_{c,t}$ is a vector of non-parametric controls for time-varying zip code level lagged growth rates ($t-1 \rightarrow t$) of: home prices, number of home purchases, unemployment rate (not included in column 2), small business loan volume (not included in column 1), and the total number of establishments. Regressions are weighted by the county's 2010 Census population. In column 1, the dependent variable is growth in small business loan volume (business loans for less than \$100k) between t and $t+1$. This dependent variable is between t and $t+1$ since the potential concern is impacts on business formation between t and $t+1$ due to business loan funding (which is contemporaneous). Loan volume data are from the Community Reinvestment Act (CRA) dataset. In column 2, the dependent variable is growth in the unemployment rate between $t-1$ and t from BLS LAU. In columns 3 and 4, the dependent variables are growth in the share of residents who are white (and non-Hispanic) and who have less than high school education, respectively, from Census ACS between $t-1$ and t.

Chapter 2

Personal Wealth Shocks and Investment Manager Overconfidence

2.1 Introduction

One of the predominate questions in finance is if professional investment managers add value. Given that active equity mutual funds invest in over $5 trillion of assets, this issue has large effects on both returns to clients and the overall efficiency of the market. Most of the prior literature has focused on ex-ante static measures that are used to determine which fund managers have skill, and has largely assumed rationality. Studies of deviations from rational behavior, such as overconfidence, have largely been restricted to retail investors (Barber and Odean, 2001; Kumar, 2009; Grinblatt and Keloharju, 2009). This paper investigates positive shocks to the personal housing wealth of fund managers as a channel of time-varying overconfidence for mutual fund managers.

Home prices vary dramatically within narrow geographical areas, which provides large idiosyncratic variation in returns. If fund managers infer their professional skill from returns to their personal housing wealth, positive shocks to housing wealth may lead to overconfidence beliefs. Local home price growth should have no effect on fund performance based on rational models. Surprisingly, a strong and robust relationship is found between positive lagged changes in zip code level housing prices and forecasts of risk adjusted performance. This alone does not signal overconfidence. To disentangle the channel driving the anomaly, the trading behavior of fund managers is analyzed. A stunning pattern emerges from studying how fund manager trading changes as home price shocks evolve.

Following positive home price shocks, fund managers become more likely to lever up on existing stock holdings that have performed poorly, shift their buying towards levering up on currently held stocks and away from selecting new stocks, make worse choices in choosing existing positions to liquidate, and become less likely to fully liquidate existing positions. Interestingly, fund managers do not make worse decisions in selecting stocks to buy that are not currently held in their portfolio after a positive home price shock. Instead, they become biased towards their past choices and become more likely to believe their poor past stock picks will become winners. Akepanidtaworn, Mascio, Imas, and Schmidt (2018) find that fund managers underperform because of poor selling choices. The overconfidence channel shown in this paper makes the selling choices of fund managers even worse.

In addition to having a short-run impact on trading, overconfidence also has a long-run effect. To show this, the fund's own ex-ante portfolio holdings are used as a benchmark, similar to the approach in Barber and Odean (2001) for retail investors. As home price growth positively accelerates, the fund's active return decreases and becomes worse than their passive portfolio. Additionally, trading costs increase, which is partially to blame for the decline in performance. To test this further, the fund's annual turnover is analyzed and found to increase as home price growth increases—a common sign of overconfidence.

To further show that the result is driven by overconfidence, this paper borrows from the retail investor and corporate finance literatures. The effect is strongest among groups of fund managers who are more likely to be susceptible to overconfidence beliefs. This includes fund managers who are male (Barber and Odean, 2001), less experienced (Greenwood and Nagel, 2009; Chernenko, Hanson, and Sunderam, 2016), have greater housing leverage (Cronqvist, Makhija, and Yonker, 2012; Liu and Yermack, 2012), who bought their home more recently, and have less education.

Past research has shown that fund flows affect returns (Berk and Green, 2004; Song, 2019) and that investors invest locally (Coval and Moskowitz, 1999; Ivković and Weisbenner, 2005; Pool, Stoffman, and Yonker, 2012). To rule out that the decline in performance is driven by local increases in fund flows caused by local housing wealth, three tests are performed. First, it is shown that fund flows are not affected by home price shocks. Second, flows are included as an additional control to the performance regression, which is found to not affect the results. Lastly, past performance is added as an additional control to the performance regression and is also found to not affect the relationship between positive home price shocks and future fund performance. This test is to rule out that future returns are decreasing due to strong prior returns. If lagged performance was strong, fund flows might increase, which subsequently would decrease returns.

To rule out a spurious relationship, three placebo tests are performed. First, it is shown that index funds do not respond to these home price shocks (Pástor, Stambaugh, and Taylor, 2017). In addition to index funds, funds with a low tracking error or Active Share should also not respond to home price shocks. The logic being that if a fund manager is a closet indexer, they will maintain the same passive strategy before and after the shock. It is found that both definitions of closet indexers do not respond via this channel, and the effect is stronger for truly active fund managers. These tests rule out the possibility that the result is driven by an unobserved local factor that influences fund returns.

One concern with the results in this paper is that fund managers might not be paying attention to changes in their home value. To provide evidence against this concern, fund managers who extract personal home equity via cash out refinancing are shown to have a decline in performance relative to their performance prior to extracting home equity. Cash-out refinancing has a large effect on the household balance sheet (Greenspan and Kennedy, 2008; Bhutta and Keys, 2016; DeFusco, 2018) and these results show that there is a link to professional performance as well. With this result, home price growth is used as an instrumental variable for the decision to extract home equity to show that fund managers who extract home equity decide to refinance when home prices are higher. The use of the instrumental variable is to provide evidence that fund managers pay attention to their personal home value and is not intended for use as an instrument in the traditional sense.

This paper provides results showing how professional behavior can be affected by home price shocks, focusing on the effect from positive housing wealth shocks. This is related to recent work by Stoffman et al. (2018) and Bernstein et al. (2018), who study the impact of negative housing wealth shocks for mutual fund managers and innovative workers, respectively. Stoffman et al. (2018) find that in response to negative housing wealth shocks, fund managers decrease their portfolio risk due to career concerns. They find no effect on risk-adjusted returns and a marginally significant effect on raw returns. This paper finds a similar effect on returns and risk taking during periods of negative home price shocks. When extended to a broader time-series, the effect for changes in risk during periods of positive home price growth survives, while the result during periods of negative home price growth does not.

This paper builds on the literature for sorting institutional investment managers by ex-ante measures of skill. Most of the prior work in this body of research has focused on static measures (Khorana, Servaes, and Wedge, 2007; Chen, Goldstein, and Jiang, 2008; Evans, 2008; Cremers and Petajisto, 2009; Cremers, Driessen, Maenhout, and Weinbaum, 2009; Greenwood and Nagel, 2009; Hong and Kostovetsky, 2012; Pool, Stoffman, and Yonker, 2015; Chernenko et al., 2016; Cremers and Pareek, 2016). Recent work, such as Gupta and Sachdeva (2017) and Pástor et al. (2017), has found dynamic measures that forecast performance. This paper focuses on a novel dynamic measure that has not been explored previously. To study housing wealth shocks, data on the personal real estate holdings and transactions of mutual fund managers are merged to data on their professional portfolio holdings and returns of their funds for a sample of over 1,200 actively managed domestic equity mutual funds with data between 2001 and 2018.

The remainder of this paper is organized as follows. Section 2 provides an overview on home price variation and the approach used in this paper. Section 3 describes the datasets used in this paper, as well as how they are merged. Section 4 describes the empirical methodology and results for the effect on performance and related tests. Section 5 studies how trading behavior is affected by housing wealth shocks. Section 6 presents evidence that fund managers pay attention to the values of their homes. Lastly, Section 7 concludes.

2.2 Home Price Variation

For the vast majority of households, the home is the single largest investment (Campbell, 2006). Fund managers are a wealthier subset of the population, but they own larger homes than the average household and their home values dwarf their annual salary (Figure 2.1).

The average home value of fund managers is generally over $800k, while their average annual salary is between $200-300k. A 20% idiosyncratic home price shock can equal more than one year of salary. If home price growth is abnormally large then fund managers might use this as feedback reinforcing a belief in their superior ability to pick assets. At the same time, this represents a large wealth shock. While the wealth effect cannot be ruled out, tests showing which types of fund managers respond to these shocks and the change in their trading behavior from these shocks lend credence to the overconfidence channel.

To test the hypothesis that home price shocks affect fund manager performance, lagged 3-year zip code level home price changes of the primary home of fund managers are regressed on forecasts of performance and investment selections.[1] In order to observe the real estate data for fund managers, a merge is utilized between fund managers in Morningstar and real estate transaction and assessor data from ATTOM (discussed in the Data section). Fund managers are included in the sample when at least three years has passed since the purchase of their home. This restriction ensures that fund managers realized the home price gains.[2] The primary regression model in this paper utilizes fund fixed effects (for each merged non-disjoint period of time) and monthly time fixed effects. It is possible that the merge between Morningstar and ATTOM produces disjoint samples across time for a fund. For these funds, each group of merged overlapping fund managers that produces a non-disjoint sample across time is separately studied. If more than one fund manager is merged for a given month, the home price growth of the fund managers is averaged. These fixed effects remove time invariant preferences of fund managers, including housing location choice (thereby controlling for sorting into home zip codes within a geographical area). With these preferences removed and aggregate economic conditions absorbed by the time fixed effects, the remaining variation in home price growth is unlikely to be endogenous.[3]

Home price growth is measured at the zip code level. This micro-level measurement of home price growth provides dramatic variation within small geographical areas. Figure 2.2 shows home price growth at the zip code level for the Boston area, a popular region for mutual fund managers. The volatility in home price growth within small geographical areas makes it unlikely that micro level home price growth is correlated with investment choices due to an omitted variable of local economic activity. To further show the micro level variation in home price growth, Figure 2.3 shows changes in home price growth for two neighboring zip codes in the Philadelphia suburbs. These two zip codes are popular for fund managers.

[1] For robustness, 2-year and 4-year home price changes and 3-year home price growth on their secondary homes are also tested.

[2] Fund managers who bought fewer than three years ago are included in the sample if they owned another property that was merged and spent, in aggregate, at least three years in the two properties. In this case, the home price growth of the two properties is averaged based on the amount of time each property was lived in.

[3] In the appendix, a second model utilizing commuting zone by time fixed effects and a battery of controls is estimated, and a similar effect is found. The additional controls attempt to correct for fund manager preferences, including location preference within a commuting zone. This leaves idiosyncratic home price growth within a commuting zone. A similar method is utilized by Stoffman et al. (2018) and Bernstein et al. (2018).

The first observation is that idiosyncratic home price variation is not persistent. Initially zip code 19087 outperforms 19405, but after a couple years the relationship flips and zip code 19405 outperforms 19087. Over the 2001 to 2018 period, the winner between these two zip codes switches multiple times. It is not that one zip code always outperforms. The dynamic nature of which zip codes outperform makes it less likely that fund managers bought in a zip code in anticipation that it will outperform a neighboring zip code. Second, variation between neighboring zip codes is substantial. In 2003, zip code 19405 had lagged 3-year home price growth of 20%, while zip code 19087 had growth of 40%. In other words, a home a short distance away experienced double the amount of home price growth. Given that many fund managers have home values in excess of $1 million, this creates very large idiosyncratic shocks.

2.3 Data

To study the effect of personal wealth shocks on professional behavior, data on both the household balance sheet and changes in performance on the job are needed. For confidentiality reasons, this combination of data is challenging to obtain. Mutual fund managers and their personal real estate are the perfect set-up to study this. Securities laws make the professional behavior of mutual fund managers public information. A fund's performance, management team, and holdings are public knowledge. This combination of data presents extremely rich data on professional behavior. In the United States, real estate transactions and the associated buyers are also public record and are collected by county recorder offices. When merged together this combination of data presents a unique opportunity to dynamically study changes in professional behavior from personal wealth shocks.

Fund characteristics and performance data are from CRSP and information on who manages a mutual fund is obtained from Morningstar. Using a method similar to Loutskina and Strahan (2015), CRSP and Morningstar are merged on CUSIP and ticker. Roughly 70% of the funds in CRSP are merged to Morningstar. Quarterly holdings data are from Thomson Reuters, which is merged to CRSP on WFICN using the mflink files.[4] This merge is only used for analysis on mutual fund holdings; for the majority of the results, the data are not restricted to funds merged to Thomson Reuters.[5] Funds are only included in the sample if

[4]The holdings data are supplemented with CRSP holdings data when the data are missing in Thomson Reuters.

[5]Additional data on stock characteristics are obtained from CRSP. Factor data is obtained from Fama-French-Carhart in WRDS and benchmark return data is obtained from FRED.

they are actively managed domestic equity funds that predominantly invest in US equities.[6] Index funds are identified by the index fund flag in CRSP being populated or having the word "index" in the name of the fund and are removed from the actively managed population. The population is further restricted to fund months with at least $15 million in total net assets (TNA, measured in inflation adjusted 2000 dollars) and with 10 or fewer managers.

Real estate transaction data is obtained from ATTOM, which is an administrative dataset of real estate transactions in the United States. ATTOM collects and digitizes the data from county recorder offices. The dataset includes the location of the home, loan amounts, and the value of the home at time of purchase.[7] To obtain changes in home value over time, monthly zip code level home price growth data from Zillow are used. The data start in 1997, with data for some zip codes not available until later. As a result, fund performance data from 2001 to June 2018 are utilized. The relatively unexplored data for mutual fund manager personal real estate is obtained by a merge between ATTOM and Morningstar. ATTOM includes the (non-standardized) names of homebuyers, which are merged to fund manager names in Morningstar. Names are standardized to try to account for spelling differences (i.e., Christopher is mapped to Chris).

Two passes are attempted to merge these two datasets. The first pass merges on the full name (including middle initial, if available). While, the second pass merges on the initials of the first and middle names and the full last name. Other attempts to merge on partial names or not using middle initial yielded limited additional successful merges. A merge is only kept if the fund manager name merged to a unique homebuyer name within the commuting zone of the fund location. Fund location is obtained from the zip codes of the fund offices listed in either CRSP or Morningstar. The algorithm is similar to the one used in Bernstein et al.

[6]These types of funds are identified by restricting the Lipper Prospectus objective code, Strategic Insight objective code, and the Weisenberger objective code to the values listed in Cremers and Pareek (2016). The restriction on the CDA/Spectrum code is not used as this value is in Thomson Reuters, which would require an additional merge that would limit the population size. Additional restrictions are included to be sure that only actively managed domestic equity funds are included. The CRSP objective code is restricted to domestic equity funds of cap-based or style funds (the first three characters set to EDC or EDY). This restricts to domestic equity large/mid/small/micro cap, growth, income, hedged, short, and income funds. Additionally, the Morningstar category has to be in the 3-by-3 size/value grid, as in Stoffman et al. (2018), or have a category type of 85%+ equity allocation. The Morningstar category is also used to assign funds a benchmark. The benchmarks are assigned as follows: US Fund Large Growth (Value) is Russell 1000 Growth (Value) Total Return, US Fund Large Blend is Russell 1000 Total Return, US Fund Mid-Cap Growth (Value) is Russell Mid-cap Growth (Value) Total Return, US Fund Mid-Cap Blend is Russell Madcap Total Return, US Fund Small Growth (Value) is Russell 2000 Growth (Value) Total Return, US Fund Small Blend is Russell 2000 Total Return, and US Fund Allocation–85%+ Equity is Russell 3000 Total Return.

[7]An additional feature of ATTOM is information on the tax address of properties (based on annual files from county assessor offices). With this information, secondary homes are linked to the primary homes of the fund managers. Roughly 16% of fund managers in the sample own more than one home, and at most four homes are owned by a single fund manager at any given point in time. The top 25 zip codes for these secondary properties are overwhelmingly in vacation destinations such as Naples Florida, Lake Tahoe California, ski towns in Colorado and the Berkshires, and various lake and beach towns (such as the Hamptons and Lake Winnipesaukee). The locations of these homes provides validation that these are vacation properties owned by the fund manager.

(2018). Merges are dropped if they result in a non-disjoint period of less than 24 months of fund data for which at least one fund manager is merged to ATTOM. This is in order to accurately estimate risk-adjusted returns (robustness is provided around this restriction).

40% of the managers in Morningstar are successfully merged to ATTOM, which is lower than the 52% rate reported in Bernstein et al. (2018) for innovative workers who produce patents. Their patent data includes the person's home zip code, while Morningstar only includes the zip code of the fund and does not include information on the location of the fund manager's home. It is assumed that a fund manager lives within the same commuting zone as their fund. Managers who were not merged either rent, have a common name, live outside the commuting zone of their office, used a trust to purchase their home (which usually partially or fully shields their identity), or had a misspelling in either dataset. The final population comprises 1,241 funds based on information for 1,368 fund managers, for a total population of 87,719 monthly fund pairs.

Figure 2.4 shows the percentage and number of managers merged for each fund in the final population. Many funds are managed by a team of fund managers, but for the majority of funds only one manager is merged. Generally, at least 50% of the managers are merged for each fund. In cases when more than one manager is merged, the fund manager characteristics and housing growth information is averaged. Not having all of the managers merged leads to noise on the right hand side, which should bias the results towards zero—robustness is provided around the merge rate.

Table 2.1 provides summary statistics comparing the population of active domestic equity mutual funds merged to ATTOM against the population not merged. On average, both populations have a negative alpha after fees of about 10bps per month. The merged population is biased towards larger funds (both in terms of TNA and number of managers) with lower turnover. On the fund manager characteristics qualities, almost all managers are male (92%), most have an MBA or CFA (57 and 56%, respectively), and most had substantial career experience at the time they became fund manager (18 years on average). Experience data are only available for a subset of managers and is supplemented with year of birth+22 or the year of undergraduate graduation for the career start date (when these variables are available).

2.4 Empirical Results on Performance

2.4.1 Main Specification

The primary hypothesis of this paper is that exogenous positive returns to the personal real estate of fund managers affects fund managers via overconfidence. To test this, a regression model of lagged zip code level home price growth on forecasts of performance is estimated. The baseline model analyzes 1-month ahead forecasts of risk adjusted returns on lagged 3-year home price changes for the values of the primary homes of fund managers. Additional tests are performed to link this effect to overconfidence.

The regression framework is similar to the one utilized by Gupta and Sachdeva (2017). The four factor Fama-French-Carhart model is estimated at the fund level:

$$R_{it} - R^f_t = \beta_{MKT,i} MKT_t + \beta_{SMB,i} SMB_t + \beta_{HML,i} HML_t + \beta_{UMD,i} UMD_t + \epsilon_{it} \tag{2.1}$$

Returns are net of fees (robustness is provided around this), which represents the returns experienced by investors. The factor loadings are used to obtain monthly alphas:

$$\widehat{\alpha}^{FFC}_{it} = R_{it} - R^f_t - \widehat{\beta}_i * \lambda^{FFC}_t \text{ where } \lambda^{FFC}_t = (MKT_t \; SMB_t \; HML_t \; UMD_t) \tag{2.2}$$

The $\widehat{\alpha}^{FFC}_{it}$ are used in panel regressions to estimate the predictive ability of lagged home price growth to forecast 1-month ahead risk adjusted performance:

$$\widehat{\alpha}^{FFC}_{i,t,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita} \tag{2.3}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects, and ρ_a are non-parametric fixed effects for TNA. Standard errors are clustered at the level of fund by each merged non-disjoint period of time (robustness is provided around this), which allows for arbitrary correlation of the error terms within fund over time. Unbounded variables are winsorized at the 1% and 99% level.

To test the effect over periods of positive and negative home price growth, two approaches are used. The first simply divides the sample into two periods, one where $\Delta 3\text{yr HPI}_{i,t-1} \geq 0$ and the other where $\Delta 3\text{yr HPI}_{i,t-1} < 0$. The second approach uses the entire population and estimates the coefficients for positive and negative home price growth separately in a single regression. $ln(\Delta 3\text{yr HPI}_{i,t-1})$ is interacted with indicator variables for positive and negative home price growth, denoted as $ln(\Delta 3\text{yr HPI}^{+}_{i,t-1})$ and $ln(\Delta 3\text{yr HPI}^{-}_{i,t-1})$, respectively. For this regression, the time fixed effects are also interacted with indicator variables for positive and negative home price growth to control for variation in returns among funds managed by managers experiencing positive and negative home price growth in a given month, denoted by σ^{+}_t and σ^{-}_t, respectively.

The results for the main regression model are presented in Table 2.2. Overall, home price growth significantly negatively impacts future performance when using data from the entire 2001 to 2018 period (column 1). In columns 2 and 3, the sample is divided into periods of positive and negative home price growth, respectively. From this segmentation, it is shown that the effect is isolated to periods of positive home price growth, with the magnitude increasing when periods of negative home price growth are removed. Interestingly, during periods of negative home price growth there is no effect. This is in line with the findings of Stoffman et al. (2018), who also find no impact on risk adjusted returns from home price shocks during the Great Recession. This also shows how the two shocks operate via different channels, overconfidence during periods of positive home price growth and, as shown by Stoffman et al. (2018), career concerns during periods of negative home price growth. When the regression is estimated using the entire sample and the home price growth variable is bifurcated into periods of positive and negative home price growth, similar results are found (column 4).

The average standard deviation in 3-year home price growth within fund is 12%, which yields a decline in alpha of 37bps per year for a one standard deviation positive home price shock. Given the dramatic rise in home price growth over the mid-2000s and late 2010s, there is the potential for a dramatic decrease in performance from exogenous positive shocks to home price growth. The results are significant at the 1% level and the t-statistics are generally above 3 to pass the threshold of Harvey, Liu, and Zhu (2016). In column 5, 1-year lagged alpha (calculated from a regression of Fama-French-Carhart on the preceding 12 months) and 1-month forward fund flows are added as controls. With these variables included, the sample size is reduced since there have to be 12-months of lagged data available. The result is strengthened with the inclusion of these controls.

To further show that fund flows are not affected by home price shocks, the regression is estimated with the left hand side replaced with 1-month and 1-year forecasts of fund flows, calculated using gross returns (before fees). Table 2.3 shows that the home price shocks do not have an impact on fund flows. Combined with the earlier result showing that the relationship between home price shocks and returns is not affected by controlling for fund flows, fund flows are not a concern for the findings in this paper. Additional robustness tests are provided in the Appendix.

A common placebo test is to re-estimate the effect on performance for a sample of index funds. Table 2.4 shows that when the sample of active funds is replaced with a sample of index funds, the result becomes insignificant. In addition to using index funds as a placebo, closet index funds represent another potential placebo test. For robustness, two approaches are used to define a fund manager as being a closet indexer. The first approach labels a fund manager as a closet indexer if their fund's tracking error is in the lowest quantile. Column 4 shows that for this population of closet index funds no effect is found. This approach can also be used to restrict the sample to the most active funds. When restricted to very active funds, a larger response is found (column 5). For the second approach, a fund is defined as a closet index fund if the fund's Active Share is less than 60%, following Chernenko et al. (2016). Active Share is provided by Cremers and Petajisto (2009). Similarly, when using the Active Share definition of a closet indexer no effect is found (column 6). Using Active Share to restrict to the most active funds (those with an Active Share > 60%), a stronger effect is found once again.

2.4.2 Which fund managers/funds respond to home price growth?

This section provides subsample analysis to see which types of fund managers and funds respond to home price shocks. If overconfidence is driving the result then fund managers with less education, less experience, who bought their homes more recently, with greater housing leverage, and who are male should respond more to home price shocks. To test this, the regression model is estimated with the home price growth variable interacted with indicator variables for different measures of characteristics of funds/fund managers during periods of positive home price growth. The omitted category is the group of funds/fund managers who are least likely to be affected by home price shocks (i.e. those with an MBA when splitting the sample by education).

Figures 2.5, 2.6, and 2.7 present these results graphically. Figure 2.5 shows that not having an MBA, having less experience, and buying a home more recently all produce a stronger effect. These are exactly the types of fund managers one would expect to buy into overconfidence from home price shocks. While more experienced fund managers generally do not respond to home price shocks, experienced fund managers who bought their home more recently respond strongly to home price shocks. Buying more recently should induce a stronger response as homeowners would be more likely to pay attention to their home value shortly after they buy their home. This is due to two reasons. First, their home value is more likely to be on their mind. Second, short-run movements will have a larger effect on the overall return on their purchase. Buying in the distant past makes short-run movements in home prices less important due to already having accumulated home price gains.

Figure 2.6 also confirms that fund managers without a CFA designation (not significantly), with more expensive homes, and with greater housing leverage (as measured by loan to value [LTV] at purchase) also have a stronger response to home price shocks. This is related to the CEO literature, namely Cronqvist et al. (2012) and Liu and Yermack (2012), who show that CEOs with larger homes and higher LTV ratios engage in riskier behavior. Consistent with the overconfidence results of Barber and Odean (2001) and Lu and Teo (2018), funds with only male managers (among the merged managers) have a much stronger response.

In addition to segmenting the population on fund manager characteristics, the population is segmented on fund characteristics to identify the types of funds that are more likely to have managers affected by home price shocks (Figure 2.7). The size of the fund (as measured by TNA) does not produce a differential response, indicating that the result is not driven by small funds. Fund type and capitalization focus do, however, provide strong segmentation. Value and small cap fund managers are less affected by home price shocks, while growth/blend and large and mid-cap funds are strongly affected. Lastly, funds with low expense ratios have a stronger impact. This could be explained by the fact that funds that are able to command a high expense ratio are likely managed by senior managers with significant experience, and as such are less likely to be swayed by home price shocks.

As a last test, the sample is divided into funds managed by fund managers who have never owned a second home and fund managers who have. In Table 2.5, home price growth for secondary homes is included as an additional regressor. If the population is restricted to fund managers who ever owned a second home (columns 1 and 3) the result becomes insignificant, while focusing on fund managers who never owned a second home preserves the results (columns 2 and 4). This is additional evidence of overconfidence, as fund managers who own a second home are likely to have greater experience. While other channels cannot be ruled out, such as increased risk taking from a wealth effect, the results are consistent with overconfidence being the channel. In the next section, trading behavior is explored and the results provide additional evidence of the overconfidence channel.

2.5 Trading Behavior

In order to understand why performance declines following positive home price shocks, the trading behavior of fund managers is analyzed. First, short-run changes in trading behavior of fund managers is analyzed. Second, long run trading behavior is studied using the fund's own benchmark returns as a passive portfolio. Lastly, portfolio risk taking is analyzed.

2.5.1 Trading Return

After establishing that performance declines and trading increases from positive home price shocks, it is crucial to understand the mechanisms that are driving this. In this section, the trades performed each quarter are studied using quarterly equity holdings data.[8] With this data, stocks are identified as purchased, held constant, or sold between quarters. A trade is labeled a buy trade if the fund's position in the stock increases by at least 5%. I start by estimating the effect of home price shocks on the 1-quarter ahead return to these trades. Returns are weighted by the value of the change in the number of shares using the average of the stock's price at the beginning and end of the quarter in which the stock is traded. To estimate the effect the following regression is estimated:

$$\mathrm{Y}_{i,t+1,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{i,t,a} \tag{2.4}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are quarterly time fixed effects (interacted with indicators for positive and negative home price growth), and ρ_a are non-parametric fixed effects for TNA.

The forecasted return to the net trading decisions (the return to the buy trades less the return to the sell trades) of fund managers declines as home price growth increases (Table 2.6, column 1). This implies that fund managers are making worse trading choices from positive personal wealth shocks. A 12% positive shock to home price growth decreases the return difference by roughly 7bps, which is almost 100% of its mean. To understand why the return declines, the return is decomposed into buy and sell choices. Broadly speaking, a fund manager has five different trading choices: buy a new stock not currently held, increase the position of a stock currently held, not trade a stock currently held, fully liquidate a position, or partially liquidate a position. Using these five actions, the 1-quarter ahead returns less the benchmark return are calculated for each.

[8]The data are from Thomson Reuters s12, supplemented with data from CRSP. The data are reported as of the end of each quarter. It is possible that stocks are bought and sold within each quarter, in which case they would not be captured.

Table 2.6 column 2 shows that fund managers are not making differential choices when selecting new stocks to purchase. Instead, fund managers make worse choices in picking stocks to lever up on (column 3). A 12% positive shock decreases the return to levering up by 35% of its mean. The mean return to the buy trades of buying new stocks and levering up on existing stocks are both approximately 16bps, indicating that on average fund managers do not make bad choices in selecting stocks to buy.

For analyzing the return to sales, home price growth is not divided into periods of positive and negative home price growth due to the strong overall effect. The literature has found that a key reason for fund underperformance stems from making poor selling choices (Akepanidtaworn et al., 2018). They attribute this to inattention in the stock selling decision and following a heuristic of selling stocks with recent extreme price movements. Fund managers make worse choices in fully liquidating existing positions as home price shocks increase (column 6—the positive coefficient implies that the stock sold subsequently performs better). This result provides another channel that drives the deterioration in the choice of stocks to sell. Lastly, column 7 shows that flows are not affected, indicating that the results are not driven by fund managers trading due to variations in flow as home price shocks evolve.

The decline in returns to increasing positions in existing stocks is explored further by augmenting the above regressions to fit linear probability models of the likelihood of a stock being levered up on having ex-ante performed worse in the 2-quarter period during and prior to the trade. Table 2.7 shows that the stocks levered up become more likely to have ex-ante performed worse as home price shocks increase. The result is strongest in column 2 where the probability of buying more of stocks that have performed at least 20% worse than the benchmark in the prior 6-month period increases. As fund managers experience positive home price growth, they become more likely to increase their positions in stocks that have performed poorly, relative to both the benchmark and the stocks they hold but choose not to increase their stakes in.

Column 4 shows that this is not true for new stocks purchased that are not currently held. Fund managers do not become overconfident in poorly performing stocks that they did not previously buy. However, conditional on having already owned the stock, fund managers become more likely to double down on their position if the stock had a period of poor performance. Columns 5 and 6 estimate the effect on the share of buy (sell) trades that are increases in existing positions (full liquidations). As home price growth increases, fund managers shift their buy trades towards levering up and shift their sell trades towards partially liquidating (instead of fully selling). Not only do fund managers lever up on losers but they devote less attention to picking new stocks. In addition, they become more likely to partially hold onto their positions. In the next section, a long-run impact on the active trading returns of fund managers from their personal home price shocks is established.

2.5.2 Comparison to a Passive Strategy

In this section, evidence is presented that positive home price shocks also have a long-run impact on performance. To show this, actual returns are compared against own-benchmark

(”passive”) returns in a similar spirit of Barber and Odean (2001), who study retail investors. Each June (for every year t) the change in $ln(\Delta 3\text{yr HPI}_{i,t})$ over the prior 1-year period (the rate of change) is calculated. Figure 2.8 shows the distribution of these values. Over a one year period there can be a dramatic change in the rate of home price growth. This rate of change is interacted with indicator variables for $ln(\Delta 3\text{yr HPI}_{i,t})$ being positive or negative. The rate of change in home price growth is used, as changes in trading behavior due to changes in home price growth is studied.

To construct the performance of a ”passive” portfolio, the 2-year (24 months of data) alpha is constructed from a Fama-French-Carhart regression of the returns to the stocks held in quarter 3 of year $t-1$ between quarter 3 of year $t-1$ and quarter 2 of year $t+1$. For the active performance, an alpha over the same time period is calculated using returns to the actual quarterly holdings, updated each quarter. The sample is restricted to funds whose managers survived for the respective 24 month period, have at least 80% of their assets represented by these holdings, and have a correlation between actual returns and the constructed active returns of at least 99% (80% of the records). These last two restrictions are to ensure that the returns from the raw holdings data represent the actual returns to the funds. Without this restriction, it might be that the fund manager trades in other asset classes, for which holdings data are not available (i.e. cash, derivatives, etc.), and this would affect their positions in equities and their overall returns.

To estimate trading costs, the approach of Kiyotaki and Moore (1997) is used.[9] Since the data are annual, commuting zone by time fixed effects are used in lieu of fund fixed effects. Included are the following additional control variables: Morningstar fund category fixed effects, number of manager fixed effects, an indicator for if the manager owns a second home, an indicator for if the manager has an advanced degree (above undergraduate), and non-parametric fixed effects for TNA, home value, combined LTV (CLTV), average zip code level income (from the 2010 IRS SOI), and the percent of non-white households at the zip code level (from the 2000 Census). These variables control for fund and home location time invariant qualities. The following regression model is estimated:

$$\alpha^{\text{Active}}_{i,t-12mo\rightarrow t+12mo,cz} - \alpha^{\text{Passive}}_{i,t-12mo\rightarrow t+12mo,cz} = c + \beta(ln(\Delta\text{3-yr HPI}_{i,t}) - ln(\Delta\text{3-yr HPI}_{i,t-12mo})) + \eta_{cz,t} + Controls_{it} + \epsilon_{i,t,cz} \quad (2.5)$$

where $\eta_{cz,t}$ are annual (t is annual) time by commuting zone fixed effects (interacted with indicator variables for positive and negative home price growth).

β captures the effect of the rate of change in home price growth on the performance of the fund relative to the performance if the fund manager did not trade. A positive coefficient

[9] Kiyotaki and Moore (1997) estimate institutional stock level trading costs as a function of: trade size, market capitalization, if the fund is a technical fund, if the stock is Nasdaq listed, and if the fund is an index fund. This last indicator variable is set to zero for the population of active managers in this paper. Following Busse, Chordia, Jiang, and Tang (2016), the technical fund indicator is set to 0.45 for buys and 0.6 for sales. For obvious reasons trading costs are not subtracted from the returns to the passive portfolio.

would imply that increases in home price growth lead to the fund manager's trading adding value relative to doing nothing. Standard errors are clustered by fund (for each merged non-disjoint period of time).

Table 2.8 column 1 estimates the effect on active returns less estimated trading costs and their own-benchmark passive return. The result is negative and significant at the 5% level for periods where home price growth is positive. A one standard deviation shock to the rate of change in home price growth is 0.1, which equates to a decline in active risk-adjusted performance of 45bps per year relative to their passive benchmark. There is no effect when home price growth is negative. Column 2 replaces the left hand side with an indicator variable for if the active alpha, net of trading costs, beats the passive alpha. A one standard deviation shock to the rate of change in home prices decreases the probability of beating the passive benchmark by 8%. This implies that not only does the active performance decline but it also makes the manager less likely to add value over their own-benchmark passive holdings. On average, only 35% of fund managers beat their passive benchmark when accounting for trading costs. Again, there is no effect from negative home price shocks.

When trading costs are not included, the result for periods of positive home price growth loses magnitude and significance but remains significant at the 10% level. Part of the reason for the decline in performance is due to an increase in trading costs. Column 4 shows that trading costs increase as home price growth increases. An increase in trading is a common sign of overconfidence in the literature. To confirm that trading is increasing, the effect on annual turnover data from CRSP is estimated. For this regression, the full sample of merged funds is used (not only the ones that further merged to Thomson Reuters). Since the data are annual, the regression uses the specification above with commuting zone by time fixed effects and additional control variables. Lagged log turnover is included as a control variable as well. Column 5 shows that turnover does indeed increase as home price growth increases.

For retail investors, it is well established that excessive trading leads to worse returns (Odean, 1999; Odean and Barber, 1999; Barber and Odean, 2001), while for institutional investors the results are mixed. Overall, Pástor et al. (2017) find that higher trading leads to higher returns, with the effect being stronger for high fee funds. Conversely, Cremers and Pareek (2016) find that patient investment strategies of institutional investors leads to higher returns. This paper argues that if turnover is driven by overconfidence due to home price shocks, performance will suffer. This is due to the increase in trading being driven by irrational reasons.

2.5.3 Risk Taking

In this section, the link between home price shocks and risk taking is investigated. This section relates to Stoffman et al. (2018) and extends their findings to a broader timeseries. In their work, the primary focus is on the risk taking measure of the standard deviation of returns during the Great Recession. Using a longer period (2001 to 2018), the standard deviation of returns is calculated using monthly returns for the subsequent 12-month period following each monthly observation of lagged 3-year home price growth. Additionally, results for tracking error and CAPM beta (market risk) are shown. The regression uses the baseline specification in this paper and includes fund (for each merged non-disjoint period of time) fixed effects, time fixed effects, and non-parametric TNA fixed effects.

Table 2.9 columns 1 and 2 show that positive home price growth has the opposite effect on the standard deviation of returns compared to negative home price growth. For periods of positive home price growth, the effect is negative. While for periods of negative home price growth, the effect is positive. Both this paper and Stoffman et al. (2018) find a positive and significant effect on risk taking during periods of negative home price growth. When running a horse race on the broader 2001 to 2018 period, the effect from positive home price growth survives while the effect from negative home price growth does not.

In Stoffman et al. (2018) a positive effect is found for tracking error and no effect is found on market risk. Similarly, this paper finds a positive effect for tracking error (column 5) and no effect for market risk (column 8) during periods of negative home price growth. Similar to the effect for the standard deviation of returns, in a horse race the effect on market risk only remains for periods of positive home price growth. For periods of positive home price growth, tracking error and market risk decline (columns 4 and 7). Overall, positive home price shocks lead to a decline in risk taking. Given the earlier results, this is most likely reconciled with a decline in taking positions on new stocks and relying instead on existing positions.

2.6 Attention to Home Price Growth

A key assumption of the results in this paper is that fund managers pay attention to home price growth. To test this, personal home equity extractions of fund managers are studied. Cash out refinancing was common during the housing boom of the mid-2000s, with many papers written about this in the household finance literature.[10] For a description of how the amount of home equity extracted is constructed using ATTOM data please see the Online Appendix to Chapter 1.

This paper uses home equity extractions to study if fund managers are paying attention to home prices, among the selected group of homeowners who extract home equity. If fund managers are more likely to extract equity when home prices are growing at a faster rate, this is evidence that managers are aware of the changes in their home value. Fund managers are a wealthier segment of the population but they also heavily engage in cash out refinancing. It is left as an open question as to why fund managers extract home equity (i.e., is it to supplement income, for consumption, to start a small business, to invest, etc.). Around 5-15% of fund managers extract home equity in any given year, with the proportion of fund managers extracting home equity decreasing over time (Figure 2.9a). Among fund managers who extract home equity, the median fund manager extracts roughly \$36,000 and 10% of fund managers extract over \$340,000 (Figure 2.9b). To test the impact of home equity extraction on fund performance, the following regression model is estimated:

[10] See Greenspan and Kennedy (2008), Bhutta and Keys (2016), and DeFusco (2018), among others.

$$\overline{\alpha_{i,t+1,a}} - \overline{\alpha_{i,t-1,a}} = c + \beta\mathbb{I}(\text{Extract } \widehat{\text{Home}} \text{ Equity}_{i,t}) + \lambda_i + \sigma_t + \rho_a + \overline{\alpha_{i,t-1,a}} + \epsilon_{i,t,a} \quad (2.6)$$

where $\overline{\alpha_{i,t,a}}$ is the average monthly alpha in quarter t and $\mathbb{I}(\text{Extract Home Equity}_{i,t})$ is an indicator equal to 1 if a fund manager extracted home equity in quarter t. λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are quarterly time fixed effects, and ρ_a are non-parametric fixed effects for TNA. t is measured as quarterly observations and standard errors are clustered at the fund level (for each merged non-disjoint period of time).

If β is negative, home equity extraction in quarter t forecasts a decline in risk-adjusted performance in quarter $t+1$ relative to performance in quarter $t-1$. These results are forecasts and provide ex-ante information that can be used to screen funds. The change in lagged 3-year zip code level home prices is used as an instrument. The purpose of this instrument is not to correct for endogeneity, but to show that variation in home price growth affects the probability of home equity extraction. Table 2.10, column 1 shows overall there is no significant effect for home equity extractions on future performance. When the choice of home equity extraction is driven by variation in home price growth (columns 3 and 5) there is a large negative effect on forecasted fund performance. Additionally, lagged home price growth predicts a higher propensity to extract home equity in the first stage (columns 2 and 4). This is evidence that fund managers are aware of home price growth and that extractions of home equity signal a subsequent decline in fund performance.

2.7 Conclusion

Numerous papers have focused on overconfidence behavior of retail investors, but the majority of these studies have only looked at static measures, and the literature on overconfidence behavior of institutional investors is scant. This paper provides in-depth and highly robust evidence of a time varying measure of overconfidence and detailed findings on how this translates into lower performance among institutional investors. Using mutual fund holdings data, the trading behavior for fund managers is backed out, which provides detailed evidence explaining how fund managers react to positive home price shocks. Fund managers interpret changes in the value of their personal homes as feedback on their ability to pick investments. This paper finds that positive shocks to the value of the personal real estate of fund managers forecasts a decline in performance. Fund managers who are more susceptible to overconfidence have a much stronger response, particularly inexperienced, less educated, and male fund managers.

The decline in performance stems from an increase in trading (and the associated trading costs) and making worse choices in picking stocks to lever up on and to sell. Interestingly, following positive home price shocks, fund managers do not make worse choices when picking new stocks but instead become more likely to buy more of their existing positions that have underperformed. Fund managers do not engage in this type of trading behavior during periods of negative home price shocks. Overall, housing shocks affect behavior but via different channels depending on the sign of the shock. The channel of overconfidence during periods of positive shocks leads to a strong effect on risk-adjusted returns. Home equity extractions signal a significant decline in future performance as well, highlighting the importance of screening funds on the personal qualities of the underlying fund managers.

While this paper focuses purely on personal real estate, other personal finance measures such as FICO and overall household debt are other potentially relevant metrics. Although these measures could potentially raise privacy issues as they are not public information, unlike real estate. Overall, this paper documents a strong feedback loop of fund managers becoming overconfident in professional behavior from the investment selection of personal real estate. This effect very likely generalizes to the assets bought by the fund manager for their fund(s) and other asset classes (such as art), thereby potentially indicating feedback loops that could lead to a spiral of overconfidence.

2.8 Figures and Tables

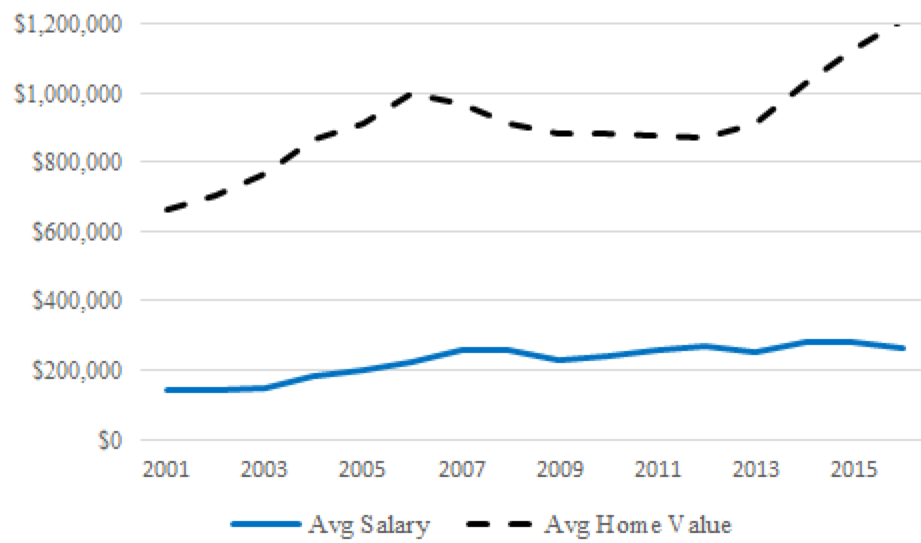


Figure 2.1: Comparison of Average Salary and Home Values for Investment Managers

A comparison of average salary and average home value for investment managers between 2001 and 2016. Data on average annual salary is from the QCEW and average home value is from ATTOM for the sample merged to Morningstar.

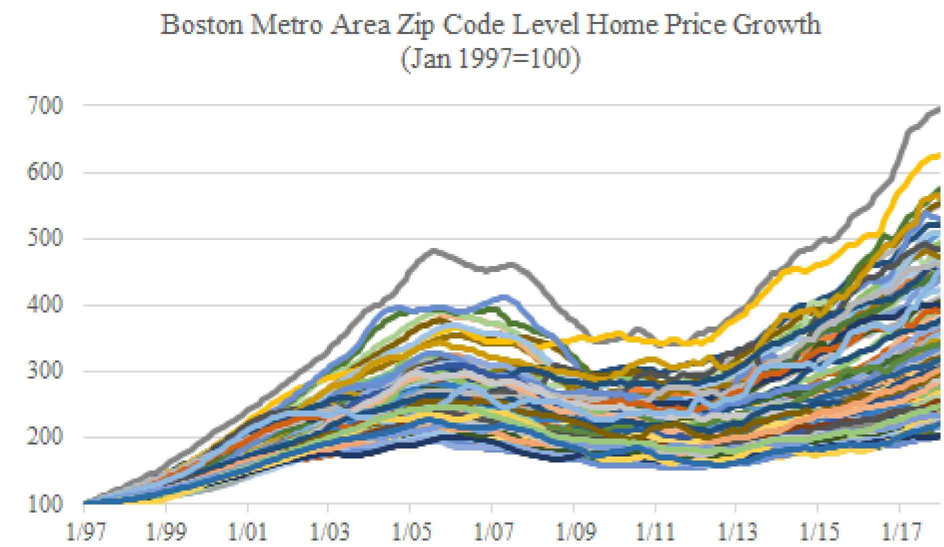


Figure 2.2: Boston Metro Area Zip Code Level Home Price Growth

Median home price growth at the zip code level for zip codes in the Boston metro area between January 1997 and December 2017. Data are from Zillow and use single-family residential and condo/co-op. Values are normalized to the value in January 1997.

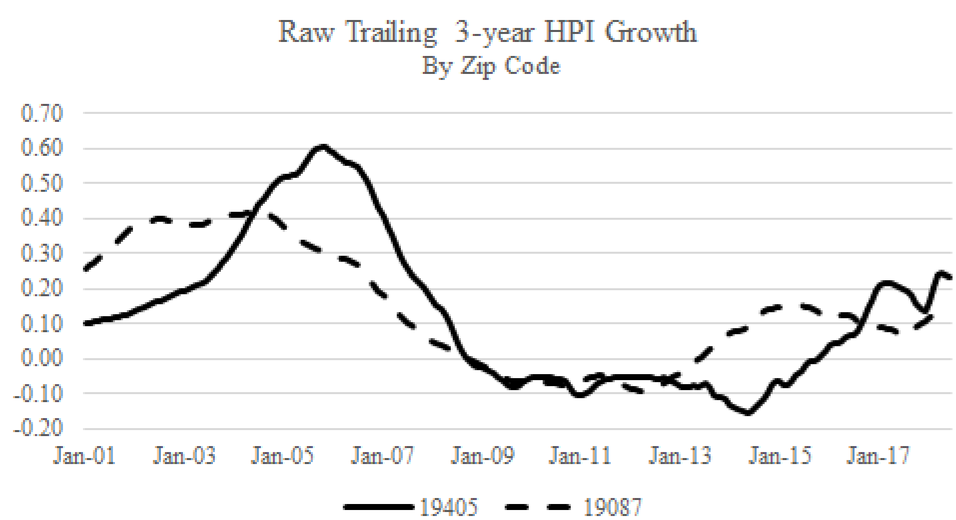
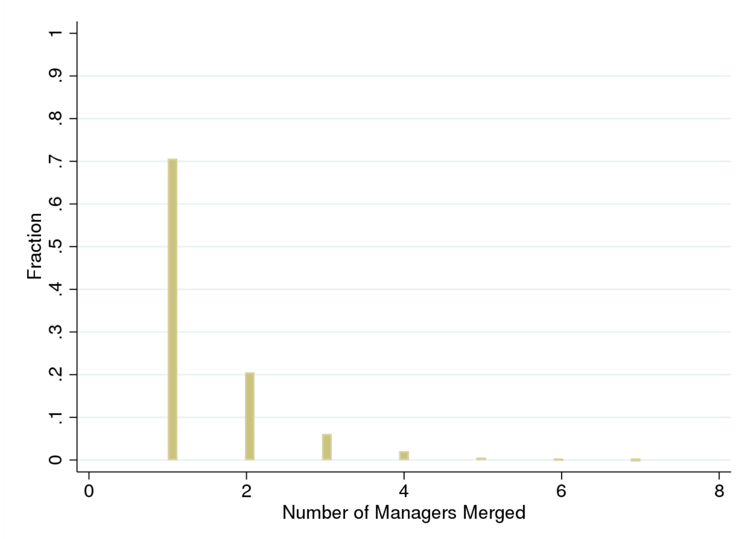
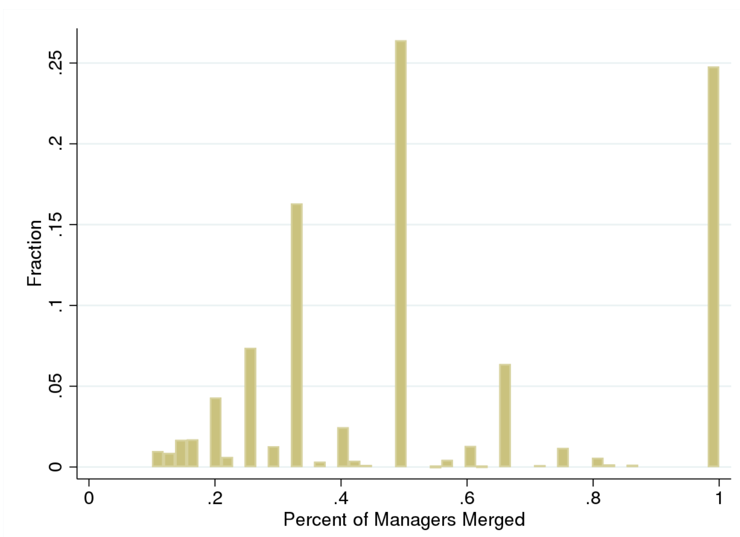


Figure 2.3: Comparison of 3-year Home Price Growth for Neighboring Zip Codes

Comparison of 3-year lagged zip code level home price growth for neighboring zip codes 19405 and 19087. Both zip codes are in the suburbs of Philadelphia.



(a)



(b)

Figure 2.4: Number and Share of Fund Managers Merged

Figure A (B) shows the number (percentage) of fund managers merged for each fund in the final population.

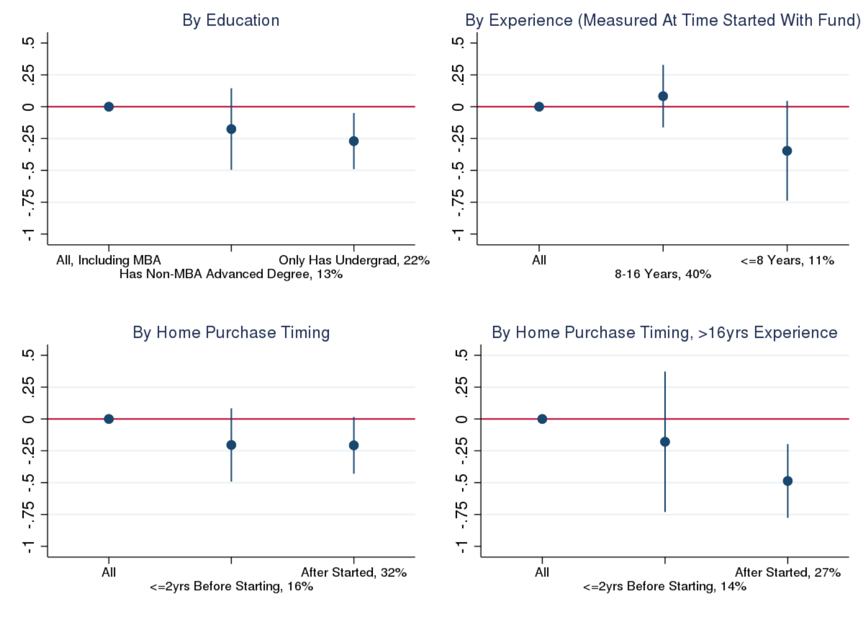


Figure 2.5: Heterogeneous Effect of Home Price Growth on Fund Alpha: 1

Estimation of the regression model:

$$\widehat{\alpha}^{FFC}_{i,t,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \sum_i \beta_i ln(\Delta 3\text{yr HPI}_{i,t-1}) * \mathbb{I}(\text{Experience Measure } i) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. $ln(\Delta 3\text{yr HPI}_{i,t-1})$ is interacted with indicator variables for measures of education, experience, or home purchase timing (measured between purchase and when the fund manager started at the fund). The omitted category is the one for the group of managers who are least likely to be affected by home price growth (i.e. those with an MBA). Percentages listed next to values on the x-axis represent the percent of the population that falls into the category listed. $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers, if TNA, measured in 2000 dollars, falls below $15million, or if home price growth is negative (only months with positive home price growth are kept). Only actively managed US domestic equity funds are included.

Figure 2.6: Heterogeneous Effect of Home Price Growth on Fund Alpha: 2

Estimation of the regression model:

$$\widehat{\alpha}^{FFC}_{i,t,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \sum_i \beta_i ln(\Delta 3\text{yr HPI}_{i,t-1}) * \mathbb{I}(\text{Experience Measure } i) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. $ln(\Delta 3\text{yr HPI}_{i,t-1})$ is interacted with indicator variables for measures of certification (CFA status), if all of the merged fund managers are male, home value at time of purchase, or LTV (including second liens) at time or purchase. The omitted category is the one for the group of managers who are least likely to be affected by home price growth (i.e. those with a CFA). Percentages listed next to values on the x-axis represent the percent of the population that falls into the category listed. $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers, if TNA, measured in 2000 dollars, falls below $15million, or if home price growth is negative (only months with positive home price growth are kept). Only actively managed US domestic equity funds are included.

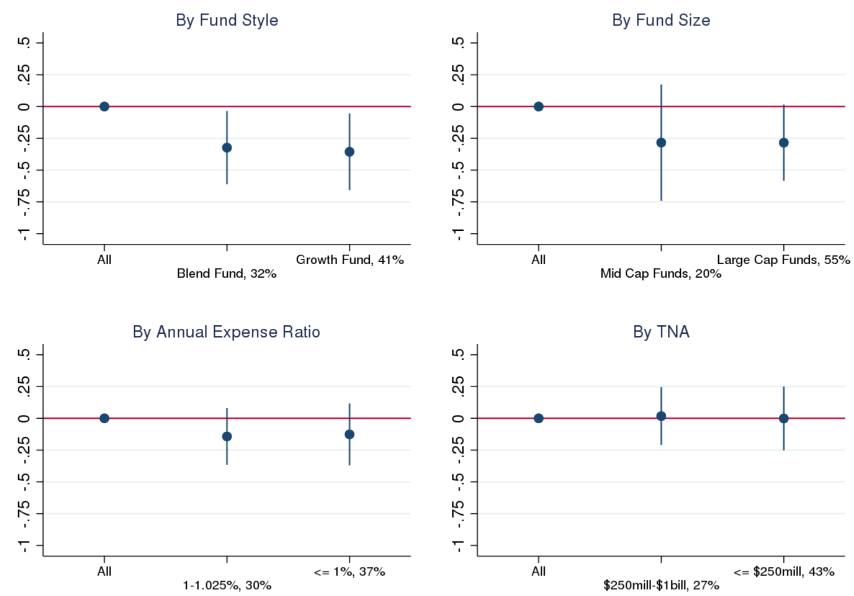


Figure 2.7: Heterogeneous Effect of Home Price Growth on Fund Alpha: 3

Estimation of the regression model:

$$\widehat{\alpha}_{i,t,a}^{FFC} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \sum_i \beta_i ln(\Delta 3\text{yr HPI}_{i,t-1}) * \mathbb{I}(\text{Fund Measure } i) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. $ln(\Delta 3\text{yr HPI}_{i,t-1})$ is interacted with indicator variables for measures of fund style, fund size focus, expense ratio, or TNA. The omitted category is the one for the group of managers who are least likely to be affected by home price growth (i.e. low expense ratio), except for fund size where the omitted category is randomly picked. Percentages listed next to values on the x-axis represent the percent of the population that falls into the category listed. $\widehat{\alpha}_{i,t,a}^{FFC}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers, if TNA, measured in 2000 dollars, falls below $15million, or if home price growth is negative (only months with positive home price growth are kept). Only actively managed US domestic equity funds are included.

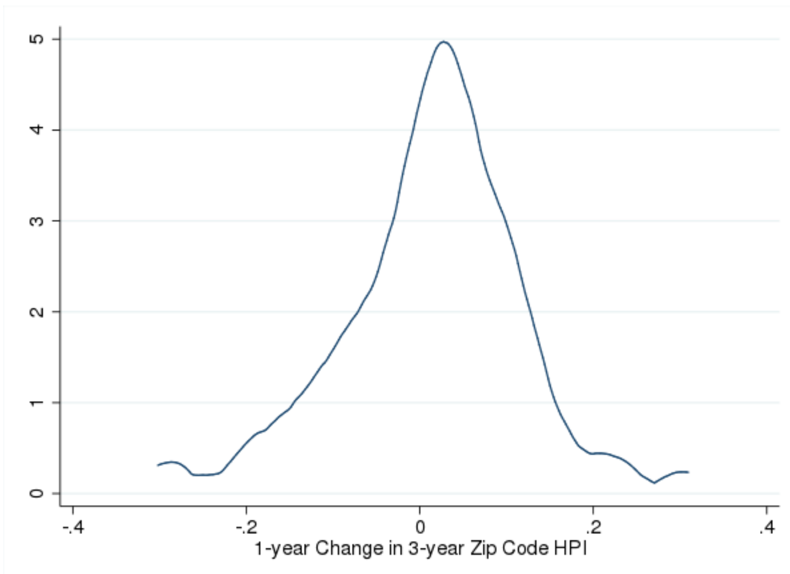
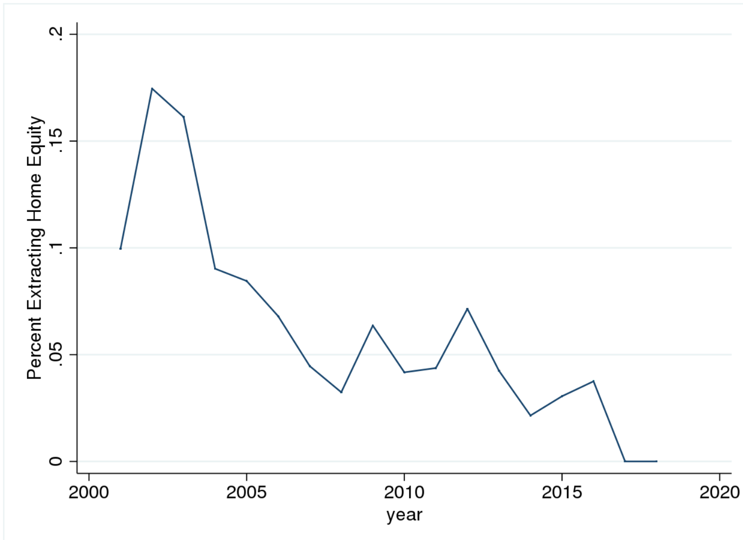
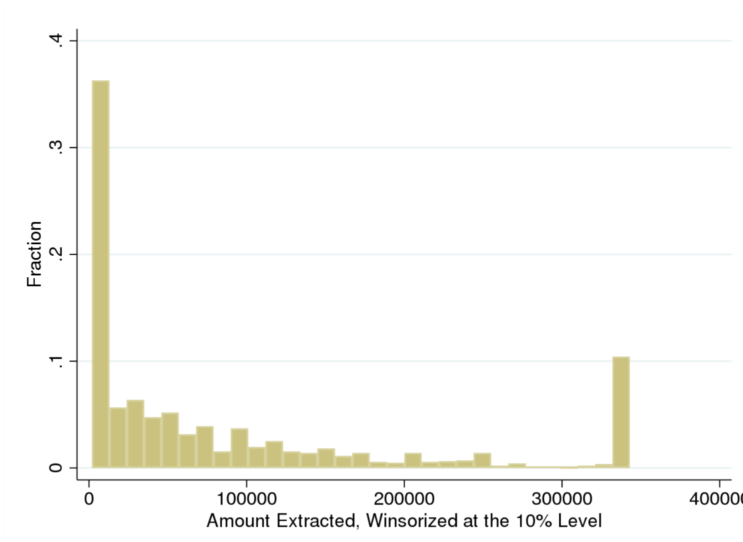


Figure 2.8: Distribution of 1-year Changes in 3-year Home Price Growth

Distribution of the 1-year change in $ln(\Delta 3\text{yr HPI}_{i,t})$, measured as of June each year.



(a)



(b)

Figure 2.9: Share of Fund Managers Extracting Home Equity and Amount of Home Equity Extracted

Figure A shows the proportion of fund managers extracting home equity in a given year. Figure B is a histogram of the amount of money extracted among fund managers who extract home equity via a cash out refinance. Values are winsorized at the 10% level, for readability.

Table 2.1: Summary Statistics Comparing the Merged and Not Merged Populations

	N	Mean	P1	P25	P50	P75	P99	Std Dev
Merged Population								
Manager Statistics								
Years Experience When Started	77619	18	5	13	17	22	39	7
Percent Female	88998	0.08	0	0	0	0	1	0.17
Percent With CFA	88998	0.56	0	0.33	0.57	1	1	0.36
Percent With MBA	86207	0.57	0	0.33	0.50	1	1	0.36
Fund Statistics								
TNA ($millions, 2000 dollars)	88998	1611	17	78	271	978	32055	5515
Annual Expense Ratio	86201	1.12%	0.20%	0.91%	1.11%	1.33%	2.15%	0.37%
Turnover	10623	0.63	0.02	0.26	0.47	0.81	2.92	0.61
Monthly α	88962	-0.09	-3.88	-0.79	-0.08	0.62	3.50	1.41
Number of Managers	88998	3.26	1	2	3	4	10	2.05
Percent of Managers Merged	88998	0.56	0.13	0.33	0.50	0.83	1.00	0.29
Not Merged Population								
Manager Statistics								
Years Experience When Started	191584	18	4	13	17	22	39	7
Percent Female	236005	0.10	0	0	0	0	1	0.23
Percent With CFA	236005	0.59	0	0	0.67	1	1	0.41
Percent With MBA	222626	0.58	0	0.14	0.63	1	1	0.40
Fund Statistics								
TNA ($millions, 2000 dollars)	236005	1040	17	74	226	781	14370	3230
Annual Expense Ratio	228436	1.18%	0.29%	0.95%	1.15%	1.39%	2.25%	0.38%
Turnover	27624	0.75	0.03	0.30	0.55	0.97	3.48	0.82
Monthly α	231528	-0.10	-4.16	-0.84	-0.08	0.66	3.83	1.46
Number of Managers	236005	2.36	1	1	2	3	8	1.59

Statistics comparing the merged population to the un-merged population for active domestic equity mutual funds. Data are aggregated from the fund by month level.

Table 2.2: Effect of Home Price Growth on Fund Alpha

	(1) All Periods	(2) HPI≥0	(3) HPI<0	(4) All Periods	(5) All Periods
$ln(\Delta 3\text{yr HPI}_{i,t-1})$	-0.171***	-0.236***	-0.0286		
	(-3.03)	(-2.73)	(-0.21)		
$ln(\Delta 3\text{yr HPI}^{+}_{i,t-1})$				-0.258***	-0.301***
				(-3.46)	(-3.45)
$ln(\Delta 3\text{yr HPI}^{-}_{i,t-1})$				-0.148	-0.129
				(-1.35)	(-1.01)
# Obs	87719	59799	27920	87719	71581
R-squared	0.0877	0.0915	0.114	0.0901	0.0903
Controls for Flow_t and $\alpha_{t-13\to t-1}$?	N	N	N	N	Y

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Baseline model for monthly forecasts of alpha, with fund and time fixed effects. Estimation of the regression model:

$$\widehat{\alpha}^{FFC}_{i,t,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table 2.3: Effect of Home Price Growth on Fund Flow

	(1)	(2)	(3)	(4)	(5)
		1-month Flow Forecast			1-year Flow Forecast
	All Periods	HPI≥0	HPI<0	All Periods	All Periods
$ln(\Delta 3\text{yr HPI}_{i,t-1})$	0.00499	0.00465	0.00887		
	(0.98)	(0.72)	(1.06)		
$ln(\Delta 3\text{yr HPI}^{+}_{i,t-1})$				0.00521	0.0435
				(0.82)	(0.43)
$ln(\Delta 3\text{yr HPI}^{-}_{i,t-1})$				0.00917	0.0163
				(1.03)	(0.11)
# Obs	83831	57202	26629	83831	71833
R-squared	0.182	0.229	0.234	0.184	0.424

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The effect of home price shocks on fund flow. Estimation of the regression model:

$$\text{Flow}_{i,t,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. Flow is calculated using gross returns and is calculated both 1-month and 1-year ahead. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table 2.4: Effect of Home Price Growth on Fund Alpha for Index and Passive Funds

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Index Funds			Lowest Quantile Tracking Error	Highest Quantile Tracking Error	Active Share $\leq 60\%$	Active Share $> 60\%$
	HPI≥0	HPI<0	All Periods	All Periods	All Periods	All Periods	All Periods
$ln(\Delta 3\text{yr HPI}_{i,t-1})$	0.0746	-0.0259					
	(0.42)	(-0.07)					
$ln(\Delta 3\text{yr HPI}^{+}_{i,t-1})$			-0.0206	-0.217	-0.585**	-0.127	-0.365***
			(-0.13)	(-1.28)	(-2.56)	(-0.79)	(-3.18)
$ln(\Delta 3\text{yr HPI}^{-}_{i,t-1})$			-0.000272	-0.113	0.0799	-0.284	-0.0198
			(-0.00)	(-0.49)	(0.19)	(-1.39)	(-0.13)
# Obs	6326	3362	9688	17616	17467	10806	47913
R-squared	0.177	0.150	0.147	0.116	0.152	0.116	0.119

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Placebo test for the baseline model for monthly forecasts of alpha using index funds and splits by tracking error and Active Share. Estimation of the regression model:

$$\widehat{\alpha}^{FFC}_{i,t,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only US domestic equity funds are included.

Table 2.5: Effect of Home Price Growth on Fund Alpha by Second Home Ownership Status

Home	(1) HPI$\geq$0 Has Second Second Home	(2) HPI$\geq$0 Does Not Have Home	(3) HPI$<$0 Has Second Second Home	(4) HPI$<$0 Does Not Have
$ln(\Delta$3yr HPI, Primary Home$_{i,t-1})$	-0.313	-0.213**	-0.334	0.0796
	(-1.47)	(-2.18)	(-0.97)	(0.54)
$ln(\Delta$3yr HPI, Second Home(s)$_{i,t-1})$	0.241		0.159	
	(1.47)		(0.87)	
# Obs	8931	50868	4751	23169
R-squared	0.124	0.0923	0.131	0.116

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robustness of the baseline model for monthly forecasts of alpha, with fund and time fixed effects, using data for all residential real estate holdings of fund managers. Estimation of the regression model:

$$\widehat{\alpha}^{FFC}_{i,t,a} = c + \beta_1 ln(\Delta 3\text{yr HPI Primary Home}_{i,t-1}) + \beta_2 ln(\Delta 3\text{yr HPI Secondary Home(s)}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. Home price growth is calculated separately for primary and secondary home(s). $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table 2.6: Effect of Home Price Growth on Returns for Buy and Sell Trades: 1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	$R^{\text{Buy}}_{i,t+1} - R^{\text{Sell}}_{i,t+1}$	$R^{\text{New Purchases}}_{i,t+1} - R^{\text{Bench}}_{i,t+1}$	$R^{\text{Buys of Existing}}_{i,t+1} - R^{\text{Bench}}_{i,t+1}$	$R^{\text{Not Traded}}_{i,t+1} - R^{\text{Bench}}_{i,t+1}$	$R^{\text{Partially Sell}}_{i,t+1} - R^{\text{Bench}}_{i,t+1}$	$R^{\text{Fully Sell}}_{i,t+1} - R^{\text{Bench}}_{i,t+1}$	Flow
$ln(\Delta\text{3yr HPI}^{+}_{i,t})$	-0.578**	0.0132	-0.502*	-0.0767			.0393
	(-2.43)	(0.05)	(-1.79)	(-0.32)			(1.35)
$ln(\Delta\text{3yr HPI}^{-}_{i,t})$	-0.349	-0.151	0.185	-0.242			-.014
	(-0.85)	(-0.32)	(0.49)	(-0.61)			(-0.27)
$ln(\Delta\text{3yr HPI}_{i,t})$					-0.226	0.613***	
					(-1.13)	(2.78)	
# Obs	27460	23384	22813	23519	22981	23311	28383
R-squared	0.0716	0.183	0.204	0.186	0.187	0.165	0.279
Mean of Dep Var	0.0748	0.152	0.168	0.201	0.125	0.128	-0.0057
SD of Dep Var	2.525	2.762	2.578	2.524	2.537	3.020	0.181

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model of effect of home price growth on trading behavior and performance. Estimation of the regression model:

$$Y_{i,t+1,a} = c + \beta ln(\Delta\text{3yr HPI}_{i,t}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{i,t,a}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are quarterly time fixed effects (interacted with indicators for positive and negative home price growth), and ρ_a are non-parametric fixed effects for TNA. $Y_{i,t+1,a}$ are: the difference in return between buy and sell trades less the benchmark return, the return on purchases of stocks in t that were not held in quarter $t-1$ less the benchmark return, the return on purchases of stocks in t that were held in quarter $t-1$ less the benchmark return (at least 5% increase in number of shares), the return on stocks currently held but not traded in quarter t less the benchmark return, the return on stocks completely sold in quarter t in $t+1$ less the benchmark return, the return on stocks partially sold in quarter t in $t+1$ less the benchmark return, and flow between quarters $t-1$ and t. Data are quarterly due to the the regressions using quarterly holdings data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund quarters are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table 2.7: Effect of Home Price Growth on Returns for Buy and Sell Trades: 2

	(1) $\mathbb{I}(R^{\text{Buys of Existing}}_{i,t-1\to t} < R^{\text{Bench}}_{i,t-1\to t})$	(2) $\mathbb{I}(R^{\text{Buys of Existing}}_{i,t-1\to t} < R^{\text{Bench}}_{i,t-1\to t} * .80)$	(3) $\mathbb{I}(R^{\text{Buys of Existing}}_{i,t-1\to t} < R^{\text{Not Traded}}_{i,t-1\to t})$	(4) $\mathbb{I}(R^{\text{Buys of New}}_{i,t-1\to t} < R^{\text{Bench}}_{i,t-1\to t})$	(5) $\frac{\text{Existing Buys}}{\text{All Buys}}$	(6) $\frac{\text{Full Liquidations}}{\text{All Sales}}$
$ln(\Delta 3\text{yr HPI}^{+}_{i,t})$	0.0947**	0.114**	0.0867*	0.0306	0.0744**	-0.0711**
	(1.99)	(2.36)	(1.76)	(0.71)	(2.03)	(-2.30)
$ln(\Delta 3\text{yr HPI}^{-}_{i,t})$	-0.0887	-0.0762	-0.0709	-0.00181	-0.0241	-0.00613
	(-1.18)	(-1.01)	(-1.02)	(-0.03)	(-0.45)	(-0.13)
# Obs	29778	29778	29778	29778	29778	29778
R-squared	0.280	0.287	0.129	0.251	0.243	0.219
Mean of Dep Var	0.352	0.314	0.535	0.279	0.487	0.344
SD of Dep Var	0.478	0.464	0.499	0.448	0.334	0.313

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model of effect of home price growth on trading behavior and performance. Estimation of the regression model:

$$Y_{i,t+1,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{i,t,a}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are quarterly time fixed effects (interacted with indicators for positive and negative home price growth), and ρ_a are non-parametric fixed effects for TNA. $Y_{i,t+j,a}$ are: indicator variables for if the return on existing stocks purchased in quarter t between quarters $t-1$ and t is less than the benchmark return over the same period, 80% of the benchmark return, or the return to stocks held but not traded in quarter t, indicator variable for if the return on new stocks purchased in quarter t between quarters $t-1$ and t is worse than the benchmark return over the same period, the percent of buy trades that were increases in existing positions, and the percent of sell trades that were full liquidations. Data are quarterly due to the the regressions using quarterly holdings data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund quarters are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table 2.8: Effect of Changes in Home Price Growth on Active Trading Performance

	(1) α, Net of Trading Costs	(2) $\mathbb{I}(\alpha^{\text{Active}} > \alpha^{\text{Passive}})$	(3) Gross α	(4) $ln\frac{\text{Trading Costs}_{i,t-1\to t+1}}{TNA_{it}}$	(5) $ln(\text{Turnover}_{i,t\to t+1})$
$ln(\Delta\text{3-yr HPI}^{+}_{it}) - ln(\Delta\text{3-yr HPI}_{i,t-12mo})$	-0.374**	-0.763*	-0.275*	0.991**	
	(-2.15)	(-1.92)	(-1.67)	(2.09)	
$ln(\Delta\text{3-yr HPI}^{-}_{it}) - ln(\Delta\text{3-yr HPI}_{i,t-12mo})$	0.0680	0.208	0.0314	-0.0285	
	(0.39)	(0.63)	(0.19)	(-0.08)	
$ln(\Delta\text{3yr HPI}^{+}_{i,t})$					0.254*
					(1.96)
$ln(\Delta\text{3yr HPI}^{-}_{i,t})$					-0.0966
					(-0.66)
# Obs	1964	1964	1964	1964	6160
R-squared	0.380	0.361	0.377	0.926	0.793
Mean of Dep Var	-0.0778	0.351	-0.00182	-13.80	-0.734
SD of Dep Var	0.235	0.477	0.226	1.768	0.908

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model of effect of changes in home price growth on the long-run active trading performance relative to the fund's own benchmark. Estimation of the regression model:

$$\alpha^{\text{Active}}_{i,t-12mo\to t+12mo,cz} - \alpha^{\text{Passive}}_{i,t-12mo\to t+12mo,cz} = c + \beta(ln(\Delta\text{3-yr HPI}_{i,t}) - ln(\Delta\text{3-yr HPI}_{i,t-12mo})) + \eta_{cz,t} + Controls_{it} + \epsilon_{i,t,cz}$$

where $\eta_{cz,t}$ are annual (t is annual) time by commuting zone fixed effects (interacted with indicators for positive and negative home price growth) and $Controls_{it}$ is comprised of: Morningstar fund category fixed effects, number of manager fixed effects, an indicator for if the manager owns a second home, an indicator for if the manager has an advanced degree (above undergraduate), and non-parametric fixed effects for TNA, home value, combined LTV (CLTV), average zip code level income (from the 2010 IRS SOI), and the percent of non-white households at the zip code level (from the 2000 Census). $\alpha^{\text{Active}}_{i,t-12mo\to t+12mo,cz}$ is the monthly alpha obtained from a 2-year regression of the active returns measured between quarter 3 of year $t-1$ and quarter 2 of year $t+1$. $\alpha^{\text{Passive}}_{i,t-12mo\to t+12mo,cz}$ is the monthly alpha obtained from a 2-year regression of the returns to holding fixed the portfolio of stocks held in quarter 3 of year $t-1$ between quarter 3 of year $t-1$ and quarter 2 of year $t+1$. The left hand side is replaced with an indicator variable for if the active alpha net of trading costs beats the passive alpha (column 3) and the log ratio of trading costs over the 2-year period relative to TNA in June of year t (column 4). Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months and did not survive the period between quarter 3 of year $t-1$ and quarter 2 of year $t+1$ are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Funds with more than 10 managers or if TNA, measured in 2000 dollars, falls below $15million are also dropped. Only actively managed US domestic equity funds are included. Column 5 uses the sample from Table 2 and estimates the effect of home price growth on 1-year ahead turnover, includes ln lagged turnover as a control, and uses annual data.

Table 2.9: Effect of Home Price Growth on Risk Taking

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		σ(Returns)			Tracking Error			CAPM Beta (Market Risk)	
	HPI≥0	HPI<0	All Periods	HPI≥0	HPI<0	All Periods	HPI≥0	HPI<0	All Periods
$ln(\Delta 3\text{yr HPI}_{i,t-1})$	-0.110***	0.0670*		-0.125***	0.0769*		-0.0932**	0.0561	
	(-2.92)	(1.92)		(-2.75)	(1.94)		(-2.47)	(1.65)	
$ln(\Delta 3\text{yr HPI}^{+}_{i,t-1})$			-0.0769**			-0.0905**			-0.0804**
			(-2.27)			(-2.16)			(-2.48)
$ln(\Delta 3\text{yr HPI}^{-}_{i,t-1})$			0.00628			0.0178			0.00586
			(0.17)			(0.40)			(0.17)
# Obs	48610	24757	73367	48623	24781	73404	48623	24781	73404
R-squared	0.879	0.941	0.900	0.820	0.922	0.866	0.576	0.705	0.568
Mean of Dep Var	1.238	1.544	1.341	0.916	1.292	1.043	0.0143	0.0355	0.0214
SD of Dep Var	0.408	0.388	0.427	0.389	0.380	0.425	0.228	0.161	0.208

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regressions showing the effect of home price shocks on risk taking. Estimation of the regression model:

$$ln(\text{Risk Taking}_{i,t\to t+12,a}) = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. Risk taking variables are: standard deviation of returns, tracking error, and CAPM beta. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table 2.10: Effect of Home Equity Extraction on Fund Performance

	(1) All Managers, OLS	(2) Extractors, 1st Stage	(3) Extractors, 2nd Stage	(4) All Managers, 1st Stage	(5) All Managers, 2nd Stage
$\mathbb{I}(\text{extracted}_{it})$	-0.0354		-2.526**		-3.179*
	(-1.64)		(-2.03)		(-1.73)
$ln(\Delta 3\text{yr HPI}_{i,t})$		0.0787***		0.0473**	
		(2.69)		(2.14)	
# Obs	26647	16088	16083	26647	26638
R-squared	0.558	0.117	0.144	0.137	0.137

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Model of effect of home equity extraction on changes in performance and tests for if fund managers are paying attention to home price growth. Estimation of the regression model:

$$\overline{\alpha_{i,t+1,a}} - \overline{\alpha_{i,t-1,a}} = c + \beta \mathbb{I}(\widehat{\text{extracted}}_{it}) + \lambda_i + \sigma_t + \rho_a + \overline{\alpha_{i,t-1,a}} + \epsilon_{i,t,a}$$

where $\overline{\alpha_{i,t,a}}$ is the average monthly alpha in quarter t, $\mathbb{I}(\text{extracted}_{it})$ is an indicator variable equal to 1 if a fund manager extracted personal home equity in quarter t. λ_i are fund fixed effects, σ_t are quarterly time fixed effects, and ρ_a are non-parametric fixed effects for TNA. t is measured as quarterly observations and standard errors are clustered at the fund for each merged non-disjoint period of time level. $\overline{\alpha_{i,t,a}}$ are the average of the monthly alphas in quarter t where the monthly alphas are obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. 2SLS regression is estimated in columns 3 and 5, where the 1st stage instruments the indicator for extracting home equity in quarter t with 3-year home price growth (the first stage results are presented in columns 2 and 4). Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Chapter 3

Duration Measurement and Hedging Channels for GSE Insured Mortgage Backed Securities[1]

3.1 Introduction

In most academic term-structure models, the interest rate risk of any particular bond can be exactly replicated with a portfolio of bonds of other maturities, so shocks to supply or demand for any bond maturity must affect the entire yield curve. However, under the "preferred habitat" view of interest rates, first described by Culbertson (1957) and modeled theoretically by Vayanos and Vila (2009), investor clienteles have preferences for particular maturities, so shocks local to a particular maturity may affect that interest rate without affecting interest rates of other maturities. This view is consistent with a large literature over the years showing the importance of local supply and demand shocks for the level of interest rates with specific maturities (e.g., Greenwood and Vayanos, 2010, 2014; Modigliani and Sutch, 1966; Ross, 1966; Wallace, 1967; Krishnamurthy and Vissing-Jørgensen, 2011; Gagnon, Raskin, Remache, and Sack, 2011).

Historically, this literature has focused on regulation changes or other government actions over the years that have significantly affected the supply or demand for bonds of a particular maturity. Due to the sheer size of the outstanding stock of mortgage-backed securities in the U.S. (about $13 trillion pre-crisis) and recent crisis-related stabilization policy initiatives on the part of the Federal Reserve Board, such as Quantitative Easing I–III and Operation Twist, that have specifically targeted the purchase of mortgage-backed securities, it is important both for our general understanding of the term structure and for evaluating and modifying these intervention policies to fully understand the exact channels through which shocks to mortgage backed securities affect the Treasury yield curve.

Recently, some authors (in particular, Hanson, 2014; Malkhozov et al., 2016) have ar-

[1]Co-authored with Aya Bellicha, Richard Stanton, and Nancy Wallace.

gued that the negative convexity of mortgages and mortgage-backed securities means that the change in aggregate interest rate risk caused by duration shifts in these securities is of comparable magnitude to that caused by government interventions, and that this can lead to an amplification of bond-market shocks. Empirically, these authors look for a relationship between bond risk premia and mortgage duration by regressing excess bond returns on mortgage duration and other controls, such as the Cochrane and Piazzesi (2005) factor. When we study these regressions in more detail, we find that the evidence is rather weaker than it first seems, being very dependent on the exact sample period chosen, and also driven at least in part by modeling changes by Barclays during the period. Additionally, Barclays duration is less anchored to reality based on having a weak relationship to underlying mortgage prepayments. However, a more fundamental problem is that the mechanism proposed to explain these results relies on MBS holders buying long term Treasuries when duration is low. Analyzing the behavior of the majority of MBS investors, we find that investors generally do not act in this way.

To identify the impact of duration and MBS holdings on the Treasury positions of banks we use call report data. We are able to estimate the change in Treasury holdings separately for banks that hedge and banks that do not hedge, identified by their holdings of interest rate derivatives, in response to duration. For foreign investors we use data from the Treasury International Capital (TIC) dataset that lists monthly Treasury holdings by country for major foreign holders. The GSEs hold a very small amount of Treasury debt, however they are known to be the most aggressive hedgers. Using quarterly data from FHFA/OFHEO regulatory reports we study the impact of the GSE hedging positions on excess bond returns. Unfortunately, for pensions and retirement funds we are not able to observe Treasury or MBS holdings. The annual comprehensive annual financial reports (CAFRs) for pensions and retirement funds only report the percentage of their assets in US bonds, which includes both Treasuries and MBS (among other assets). We estimate the effect of duration on Treasury holdings of mutual funds using data from CRSP. Lastly, using NAICS data we estimate the probability of life insurance companies buying Treasuries in a given month based on lagged MBS duration.

We find that the only investors that may follow the models of Hanson (2014) and Malkhozov et al. (2016) are life insurance firms. We also find a relation with banks however we cannot rule out that this is merely correlation. Life insurance firm market share has declined over the period, dropping below 10% since 1996 and reaching 4% in 2016. Of the investors we are not able to study, hedge funds and pensions/retirement funds are the two investor groups that may trade along the Hanson (2014) and Malkhozov et al. (2016) models. These two investor groups held almost 25% of the Agency MBS market (including households and non profit organizations) in the late 1990s, however post crisis their share has fallen below 10%.

3.2 Duration

Duration is a name given to a number of measures of interest-rate sensitivity. For a fixed-coupon bond, the first such measure, Macaulay duration (Macaulay, 1938), is a weighted average of the time to each payment, the weights proportional to the PV of each payment (discounting using the bond's yield to maturity, y). For a bond with yield to maturity y compounded n times per year, its Macaulay duration, D_{mac}, is

$$D_{\text{mac}} = \frac{1}{P}\sum_{i=1}^{n} t_i \times \frac{C_i}{(1+y/n)^{nt_i}},$$

where t_i is the time (in years) until the ith payment. A more commonly used measure in practice is modified duration, D_{mod} (Hicks, 1939):[2]

$$D_{\text{mod}} = \frac{D_{\text{mac}}}{(1+y/n)}.$$

It is a simple matter to show that

$$D_{\text{mod}} = \frac{D_{\text{mac}}}{(1+y/n)} = -\frac{1}{P}\frac{\partial P}{\partial y}, \tag{3.1}$$

so for small changes in y, the bond's price changes by approximately

$$\begin{aligned}\frac{\Delta P}{P} &\approx \frac{-D_{\text{mac}}}{1+y/n} \times \Delta y \\ &= -D_{\text{mod}} \times \Delta y.\end{aligned}$$

To hedge a portfolio, we add another security until the portfolio's overall duration is zero.

The definition of both Macaulay and modified duration requires the cash flows on the security to be fixed. Duration can, however, be extended to securities with interest-rate-dependent cash flows (e.g., securities with embedded options) by using Equation (3.1) as the *definition* of "effective duration,"

$$D_{\text{eff}} = -\frac{1}{P}\frac{\partial P}{\partial y}.^{3}$$

All of these measures relate a bond's price to changes in its own yield, which can cause problems when aggregating to the portfolio level, since the yields on different bonds do not necessarily move exactly together. More consistent is to measure the sensitivity of all bonds to movements in the *same* underlying state variable. Fisher and Weil (1971) duration

[2]Note that Macaulay and modified duration coincide with continuous compounding, where $n \to \infty$.

[3]When modified duration can be calculated, it is always equal to effective duration. However, effective duration can be calculated for a wider range of securities.

is similar to Macaulay duration, but each cash flow is discounted using the appropriate-maturity spot interest rate. It thus measures a bond's sensitivity to parallel shifts in the entire (not necessarily flat) yield curve. Similarly, in models based on the dynamics of the short-term riskless rate, r (e.g., Vasicek, 1977; Cox, Ingersoll, and Ross, 1985), we can calculate the duration of a security relative to r,

$$-\frac{1}{P}\frac{\partial P}{\partial r},$$

and in models with more than one factor, we can calculate durations with respect to each of the underlying state variables.[4]

Given an interest-rate model, the derivatives above can be evaluated either in closed form or numerically. However, any errors in the model will give rise to errors in the resulting durations and hedge ratios. An alternative, model-free, technique is to calculate a security's "empirical duration" (see, for example, Hayre, 2001, Chapter 14), in which returns on the security are regressed on changes in one or more state variables to estimate directly the average change in price for a given change in the underlying variable.[5] An extension of this idea is key-rate duration (Ho, 1992), where a multivariate regression is run of returns against changes in several different interest rates, thus estimating the sensitivity of the security to changes in each of these "key rates" keeping the other rates constant.

3.2.1 Measuring the duration of GSE MBS

Hanson (2014) and Malkhozov et al. (2016) focus on the role of prepayment related shocks to the duration of outstanding residential mortgage backed securities that, in turn, lead to large-scale shocks to the quantity of interest rate risk borne by professional bond investors. Both papers use duration measures obtained from Datastream that are the product of proprietary prepayment models developed Barclays Capital, formerly Lehman Brothers. The Barclays U.S. MBS index covers mortgage backed pass-through securities guaranteed by Government National Mortgage Association (GNMA), the Federal National Mortgage Association (FannieMae), and the Federal Home Loan Mortgage Corporation (FreddieMac), collectively known as U.S. Agency MBS. The index is composed of pass-through securities backed by conventional fixed-rate mortgages. The MBS index does not include non-agency or private-label MBS (e.g.,MBS backed by Jumbo, Alt-A, or subprime mortgages).

Malkhozov et al. (2016) use a duration-to-worst measure (LHMNBCK(DU) in Datastream) which is an MBS duration computed using the bond's nearest call date or maturity, whichever comes first. They then scale their duration-to-worst measure by the average unit price of U.S. agency MBS which produces a dollar duration per unit of MBS in the market not the aggregate MBS dollar duration. A limitation of this duration measure is that it

[4]Hedging a portfolio in a (say) two-factor world involves adding at least two additional securities until both durations equal zero.

[5]This assumes that the sensitivity remains fixed over the period of the regression (see Boudoukh, Richardson, Stanton, and Whitelaw, 1995, for a discussion and extensions).

ignores future cash flow fluctuations due to embedded optionality. Since the intent of their empirical exercise is to measure the impact of prepayment related shocks on the quantity of interest rate risk, a duration-to-worst measure would not be an accurate control for the channel of interest.

Hanson (2014) uses two measures of duration also constructed using data from Barclays Capital models and obtainable from Datastream. The first of these is an effective duration (corresponding to (LHMNBCK(DM) in Datastream) for the Barclays MBS Index and measures the percentage change in U.S. agency MBS market value following a shift in the yield curve. His second preferred duration measure is the contribution of MBS bonds to the Barclays Aggregate Index duration. This measure is constructed by weighting the effective duration measure by the ratio of the market value of MBS, using Barclays U.S. Mortgage Backed Securities – Market Value (MM), to the Barclays measure of the U.S. Aggregate – Market Value (MM). This scaled duration measure is therefore $Effective\,Dur_t \frac{MBSMV_t}{AGGMV_t}$ where $MBSMV$ is U.S. Mortgage Backed Securities – Market Value (MM) and $AGGMV_t$ is the U.S. Aggregate – Market Value (MM). The measure, captures the fact that shifts in MBS duration in the U.S. have had a growing impact on aggregate bond market duration due to the growth of the MBS market. The measure proxies for the transient component of aggregate bond market duration due to MBS and constitutes his preferred forecasting variable.

We apply a "prepayment model free" empirical duration measure using the universe of outstanding 30-year Fannie Mae (FNMA), Freddie Mac (FHLMC), and Ginnie Mae (GNMA) MBS. Our empirical duration estimates the sensitivity of daily MBS price changes to daily changes in 10-year Treasury yields. We use 10-year zero coupon Treasury yields from Gürkaynak, Sack, and Wright (2007) and TBA prices at the agency, maturity, and coupon level from EMBS to estimate the following equation:

$$\frac{TBAPrice_{t,c,p} - TBAPrice_{t-1,c,p}}{TBAPrice_{t-1,c,p}} = \alpha + \beta \frac{yield_t - yield_{t-1}}{100} + \epsilon_{t,c,p} \tag{3.2}$$

where $-1 * \beta$ is the empirical duration, t is time (daily), c is coupon (in 50bps increments), and p is program (i.e. FNMA 30-year or GNMA 15-year). The following analysis uses data for FNMA, FHLMC, and GNMA 30-year MBS with a coupon between 2.5 and 10%.

For our second duration measure, we scale our empirical duration by the market value of the outstanding stock of U.S. MBS and the market value of the Barclays Aggregate following Hanson (2014).[6] Our second measure is thus, $Empirical\,Dur_t \frac{MBSMV_t}{AGGMV_t}$, where $MBSMV_t$ is the EMBS measure of the market value of the outstanding stock of U.S. agency MBS and $AggMV_t$ is the U.S. Aggregate – Market Value (MM). Our final preferred duration measure is our empirical duration measure times U.S. Mortgage Backed Securities – Market Value (MM) to Barclays aggregate effective, duration obtained from DataStream, times U.S. Aggregate – Market Value, $\frac{EmpDur_t MBSMV_t}{AggDur_t AggMV_t}$. This is the relative contribution of MBS dollar duration to the aggregate dollar duration.

[6]Our EMBS measure of the market value of the outstanding stock of U.S. agency MBS exactly matches the Barclay measure U.S. Mortgage Backed Securities – Market Value (MM).

Malkhozov et al. (2016) notes that the duration channel is stronger when GSE share is higher. They base this on the correlation between the rolling R-squared from the regression of excess bond return on duration and the share of MBS held by the GSEs. This contrasts with the implications of the model in Hanson (2014). Hanson notes that the MBS buyers that delta hedge bear a constant amount of interest rate risk and are not important for the channel. As shown in the GSE section, the GSEs do not influence excess bond returns by their hedging activity. The GSEs primarily use swaps, swations, Treasury futures options, Eurdollar futures options, and other interest rate derviaties as well as their issuance of Agency debt to hedge their duration exposure. They do not use Treasury debt for this purpose. Consequently an increase in GSE share should reduce the importance of the duration channel on bond returns if trading in the derivatives markets does not influence excess bond returns.

On a similar note, the model in Malkhozov et al. (2016) relies on the difference between the average MBS coupon and the 5-year swap rate. Note that the average MBS coupon does not account for the change in distribution. It also does not account for various measures that are paramount in prepayment modeling, such as: SATO, the percent underwater, FICO, etc.

3.2.2 Barclays Modeling Changes

Starting in November 2008, Barclays regularly updated their prepayment model to capture changes in the market and regulatory environment (Risa, Ibanez-Meier, Fan, and Maoui (2008) and Srinivasan and Velayudham (2010)). Primarily, Barclays is interested in capturing frictions associated with mortgage terminations. The model changes are in effect structural breaks in the data. These changes can have dramatic (and persistent) effects on duration. The Barclays effective duration measure changed by 1.64 (from 1.29 to 2.93 versus a change from 1.2 to 1.48 in the empirical duration measure) between August and September 2010, primarily due to the model change in September 2010.

Figure 3.1a shows the difference between Barclays effective duration and empirical duration. The series is roughly white noise around zero until 2008. Starting in late 2008, the model starts to deviate from the zero trend, with the biggest break in September 2010 when Barclays introduced changes to their prepayment model that had large effects on their duration measures. This is also shown in figure 3.1b, which compares the levels of the two duration series. As is clear from the two graphs, Barclays effective duration measure is nearly uniformly higher (implying the bonds are longer) than the empirical duration measure between September 2010 and October 2015. These results suggest that revisions to the Barclays prepayment model led to persistent under predictions of prepayment relative to actual GSE prepayment speeds for most of the later part of the sample.

3.2.3 Influence of Extreme Observations

Given the issues with the Barclays duration measure and the relatively stable nature of duration except for a few periods of extreme and persistent changes in duration it is natural to see if the result is driven by a few extreme observations for the subsample that does not include model changes.

We first consider our two measures of effective duration, Barclays effective duration and our empirical duration. Column 1 of Table 3.1 replicates the finding of Hanson (2014) but for the period between February 1996 and October 2015. We find a similar result. Column 2 replicates this using empirical duration and we again find a similar result, albeit slightly smaller. However, given the issues with modeling changes starting in the end of 2008 we can not be certain that the result over this period is due to portfolio rebalancing or modeling changes. To account for this, we look at the period between January 1989 and August 2008 and include period fixed effects for the two refinance periods (August 1998-January 1999 and August 2002-June 2003). Using this date range and with these fixed effects, as reported in column 3 of Table 3.1, we find the Barclays effective duration measure is no longer significant. Furthermore, only the earliest August 1998 to January 1999 refinance episode is significant.

For the empirical duration series we can include the data through October 2015, however these results may be biased by the actions of the Federal Reserve. We would expect the Federal Reserve interventions to directly influence excess 10 year Treasury returns as the Federal Reserve entered the market to influence bond yields, especially during Operation Twist. Operation Twist started in September 2011 and lasted through 2012, however, unlike the other three quantitative easing programs Operation Twist's sole intention was to buy long term Treasury debt and sell short term Treasury debt. The consequences of the Fed's trading activity should be a decrease in the excess return measure since the long rate should decline as long term debt prices are bid up and the short rate should increase as short term debt prices sell off. Since it seems unlikely that the excess bond return dynamics over this period are only responding to the fluctuations in MBS duration, we also estimate a specification that introduces period controls for QE1, QE2, QE3, Operation Twist, and the tapering period, as well as the fixed effects for the two earlier refinance episodes. For the Federal Reserve controls, we include variables equal to the dollar amount of long-term Treasuries (greater than 5 years maturity) held by the Federal Reserve during the respective period (and 0 otherwise). As shown in column 4 of Table 3.1, we find that the empirical duration variable is no longer significant. As we will discuss, given the issues with model risk for the Barclays duration measure over this later period we put more emphasis on the empirical duration result for the full sample.

Columns 5-8 of Table 3.1 repeat the regressions in columns 1-4 but use the fraction of dollar MBS duration to dollar market duration to account for growth in MBS interest rate risk relative to growth in overall interest rate risk. Over the long period studied, MBS and overall debt volumes increased dramatically. We find that the empirical duration measure is not significant when including the previously mentioned period fixed effects, however Barclays duration is always significant.

3.2.4 Mortgage Terminations

A key reason that MBS duration fluctuates from changes in interest rates is that underlying mortgage holders prepay their mortgage as interest rates fall or become less likely to prepay their mortgage as interest rates rise. Based on this, we compute a measure of the percent of outstanding mortgages that prepay each month. This measure should closely align with MBS duration, based on the mechanism through which MBS duration fluctuates.

To construct a measure we use data from McDash, which provides us with monthly mortgage payment information for every outstanding mortgage from the eight largest mortgage servicers. We restrict to 30 year fixed rate first lien mortgages held in Freddie Mac, Fannie Mae, or Ginnie Mae MBS and remove construction loans. For each month between January 1992 and April 2018, we calculate the percent of mortgages that voluntarily terminate as a fraction of outstanding mortgages, weighted by the original loan amount. We do not include mortgages that terminate due to foreclosure, servicing transfers, or that go missing in the data, as these terminations do not correlate to changes in interest rates. In a figure available upon request, we show that the dynamics of this measure match the dynamics of a similar measure calculated from EMBS loan performance data. EMBS loan performance data includes data for every mortgage in a GSE or Ginnie Mae pool (not just from the eight largest servicers). However, EMBS data do not provide us with the reason for loan termination, which includes non-voluntary terminations and servicing transfers.

Figure 3.2 compares normalized versions of the empirical and Barclays duration measures against the voluntary termination measure. The termination measure moves closely with the duration measures, as expected. Interestingly, empirical duration appears to more closely align with the measure of the underlying mortgage terminations compared to Barclays duration. This is confirmed by finding an R-squared of 0.35 when regressing empirical duration onto the termination measure, compared to an R-squared of 0.19 when regressing Barclays duration onto the termination measure. This shows that even though the empirical duration measure is "model free", it more closely tracks reality compared to the model based Barclays duration measure.

Next, we replicate the earlier results using the termination measure. In addition, McDash provides us with more recent data, which allows us to extend the analysis through April 2018. We find that for the January 1992 through April 2018 period, both Barclays duration and the termination measure are significant at the 5% level. However, when we include period fixed effects for the late 1990s and early 2000s refinance cycles and the various quantitative easing measures taken by the Federal Reserve, we find that the mortgage termination measure becomes insignificant. Barclays duration remains significant at the 10% level. However, this result is partly based on data during the period for which Barclays actively adjusted their model. Lastly, when we restrict to the period after the quantitative easing measures by the Federal Reserve ended (January 2014 through April 2018), we find that the mortgage termination measures becomes insignificant, while the Barclays measure remains significant.

These regressions show the importance of the choice in a measure to reflect MBS interest rate risk. With a measure of realized mortgage terminations, we find a very weak relationship that is based on a few unique episodes and that has since become irrelevant in recent years. However, using a model derived measure of MBS interest rate risk, we find a stronger effect that is currently a factor influencing Treasury yields. Based on the inherent weaknesses of the Barclays model derived measure and the consistency of the results based on two independent measures of MBS interest rate risk that are based on fundamentals, we find limited support of the claim that MBS duration hedging affects Treasury yields. To further show this, in Section 3 we rigorously show that most classes of investors do not buy Treasuries to hedge MBS duration risk, which is the channel highlighted by previous work to explain a relationship between MBS duration and excess Treasury yields.

3.3 Evidence for MBS Investor Hedging

Using the Federal Reserve Flow of Funds, Agency MBS investors are presented in Figure 3.3. The investors we can account for are the Federal Reserve, banks, foreign investors, the GSEs, mutual funds, and life insurance companies. We do not account for brokers/dealers, federal/state/local government, property/casualty insurance companies, households and non-profit organizations (includes hedge funds), retirement and pension funds, issuers of ABS, and REITs. Overall, we are able to estimate the hedging response of investors comprising at least 70% of the MBS market since the first quarter of 2001 (barring a few months where their share dipped slightly below 70%) and in some months more than 80% of the market.

3.3.1 The Federal Reserve

Since November 2008 the Federal Reserve has purchased over $2.3 trillion dollars of mortgage backed securities through its quantitative easing programs (Federal Housing Finance Agency Officer of Inspector General, 2014). By the first quarter of 2014, the Federal Reserve held $1.5 trillion Freddie Mac PCs, Fannie Mae mortgage backed securities, and GNMA mortgage backed securities on its balance sheet (Patrabansh, Doerner, and Asin 2014). The Federal Reserve does not hedge nor do they care about portfolio duration so this important investor cannot important for the measurement of the overall U.S. GSE MBS duration. Malkhozov et al. (2016) finds that the Federal Reserve's market share is negatively correlated with the strength of the duration channel. This is attributed to the Federal Reserve abstaining from hedging. However, the period when Federal Reserve MBS holdings are higher is the period during unconventional monetary policy and thus it is difficult to attribute the reduction in the strength of the duration channel over this period to Federal Reserve MBS holdings or Federal Reserve open market operations more broadly (such as Operation Twist).

3.3.2 Government Sponsored Enterprises

Malkhozov et al. (2016) notes that the duration channel is stronger when GSE share is higher. They base this on the correlation between the rolling R-squared from the regression of excess bond return on duration and the share of MBS held by the GSEs. This contrasts with the implications of the model in Hanson (2014). Hanson notes that the MBS buyers that delta hedge bear a constant amount of interest rate risk and are not important for the channel. As shown below, the GSEs do not influence excess bond returns by their hedging activity. The GSEs primarily use swaps, swations, Treasury futures options, Eurdollar futures options, and other interest rate derivatives as well as their issuance of Agency debt to hedge their duration exposure. They do not use Treasury debt for this purpose. Consequently an increase in GSE share should reduce the importance of the duration channel on bond returns if trading in the derivatives markets does not influence excess bond returns.

Figure 3.4 shows the holders of Treasury debt. The GSE share is not visible, rising to a maximum of only 1.8% over the period. Figure 3.5 shows the distribution of the GSEs derivatives portfolio revealing that interest rate swaps are the dominant instrument used for hedging. Note, Figure 3.5 does not show the non-mortgage investments portfolio, which comprises Treasury debt.

However, Treasury debt is not broken out separately and is rolled into the Other category, showing that it is a very small investment category for the GSEs. The GSE hedging portfolio data includes holdings of interest rate derivative products by the GSEs (as shown in Figure 3.5). Using quarterly data from the FHFA/OFHEO regulatory reports we test the effect of the GSE derivatives portfolio on log excess return for 10-year zero coupon bonds. The following regression is estimated:

$$rx^{(10)}_{t+12m} = \alpha + \beta_1 \times GSEHedgingPortfolio_t + \epsilon^{(10)}_{t+12m} \tag{3.3}$$

Equations including Barclays effective and empirical duration are also estimated. The data are quarterly and standard errors are Newey West allowing 18 months of serial correlation. Regressions of changes in duration and interest rate derivative holdings on the change in excess return are also estimated.

If the GSEs buy more interest rate derivatives when duration declines and the buying of these contracts influences excess bond returns then β_1 should be negative. Tables 3.3 and 3.4 show that the GSE hedging portfolio does not have an impact on the excess return and the coefficient is positive. Including duration or estimating the regression with differences instead of levels does not change the finding.

Note, the Flow of Funds data (L.211) measures GSE holdings of agency and GSE backed securities. However, the GSEs hold a large amount of whole loans and private label securities (PLS), which are not included in the Flow of Funds data (Figure 3.6).

Further complicating matters, in Q1 of 2010 there was an accounting policy change that dramatically reduced reported GSE holdings (and increased their liabilities). The retained portfolio reporting data does not show a similar drop in Q1 2010. Using FHFA/OFHEO retained portfolio reporting data we find that before 2002 the retained portfolio dynamics followed the Flow of Funds data but diverged afterwards. Agency MBS holdings (FNMA, FHLMC, and GNMA MBS) in the FHFA data closely track the Flow of Funds data. However the FHFA data do not show a decline in 2010 from the accounting change.

Overall, the GSEs do not directly hold Treasuries, however they do hold very large interest rate derivative portfolios. We test whether these influence excess Treasury yields and find that they do not.

3.3.3 Banks

Using call report data we test the impact of duration on Treasury bond holdings (and more broadly non-mortgage debt holdings by maturity) of banks. We divide the sample into banks with MBS pass-through holdings in their non-trading accounts that have and have never held interest rate contracts in their non-trading accounts as well as banks that never held MBS securities as a comparison. Figure 3.7 shows the holdings of MBS pass-through by account for banks that have held interest rate contracts and for banks that have not held interest rate contracts. Most banks have never held an interest rate contract. The majority of the bank pass-through holdings are held by the smaller subset of banks that have held interest rate contracts and are kept in the available for sale account. However, this was not true before the early 2000s. The trading assets account is omitted, however a negligible amount of MBS is held in this account.

We use four different dependent variables. First, we look at the overall Treasury to asset ratio (the sum of RCFD0211 [held to maturity amortized cost] and RCFD1287 [available for sale fair value] over RCFD2170 [total balance sheet assets]). However, this does not allow us to decompose by maturity. For that we look at the non-mortgage debt to asset ratio (the sum of RCFDA549/550/551/552/553/554 [securities issued by the U.S. Treasury, U.S. Government agencies, and states and political subdivisions in the U.S.; other non-mortgage debt securities; and mortgage pass-through securities other than those backed by closed-end first lien 1-4 family residential mortgages, by remaining maturity or next repricing date] over RCFD2170). For this variable we also look at the ratio for debt with three or fewer years remaining maturity and for debt with more than three years remaining maturity. Figure 3.8 shows the unweighted average Treasury and non-mortgage debt (as defined above) to asset ratios for all banks (including those that do not hold MBS). We find that since 2002, banks have held a very small portfolio of Treasuries, however they do have sizable holdings of other non-mortgage debt (such as agency debt and municipal debt). This is logical as these forms of debt are also very safe but provide a higher yield. This is also confirmed by the flow of funds data (Figure 3.4) which showed that banks have held a negligible amount of Treasuries since 2002.

The following equation is estimated:

$$\frac{Debt_{t,i}}{Assets_{t,i}} = \alpha + \beta_1 \times Duration_{t-12m} + \beta_2 \times \frac{MBSSecurities_{t-12m,i}}{Assets_{t-12m,i}} + \beta_3 \times \frac{MBSSecurities_{t-12m,i}}{Assets_{t-12m,i}} \times Duration_{t-12m} + \mathbb{X}_{i,t-12m} + \epsilon_{t,i} \quad (3.4)$$

where i is a bank, and $\mathbb{X}_i$ are bank fixed effects and time varying controls for the ratio of residential loans to total loans, loan to asset ratio, capital and liquidity ratios, deposit and funding costs, and deposit to asset ratio. The measure of MBS securities is the sum of RCFDG300/304/308/303/307/311 (held to maturity amortized cost and available for sale fair value residential pass-through securities issued by GNMA/FNMA/FHLMC/other). The data are quarterly between June 1997 and December 2013 and standard errors are clustered at the bank level allowing for correlation in the standard errors within banks. We report results using only empirical duration, as it does not suffer from model risk.

If banks increase Treasury holdings following a decline in duration we would expect β_1 to be negative. We would also expect β_3 to be positive if banks that hold more MBS are more sensitive to duration (β_3 is expected to be positive in this case as β_2 is negative, as will be shown in the regressions below). The effect should be much stronger for banks that do not have interest rate derivative holdings and for Treasuries with longer duration (assuming banks holding interest rate derivatives hedged their MBS positions).

Table 3.5 shows results using data for banks that have held MBS in their non-trading accounts and that have never held interest rate derivatives in their non-trading accounts in columns 1-4 and banks that have held interest rate derivatives in their non-trading accounts in columns 5-8. Columns 1 and 5 show that Treasury holdings overall increase relative to assets when duration declines. Columns 2 and 6 show a different pattern for all non-mortgage debt relative to assets. Banks that do not hold interest rate derivatives decrease this variable when duration declines, however long duration debt still increases following declines in duration (column 3). Short duration bonds seem to dominate leading to the overall positive relationship found in column 2. Banks that have held interest rate derivatives increase this variable in response to declines in duration, with a negative coefficient for long duration non-mortgage debt to assets (column 7). The interaction term is not significant in columns 1 and 6-8 indicating mixed results for the sensitivity of banks relative to the size of their MBS holdings. Overall, we find a negative relation between MBS holdings and Treasury holdings indicating a possible crowding out effect or substitution effect.

As a placebo test, we replicate these regressions for the population of banks that never held MBS securities in their non-trading accounts. MBS duration should have no impact on debt holdings for these banks, however we find similar results. Table 3.6 shows that these banks do in fact increase their Treasury and long duration non-mortgage debt when duration declines (columns 1 and 3). This finding, along with the insignificant interaction terms in table 3.5 indicates that there is potentially another channel driving these results. When duration declines it is typically when the Federal Reserve is loosening policy thus indicating heightened risk aversion and low inflation. This is an environment conducive to buying safe long duration assets. Based on these regressions we cannot conclude that MBS duration results are causal.

3.3.4 Major Foreign Holders

Since 2002, foreign investors have held between 10 and 22% of outstanding Agency MBS (Figure 3.3). Furthermore, foreign investors are major holders of US Treasury debt, holding more than 40% in recent years (Figure 3.4). This is an investor group that holds a sizable MBS portfolio and also holds a significant portion of Treasury debt.

Figure 3.9 shows that China and Japan each hold roughly 10% of total US Treasury debt.

Using data from the Treasury International Capital (TIC) dataset the following equation is estimated:

$$\frac{\Delta TSYHoldings_{t,t+12m}}{TSYHoldings_t} = \alpha + \beta_1 \times \Delta Duration_{t-12m,t} + \beta_2 \times \Delta FXRate_{t-12m,t} + \epsilon_{t,t+12m} \tag{3.5}$$

where the foreign exchange rate is how many units of the country's currency buys one US Dollar. The data are monthly and the standard errors are Newey West allowing for 18 month lags. Regressions are estimated using percentage change in Treasury holdings instead of the difference in Treasury holdings to control for the massive increase in holdings over the period.

If these countries buy Treasuries when duration decreases to extend the duration of their portfolios we should find β_1 to be negative. Investors in these countries may also buy Treasuries when their currency appreciates, in which case we would expect β_2 to be negative. However, if the country's currency freely floats then low duration may also coincide with a more valuable foreign currency for the country as low duration may be from a low Fed Funds rate, which would depreciate the US dollar relative to foreign currencies. A more valuable foreign currency may lead to increased buying of US Treasury debt.

Table 3.7 shows results using Barclays effective duration. The change in duration variable is never significant. Using empirical duration we find similar results (Table 3.8). Thus, the largest foreign holders of US Treasuries do not buy Treasuries when duration declines.

3.3.5 Mutual Funds

Using data from CRSP, we test the impact of duration on Treasury bond holdings of mutual funds. The data are available quarterly between Q1 2010 (when the government bond holding percentages first started being reported) and Q2 2016. The following equation is estimated:

$$PerGovtBonds_{t,i} = \alpha + \beta_1 \times Duration_{t-12m} + \mathbb{X}_i + \epsilon_{t,i} \tag{3.6}$$

where i is a mutual fund, duration is either empirical or Barclays effective duration, and $\mathbb{X}_i$ are mutual fund fixed effects. Standard errors are clustered at the mutual fund level. Regressions are also estimated on changes instead of levels. We estimate the equation for all mutual funds in the sample and for mutual funds that ever had MBS holdings. Data on MBS holdings became available in October 2010.

If mutual funds increase Treasury holdings following a decline in duration we would expect β_1 to be negative. Table 3.9 reports results for the estimates of the equation on levels and shows that β_1 is positive and significant across all specifications. Table 3.10 reports results for changes and finds β_1 is positive and significant across all specifications. These results show that mutual funds do not increase their Treasury holdings when duration declines.

3.3.6 Life Insurance Companies

Life insurance company liabilities are generally long duration and require a minimum level of return with minimal risk. To satisfy their liabilities life insurance companies primarily hold fixed income assets. Berends, McMenamin, Plestis, and Rosen (2013) find that 75% of life insurance general account assets are in bonds. Of the bond holdings 60% are in corporate bonds, 18% are in MBS (both private label and GSE), and only 7% are in Treasuries. Corroborating this finding, figure 3.4 shows that over this period insurance companies, including property and casualty insurance companies, held a very small share of Treasury debt.

Figure 3.3 shows that insurance companies held roughly 10% of outstanding Agency MBS in 1985 however this share has declined over time (70–80% of these holdings are from life insurance companies). Using data from NAICS we are able to identify life insurance firms that hold MBS (either GSE or private label) and their monthly trading of Treasury securities. We use data from the 300 files, Schedule D Part 1, which reports securities held as of year end. There is also a 303/304 file, Schedule D Part 3, which reports all trades in a year. However data for the 303/304 files ends in 2007, while data for the 300 files ends in 2012.

The downside to using the 300 files is that the reporting will not capture securities held and then sold after a short period of time or securities that mature before the end of the year that they are purchased. However, since we are interested in Treasury holdings that are bought for the purpose of extending portfolio duration during periods of persistent declines in MBS duration this is likely to be less of an issue. In terms of absolute numbers of Treasury purchased, the 300 files contain 97% of the number of Treasury purchases compared to the 303/304 files, while the 300 files only contain 79% of all trades compared to the 303/304 files. Thus it seems that Treasuries are likely to be held for extended periods of time. However, comparing counts by tuples of firm and year and month of the trade only 82% of Treasuries in the 303/304 files are matched to the 300 files, thus there is variation in the high frequency data. For robustness, we report regressions for the overlapping period of 2001–2007 for both the 300 and 303/304 files.

To test if life insurance companies buy Treasuries when MBS duration is low we estimate a linear probability model of the probability of buying Treasuries in a given month based on lagged duration with firm fixed effects. The equation estimated is:

$$\mathbb{1}_{t,i} = \alpha + \beta_1 \times Duration_{t-j} + \mathbb{X}_i + \epsilon_{t,i} \tag{3.7}$$

where $\mathbb{1}_{t,i}$ is equal to 1 if firm i bought Treasuries in month t, $Duration_{t-j}$ is either Barclays effective MBS duration or empirical duration (where j is the lag relative to the month of buying Treasuries and is either 1, 6, or 12 months), and $\mathbb{X}_i$ are firm fixed effects. The standard errors are clustered at the firm level. Our identification of Treasuries includes TIPS and excludes when issued and STRIPS. The regressions are estimated using data for firms that held MBS securities (identification of MBS securities includes GSE debt) at any point in the data.

If life insurance companies buy Treasuries when duration is low we would expect β_1 to be negative. Table 3.11 shows that β_1 is negative and significant, indicating that life insurance firms buy Treasuries when MBS duration is low. However, the coefficient is small indicating a small relative increase in Treasury purchases. Many firms infrequently buy Treasuries. 13% never bought Treasuries and 65% bought Treasuries in 12 or fewer months over the 12 year period.

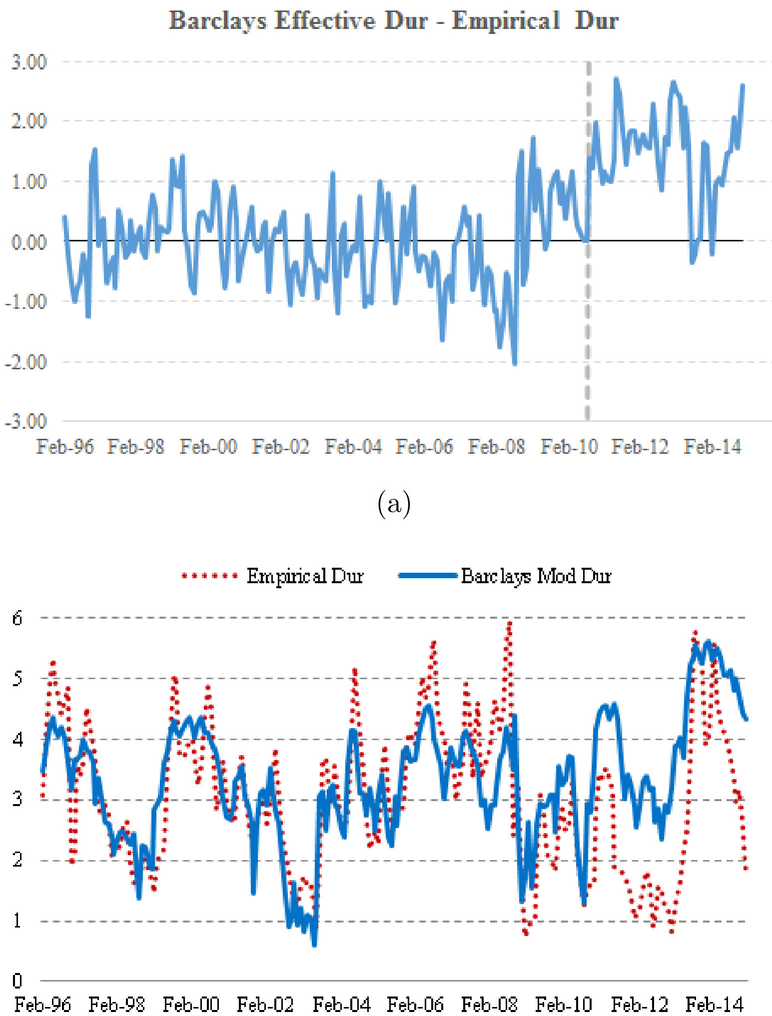
We impute the maturity of the Treasuries purchased based on the year of the trade and the stated maturity of the Treasury. Roughly 53% have a maturity of 5 or fewer years and 44% have a maturity of more than 5 years (we are missing the maturity for 3% of the bonds). Furthermore, only 33% have a maturity greater than 8 years, thus the Treasuries purchased tend to be of shorter maturity. Column 9 shows results for the linear probability model estimated with the left hand side equal to one if a Treasury with $\leq$5 years is purchased in the given month and 0 otherwise. Conversely, column 10 is estimated with the left hand side equal to one if a Treasury with >5 years maturity is purchased in the given month. Both coefficients are negative and significant.

As a robustness check, columns 11–12 show that we obtain similar results if we use either the 300 or 303/304 files over the 2001–2007 period. Table 3.12 shows results using empirical duration and results are similar.

3.4 Conclusions

We propose an empirical duration measure for the stock of U.S. Agency MBS that appears to be less prone to model risk than measures such as the Barclays Effective Duration measure. We find that this measure does not appear to have a strong effect on the 12-month excess returns of ten-year Treasuries as would be expected if shocks to MBS duration lead to commensurate shocks to the quantity of interest rate risk borne by professional bond investors (Hanson, 2014; Malkhozov et al., 2016). Given this negative reduced form result, we then explore the mortgage and treasury hedging activities of the primary MBS investors such as commercial banks, insurance companies, the agencies, the Federal Reserve Bank, mutual funds, and foreign investors. We find that the only investors that may follow the models of Hanson (2014) and Malkhozov et al. (2016) are life insurance firms and possibly banks. Life insurance firm market share has declined over the period, dropping below 10% since 1996 and reaching 4% in 2016. Of the investors we are not able to study, hedge funds and pensions/retirement funds are the two investor groups that may trade along the Hanson (2014) and Malkhozov et al. (2016) models. However, although these two investor groups held almost 25% of the Agency MBS market (including households and non profit organizations) in the late 1990s, post crisis their share has fallen below 10%.

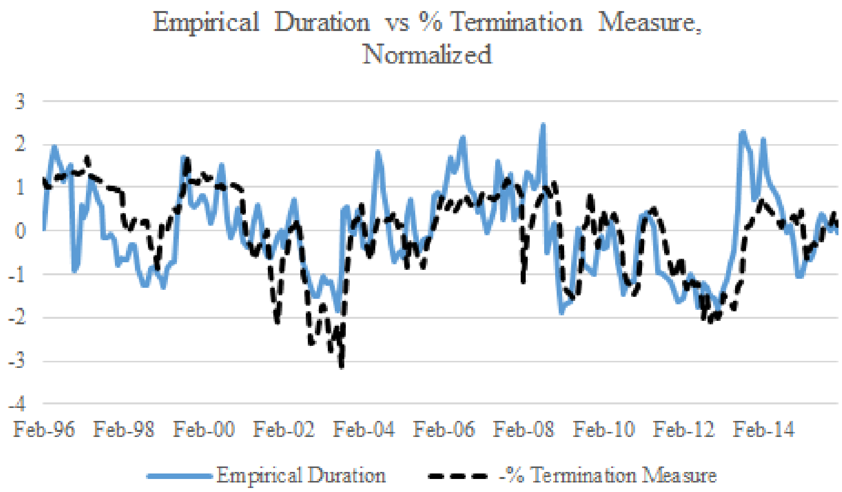
3.5 Figures and Tables



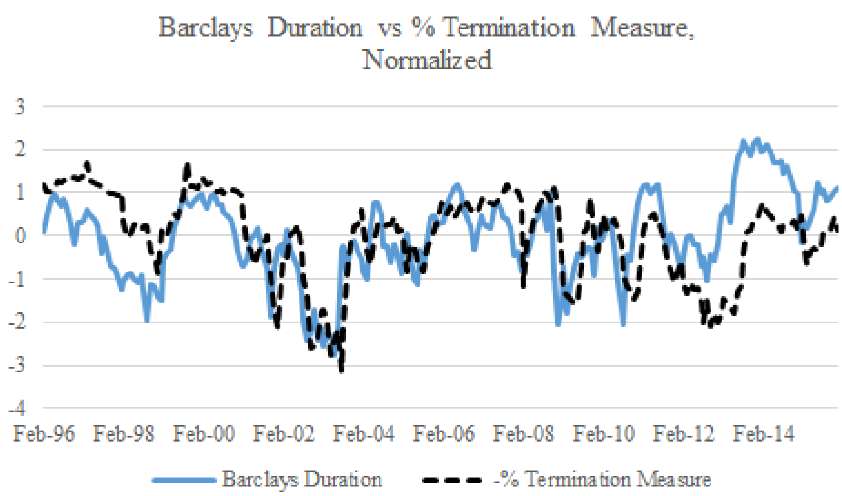
(b)

Figure 3.1: Comparison of Barclays Effective Duration and Empirical Duration

Figure A shows the difference between Barclays effective duration and empirical duration from February 1996 to October 2015. The vertical dashed gray line represents September 2010. Figure B compares the empirical duration and Barclays duration series in levels.



(a)



(b)

Figure 3.2: Comparison of Duration Measures (Empirical and Barclays effective) Against Termination Measure

Figures A (empirical duration) and B (Barclays effective duration) compare duration measures against the voluntary termination measure constructed from McDash data. The timeseries are normalized (de-meaned and divided by standard deviation) and the voluntary termination measure is multiplied by -1 to be comparable to duration.

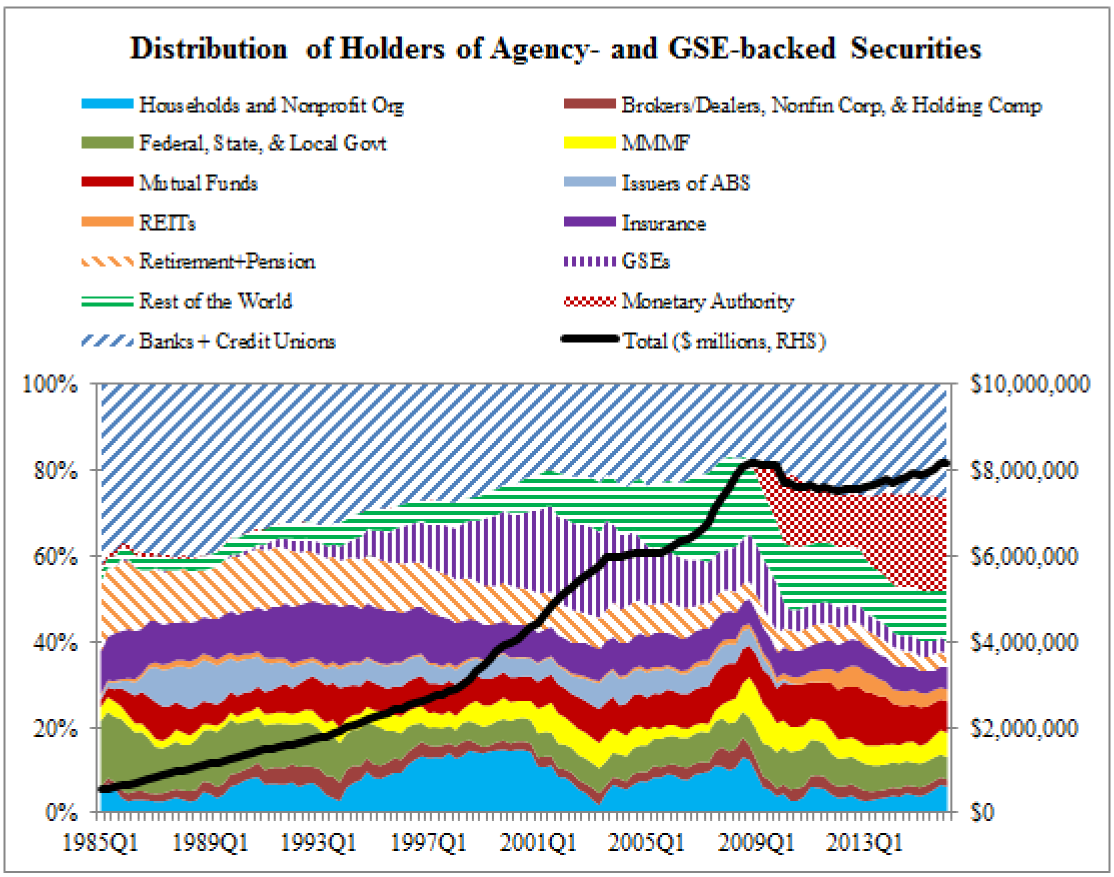


Figure 3.3: Distribution of Agency and GSE Backed Securities Holdings by Investor Group

The data are from the Federal Reserve flow of funds and are quarterly from Q1 1985 to Q1 2016 (left axis). The solid line (right axis) shows total outstanding Agency and GSE backed securities in USD millions.

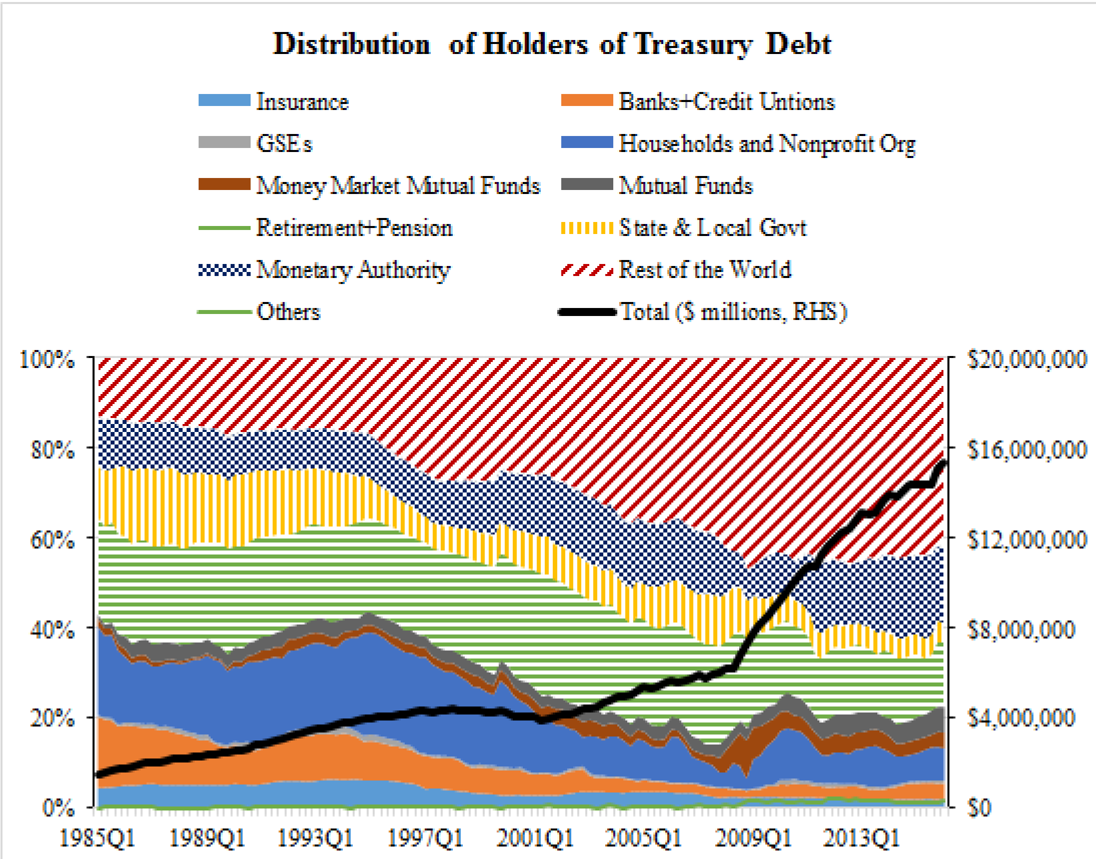
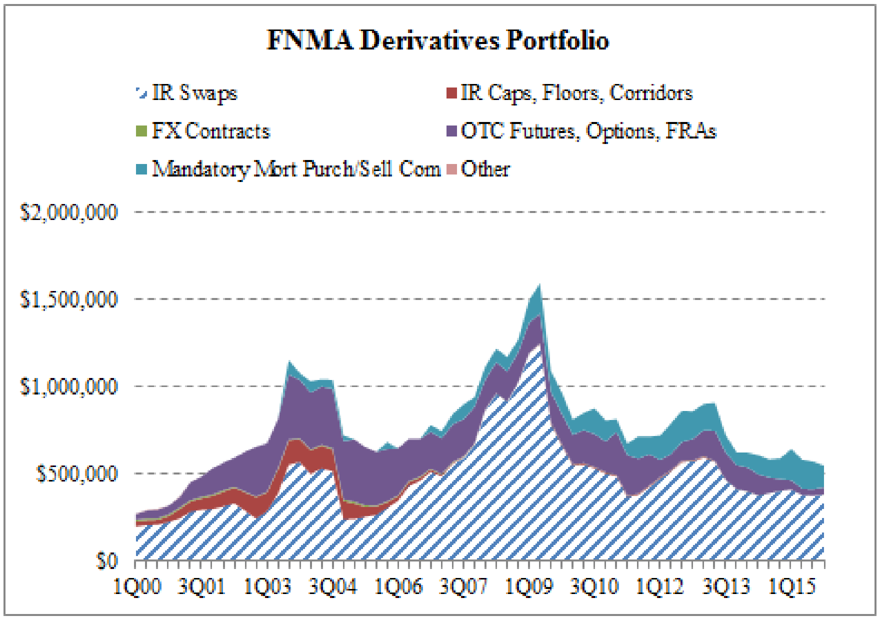
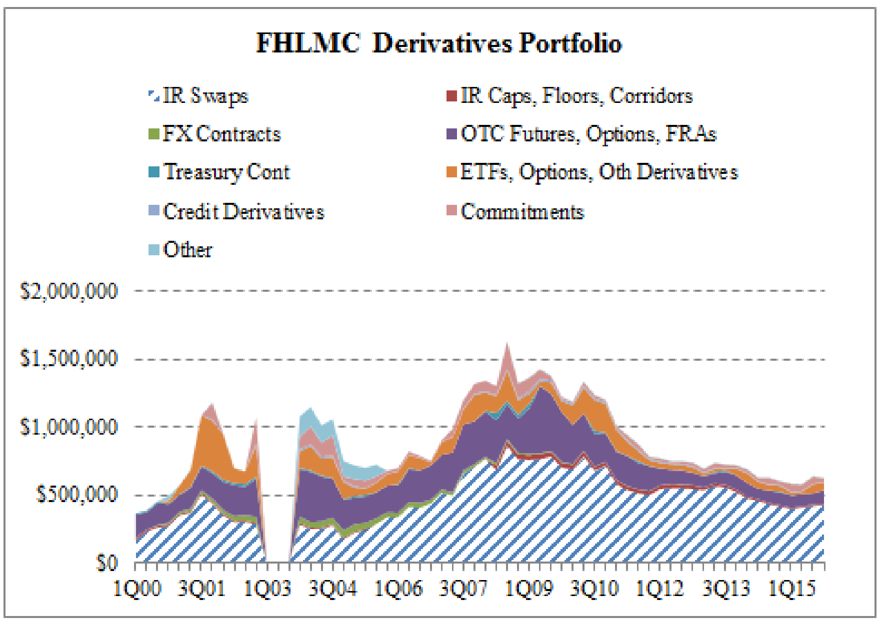


Figure 3.4: Distribution of US Treasury Holdings by Investor Group

The data are from the Federal Reserve flow of funds and are quarterly from Q1 1985 to Q1 2016 (left axis). Other includes non-financial corporate business, non-financial non-corporate business, closed end fund, exchange traded fund, issuer of ABS, security broker and dealer, and holding company holdings of US Treasuries. The solid line (right axis) shows total outstanding US Treasury debt in USD millions.



(a)



(b)

Figure 3.5: Fannie Mae and Freddie Mac Derivatives Portfolios

Figures A (Fannie Mae) and B (Freddie Mac) show the total dollar amount (notional, in millions) and composition of the financial derivatives portfolios using quarterly data from the FHFA/OFHEO annual reports to Congress between Q1 2000 and Q4 2015. The categories of securities are not consistent between Fannie Mae and Freddie Mac. The data are not available for Freddie Mac between Q1 2003 and Q3 2003.

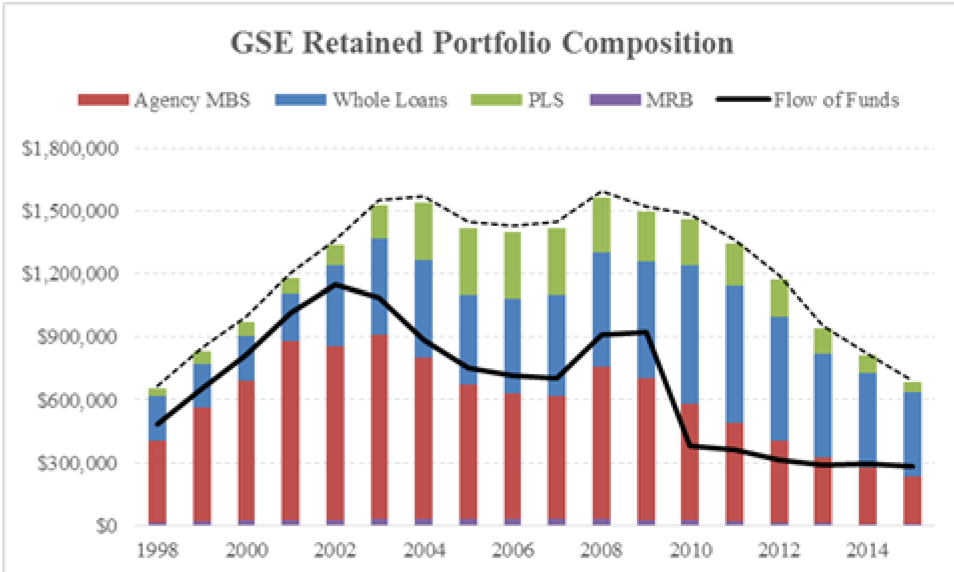


Figure 3.6: Comparison of GSE Retained Portfolio Reporting to Congress vs Federal Reserve Flow of Funds

Compares the aggregate size of the GSE retained portfolios (for Fannie Mae and Freddie Mac) as reported by the FHFA/OFHEO annual reports to Congress (dotted line) and the Agency and GSE backed holdings of the GSEs as reported by the Federal Reserve flow of funds data (solid line). Within the portfolio reported by the FHFA/OFHEO annual reports to Congress the figure shows the composition by security type (Agency MBS, whole loans, private label MBS, and mortgage revenue bonds). The data are annual between 1998 and 2015.

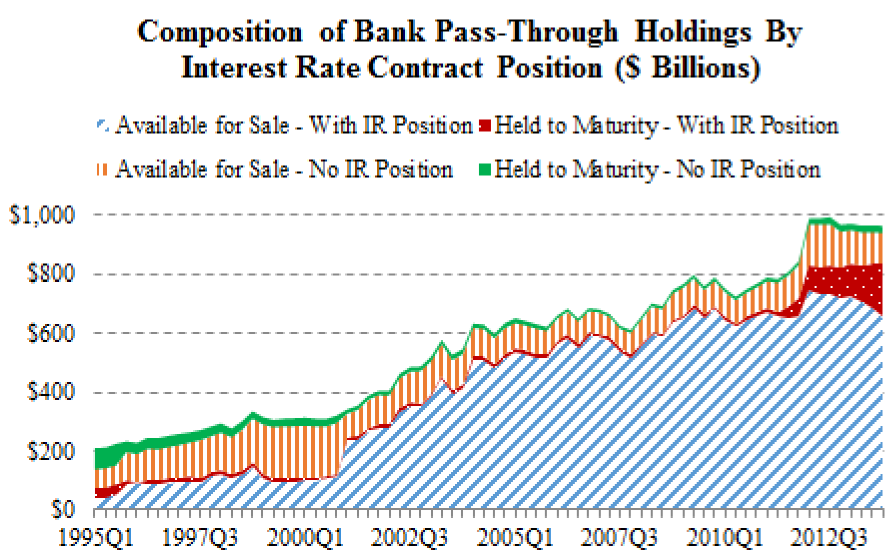


Figure 3.7: Bank MBS Pass-Through Holdings

Bank MBS pass-through holdings grouped by account and if the bank ever held an interest rate contract. The data are from the call reports and are quarterly between 1995 and 2013.

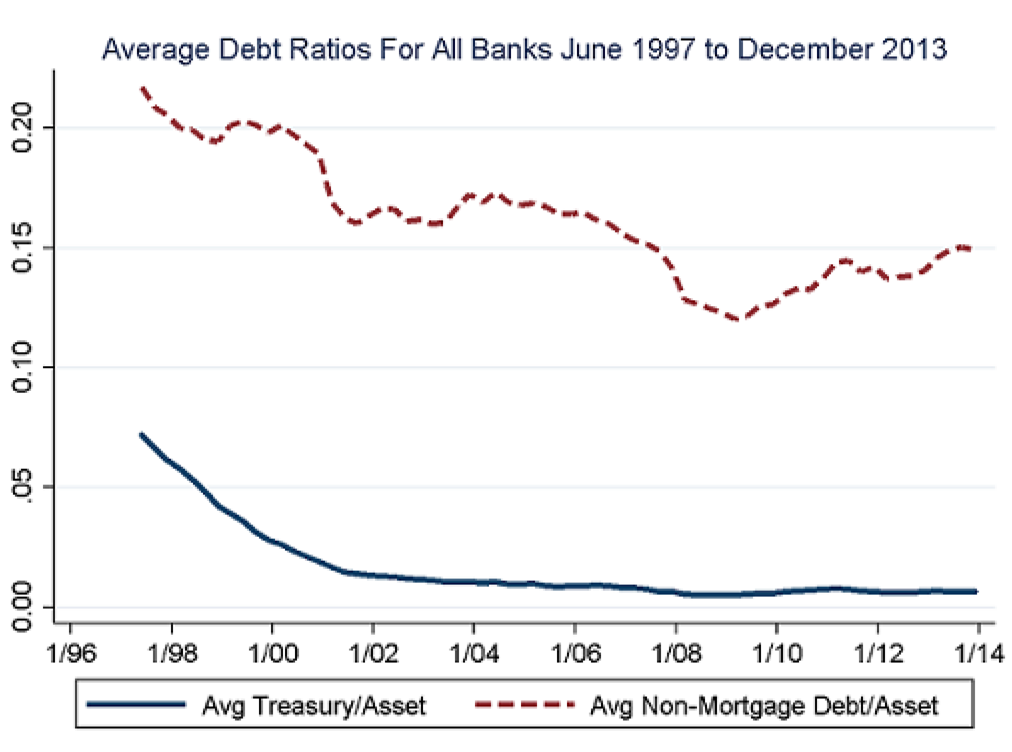


Figure 3.8: Bank Ratio of Treasuries to Assets

Unweighted average Treasury to asset ratio and non-mortgage debt to asset ratio from the call reports. The data are quarterly between June 1975 and December 2013 and includes banks that do not hold MBS.

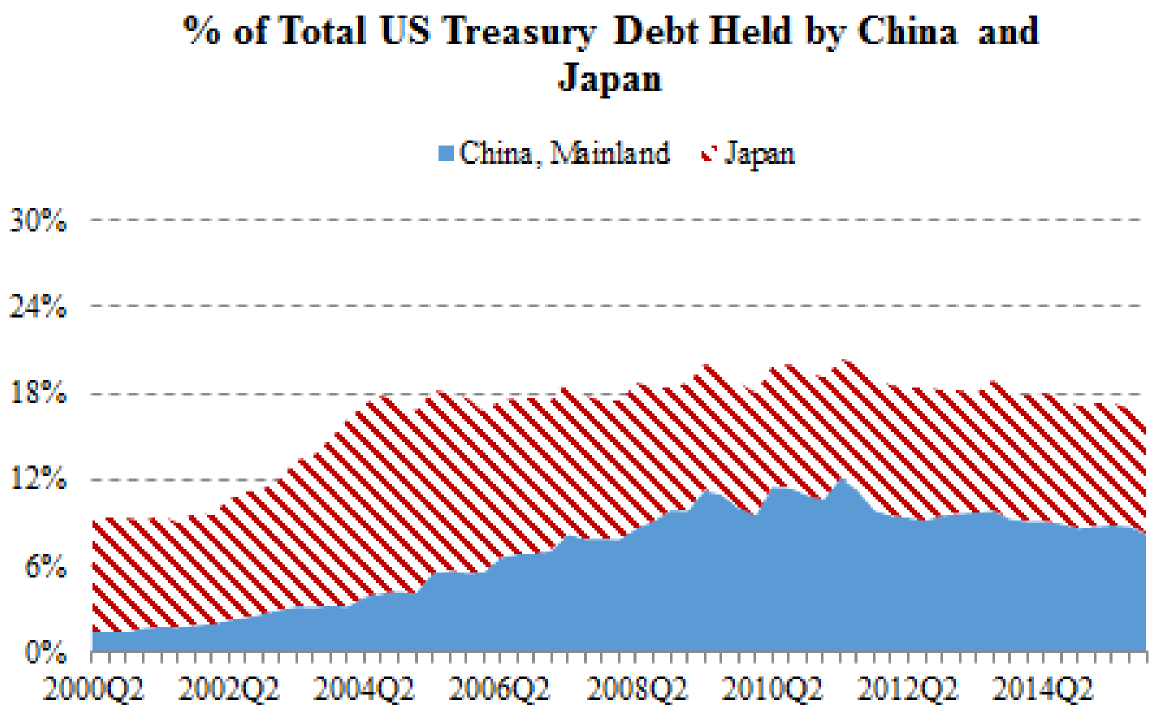


Figure 3.9: Share of US Treasury Debt Held by China and Japan

The percentage of total outstanding US Treasury debt (from the Federal Reserve flow of funds data) held by mainland China and Japan as reported in the Treasury International Capital System (TIC) data. Data are quarterly between Q2 2000 and Q1 2016.

Table 3.1: Effect of MBS Duration on Treasury Excess Bond Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Barclays $Duration_t$	2.577***		2.589					
	(3.01)		(1.66)					
Empirical $Duration_t$		1.723**		1.186				
		(2.43)		(1.87)				
Dates in Aug98-Jan99			-13.61***	-16.44***			-11.54***	-16.80***
			(-4.37)	(-9.82)			(-3.61)	(-11.01)
Dates in Aug02-Jun03			2.755	-2.045			4.688	-2.888
			(0.64)	(-1.14)			(1.15)	(-1.87)
QE1				1.474**				1.242**
				(2.44)				(2.17)
QE2				2.203***				2.233***
				(10.43)				(10.43)
QE3				-0.297**				-0.239
				(-2.11)				(-1.43)
Tapering				0.101				0.189**
				(1.21)				(2.04)
Operation Twist				-0.299				-0.356
				(-1.52)				(-1.89)
$\frac{BarDur_t*MBSMV_t}{AggDur_t*AggMV_t}$					55.91***		56.42**	
					(3.66)		(2.33)	
$\frac{EmpDur_t*MBSMV_t}{AggDur_t*AggMV_t}$						22.08		12.42
						(1.96)		(1.54)
Constant	-3.684	-0.222	-4.219	1.576	-8.121**	0.223	-8.840	2.502
	(-1.07)	(-0.08)	(-0.70)	(0.61)	(-2.02)	(0.07)	(-1.38)	(1.09)
# Obs	237	237	236	237	237	237	236	237
Date Range	Feb96-Oct15	Feb96-Oct15	Jan89-Aug08	Feb96-Oct15	Feb96-Oct15	Feb96-Oct15	Jan89-Aug08	Feb96-Oct15

t statistics in parentheses
** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of effective MBS duration and MBS interest rate risk relative to aggregate interest rate risk (the ratio of MBS dollar duration to aggregate fixed income dollar duration) on 10 year Treasury 12 month excess bond returns. For MBS duration we use either Barclays effective duration or empirical duration. For aggregate fixed income duration we use the aggregate fixed income duration provided by Barclays. MBS and total fixed income market value (AGG) are provided by Barlcays. The MBS market value only includes GSE and Ginnie Mae MBS. Barclays effective duration begins in January 1989, while the empirical duration series begins in February 1996. Both series have data through October 2015. The model is fitted on monthly data and standard errors are Newey West with 18 lags to account for the overlapping structure of the data. The regression model is: $rx^{(10)}_{t+12m} = \alpha + \beta_1 \times Duration_t + \epsilon^{(10)}_{t+12m}$

Table 3.2: Effect of Mortgage Terminations on Treasury Excess Bond Returns

	(1)	(2)	(3)	(4)	(5)	(6)
% Voluntary Payoff$_t * -100$	3.637**		3.609		4.204	
	(2.19)		(1.44)		(0.68)	
Barclays Duration$_t$		2.452***		2.545*		3.422***
		(2.69)		(1.92)		(3.73)
Dates in Aug98-Jan99			-16.25***	-13.23***		
			(-9.73)	(-5.25)		
Dates in Aug02-Jun03			1.209	3.089		
			(0.32)	(0.88)		
QE1			4.118	5.049**		
			(1.61)	(2.29)		
QE2			14.97***	11.92***		
			(9.01)	(7.43)		
QE3			-1.486	-6.090***		
			(-0.50)	(-2.86)		
Tapering			3.335**	-0.717		
			(2.55)	(-0.30)		
Operation Twist			-1.267	-1.745		
			(-0.41)	(-0.77)		
Constant	7.260***	-4.176	6.936***	-4.507	5.031	-12.96***
	(5.30)	(-1.19)	(4.75)	(-0.91)	(0.92)	(-4.67)
# Obs	316	316	316	316	52	52
Date Range	Jan 92-Apr 18	Jan 92-Apr 18	Jan 92-Apr 18	Jan 92-Apr 18	Jan 14-Apr 18	Jan 14-Apr 18

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of Barclays effective MBS duration and underlying voluntary mortgage terminations on 10 year Treasury 12 month excess bond returns. Voluntary mortgage terminations are calculated from McDash and are based on 30-year fixed rate first lien mortgages, excluding construction loans, that are securitized into GSE or Ginnie Mae MBS. The data are monthly between January 1992 and April 2018 and standard errors are Newey West with 18 lags to account for the overlapping structure of the data. The regression model is: $rx^{(10)}_{t+12m} = \alpha + \beta_1 \times Duration_t + \epsilon^{(10)}_{t+12m}$

Table 3.3: Effect of GSE Hedging Portfolios on Treasury Excess Bond Returns

	(1)	(2)	(3)	(4)	(5)
$BarclaysDur_t$	2.384***			2.880**	
	(2.95)			(2.63)	
$EmpiricalDur_t$		1.744**			1.857**
		(2.12)			(2.39)
$HedgingPortfolio_t$			1.230	3.175	1.815
			(0.87)	(1.79)	(1.33)
Constant	-2.508	0.436	3.499	-9.525	-2.946
	(-0.84)	(0.14)	(1.25)	(-1.41)	(-0.88)
# Obs	63	63	63	63	63

t statistics in parentheses

** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of the GSE hedging portfolios on 10 year Treasury 12 month excess bond returns. MBS duration is also included in columns 1, 2, 4, and 5. For MBS duration we use either Barclays effective duration or empirical duration. The GSE hedging portfolio data are quarterly and were retrieved from the historical FHFA (OFHEO prior to the creation of FHFA) reports to Congress. GSE hedging portfolio data are the sum of financial derivatives holdings of Fannie Mae and Freddie Mac. The data start in Q1 2000 and end in Q3 2015. Standard errors are Newey West with 18 lags to account for the overlapping structure of the data. The regression model is: $rx^{(10)}_{t+12m} = \alpha + \beta_1 \times HedgingPortfolio_t + \beta_2 \times Duration_t + \epsilon^{(10)}_{t+12m}$

Table 3.4: Effect of Change in GSE Hedging Portfolios on Change in Treasury Excess Bond Returns

	(1)	(2)	(3)	(4)	(5)
$\Delta BarclaysDur_{t,t-12m}$	-1.573			-1.665	
	(-1.70)			(-1.85)	
$\Delta EmpiricalDur_{t,t-12m}$		-1.446			-1.452
		(-1.75)			(-1.73)
$\Delta GSEHedgingPort_{t,t-12m}$			-1.144	-1.836	-1.237
			(-0.31)	(-0.48)	(-0.32)
Constant	-0.0806	-0.266	-0.177	-0.00385	-0.220
	(-0.04)	(-0.14)	(-0.10)	(-0.00)	(-0.11)
# Obs	56	56	56	56	56

t statistics in parentheses

** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of the 12 month change in GSE hedging portfolios on the 12 month change in 10 year Treasury 12 month excess bond returns. The 12 month change in MBS duration is also included in columns 1, 2, 4, and 5. For MBS duration we use either Barclays effective duration or empirical duration. The GSE hedging portfolio data are quarterly and were retrieved from the historical FHFA (OFHEO prior to the creation of FHFA) reports to Congress. GSE hedging portfolio data are the sum of financial derivatives holdings of Fannie Mae and Freddie Mac. The data start in Q1 2000 and end in Q4 2014. Standard errors are Newey West with 18 lags to account for the overlapping structure of the data. The regression model is: $\Delta rx^{(10)}_{t+24m,t+12m} = \alpha + \beta_1 \times \Delta GSEHedgingPortfolio_{t,t-12m} + \beta_2 \times \Delta Duration_{t,t-12m} + \epsilon^{(10)}_{t+24m,t+12m}$

Table 3.5: Effect of MBS Duration on Bank Treasury Bond Holdings for Banks With MBS Holdings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$EmpiricalDur_{t-12m}$	-0.000872***	0.000715***	-0.00441***	0.00513***	-0.00136***	-0.00155***	-0.00385***	0.00230***
	(-3.96)	(3.31)	(-21.57)	(22.36)	(-14.71)	(-5.25)	(-15.07)	(9.56)
$\frac{MBSHoldings_{t-12m}}{Assets_{t-12m}}$	-0.0783***	-0.480***	-0.188***	-0.292***	-0.0564***	-0.412***	-0.180***	-0.232***
	(-12.27)	(-29.02)	(-13.31)	(-24.64)	(-8.04)	(-16.89)	(-9.19)	(-13.44)
$\frac{MBSHoldings_{t-12m}}{Assets_{t-12m}}$	-0.000141	-0.0141***	-0.00979***	-0.00429**	0.00314***	0.00264	0.00275	-0.000102
$*EmpiricalDur_{t-12m}$	(-0.15)	(-5.02)	(-3.88)	(-2.41)	(3.43)	(0.60)	(0.79)	(-0.04)
Constant	-0.0277***	0.260***	0.165***	0.0945***	0.00120	0.173***	0.105***	0.0687***
	(-5.19)	(17.53)	(13.97)	(8.89)	(0.21)	(8.99)	(8.48)	(4.35)
# Obs	327440	327440	327440	327440	115350	115350	115350	115350
R-squared	0.572	0.816	0.712	0.672	0.465	0.759	0.680	0.619
Maturity of Dep Variable	Treas	All	$> 3yrs$	$\leq 3yrs$	Treas	All	$> 3yrs$	$\leq 3yrs$
Bank Has Held IR	No IR	No IR	No IR	No IR	Has IR	Has IR	Has IR	Has IR

t statistics in parentheses
** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of 12 month lagged MBS duration on the ratio of debt holdings to total assets at the bank level for banks that **have never** held interest rate derivatives in their non-trading accounts in columns 1-4 and for banks that **have** held interest rate derivatives in their non-trading accounts in columns 5-8. All banks in this sample **have** held MBS in their non-trading accounts. The bank level data are retrieved from the bank call reports and are quarterly from June 1997 to December 2013. For MBS duration we use empirical duration. Bank fixed effects and bank time varying controls are also included. Standard errors are clustered at the bank level (rssd9001). The regression model is: $\frac{Debt_{t,i}}{Assets_{t,i}} = \alpha + \beta_1 \times Duration_{t-12m} + \beta_2 \times \frac{MBSSecurities_{t-12m,i}}{Assets_{t-12m,i}} + \beta_3 \times \frac{MBSSecurities_{t-12m,i}}{Assets_{t-12m,i}} \times Duration_{t-12m} + \mathbb{X}_{i,t-12m} + \epsilon_{t,i}$

Table 3.6: Effect of MBS Duration on Bank Treasury Bond Holdings for Banks Without MBS Holdings

	(1)	(2)	(3)	(4)
$EmpiricalDur_{t-12m}$	-0.000615	0.00122**	-0.00510***	0.00632***
	(-1.70)	(2.55)	(-11.69)	(12.93)
Constant	0.0268	0.282***	0.130***	0.152***
	(1.40)	(9.83)	(7.66)	(6.77)
# Obs	47227	47225	47227	47225
R-squared	0.724	0.875	0.743	0.759
Maturity of Dep Variable	Treas	All	$> 3yrs$	$\leq 3yrs$

t statistics in parentheses
** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of 12 month lagged MBS duration on the ratio of debt holdings to total assets at the bank level for banks that **have never** held MBS in their non-trading accounts. The bank level data are retrieved from the bank call reports and are quarterly from June 1997 to December 2013. For MBS duration we use empirical duration. Bank fixed effects and bank time varying controls are also included. Standard errors are clustered at the bank level (rssd9001). The regression model is: $\frac{Debt_{t,i}}{Assets_{t,i}} = \alpha + \beta_1 \times Duration_{t-12m} + \mathbb{X}_{i,t-12m} + \epsilon_{t,i}$

Table 3.7: Effect of Change in Barclays Effective MBS Duration on Change in Foreign Treasury Bond Holdings

	(1)	(2)	(3)	(4)
$\Delta BarclaysDuration_{t-12m,t}$	-0.0142	-0.00844	-0.0198	-0.0207
	(-0.66)	(-0.40)	(-0.75)	(-0.74)
$\Delta FX = 1USD_{t-12m,t}$		-0.00461**		-0.0977
		(-2.17)		(-0.35)
Constant	0.106**	0.105***	0.243***	0.228***
	(2.60)	(2.73)	(4.86)	(3.58)
# Obs	175	175	175	175
Country	Japan	Japan	China	China

t statistics in parentheses

** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of 12 month changes in MBS Barclays effective duration on the 12 month percentage change in US Treasury holdings of mainland China and Japan. Change in FX rates are also included as a control in some specifications. Country level US Treasury holding data are from the Treasury International Capital (TIC) dataset. The data are monthly from June 2000 to December 2014. Standard errors are Newey West with 18 lags to account for the overlapping structure of the data. The regression model is: $\frac{\Delta TSYHoldings_{t,t+12m}}{TSYHoldings_t} = \alpha + \beta_1 \times \Delta Duration_{t-12m,t} + \beta_2 \times \Delta FXRate_{t-12m,t} + \epsilon_{t,t+12m}$

Table 3.8: Effect of Change in Empirical MBS Duration on Change in Foreign Treasury Bond Holdings

	(1)	(2)	(3)	(4)
$\Delta EmpiricalDuration_{t-12m,t}$	-0.0108	-0.00429	0.00238	0.00176
	(-1.08)	(-0.42)	(0.11)	(0.08)
$\Delta FX = 1USD_{t-12m,t}$		-0.00462**		-0.0881
		(-2.11)		(-0.31)
Constant	0.105***	0.104***	0.241***	0.229***
	(2.62)	(2.76)	(4.57)	(3.47)
# Obs	175	175	175	175
Country	Japan	Japan	China	China

t statistics in parentheses

** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of 12 month changes in MBS empirical duration on the 12 month percentage change in US Treasury holdings of mainland China and Japan. Change in FX rates are also included as a control in some specifications. Country level US Treasury holding data are from the Treasury International Capital (TIC) dataset. The data are monthly from June 2000 to December 2014. Standard errors are Newey West with 18 lags to account for the overlapping structure of the data. The regression model is: $\frac{\Delta TSYHoldings_{t,t+12m}}{TSYHoldings_t} = \alpha + \beta_1 \times \Delta Duration_{t-12m,t} + \beta_2 \times \Delta FXRate_{t-12m,t} + \epsilon_{t,t+12m}$

Table 3.9: Effect of MBS Duration on Mutual Fund Treasury Bond Holdings

	(1)	(2)	(3)	(4)
$BarclaysDur_{t-12m}$	0.143^{***}	0.313^{***}		
	(12.23)	(12.31)		
$EmpiricalDur_{t-12m}$			0.0556^{***}	0.118^{***}
			(7.18)	(7.16)
Constant	7.641^{***}	15.56^{***}	8.056^{***}	16.49^{***}
	(168.70)	(157.23)	(416.20)	(399.54)
# Obs	724356	250620	724356	250620
R-squared	0.934	0.910	0.934	0.910
Only Funds with MBS Positions?	N	Y	N	Y
Duration Measure	Barclays	Barclays	Empirical	Empirical

t statistics in parentheses

** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of MBS duration on the percentage of assets held in Treasury bonds by mutual funds. Mutual fund data are from CRSP and are quarterly from Q1 2010 to Q2 2016. MBS duration is either Barclays effective duration or empirical duration. Standard errors are clustered at the mutual fund level. The regression model is: $PerGovtBonds_{t,i} = \alpha + \beta_1 \times Duration_{t-j} + \mathbb{X}_i + \epsilon_{t,i}$

Table 3.10: Effect of Change in MBS Duration on Change in Mutual Fund Treasury Bond Holdings

	(1)	(2)	(3)	(4)
$\Delta BarclaysDur_{t-12,t}$	0.0664***	0.206***		
	(8.25)	(11.59)		
$\Delta EmpiricalDur_{t-12,t}$			0.0385***	0.115***
			(6.01)	(8.08)
Constant	0.149***	0.215***	0.165***	0.264***
	(67.13)	(44.42)	(400.30)	(293.21)
# Obs	571324	199230	571324	199230
R-squared	0.140	0.128	0.140	0.127
Only Funds with MBS Positions?	N	Y	N	Y
Duration Measure	Barclays	Barclays	Empirical	Empirical

t statistics in parentheses

** $p < 0.05$, *** $p < 0.01$

Regressions measuring the effect of the change in MBS duration on the change in the percentage of assets held in Treasury bonds by mutual funds. Mutual fund data are from CRSP and are quarterly from Q1 2010 to Q2 2015. MBS duration is either Barclays effective duration or empirical duration. Standard errors are clustered at the mutual fund level. The regression model is: $\Delta PerGovtBonds_{t\to t+j,i} = \alpha + \beta_1 \times \Delta Duration_{t-j\to t} + \mathbb{X}_i + \epsilon_{t\to t+j,i}$

Table 3.11: Effect of Barclays Effective MBS Duration on Life Insurance Company Treasury Bond Purchases

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$BarclaysDuration_{t-1m}$	-0.00510***			-0.00541**	-0.00837***	-0.00270***	-0.00579***	-0.00515***	-0.00562***
	(-4.36)			(-2.20)	(-3.41)	(-2.70)	(-6.59)	(-3.48)	(-3.70)
$BarclaysDuration_{t-6m}$		-0.00520***							
		(-4.34)							
$BarclaysDuration_{t-12m}$			-0.00912***						
			(-8.05)						
Constant	0.131***	0.132***	0.144***	0.130***	0.242***	0.0814***	0.0747***	0.131***	0.133***
	(37.27)	(36.13)	(40.98)	(17.51)	(32.50)	(27.08)	(28.23)	(30.23)	(29.84)
# Obs	119402	119402	119402	27741	49917	119402	119402	76988	78867
R-squared	0.154	0.154	0.155	0.0191	0.105	0.116	0.125	0.168	0.169
Years	All	All	All	All	All	All	All	2001-2007	2001-2007
Num Mo With TSY Trade	All	All	All	≤ 12	>12	All	All	All	All
File	300	300	300	300	300	300	300	300	303/304
TSY Maturity	All	All	All	All	All	STN	LTD	LTD	All

t statistics in parentheses
** $p < 0.05$, *** $p < 0.01$

Regressions estimate a linear probability model of the probability of life insurance firms buying Treasuries due to lagged MBS duration. Life insurance holding data are from NAICS and are monthly from 2001 to 2012. MBS duration is Barclays effective duration. Firm fixed effects are also included. Standard errors are clustered at the firm level. Models with 1, 6, and 12 month lags on the MBS duration variable are estimated. The regression model is: $\mathbb{1}_{t,i} = \alpha + \beta_1 \times Duration_{t-j} + \mathbb{X}_i + \epsilon_{t,i}$

Table 3.12: Effect of Empirical MBS Duration on Life Insurance Company Treasury Bond Purchases

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$EmpiricalDuration_{t-1m}$	-0.00574***			-0.00603***	-0.0116***	-0.00380***	-0.00562***	-0.00518***	-0.00534***
	(-5.35)			(-3.11)	(-5.17)	(-4.46)	(-7.09)	(-3.92)	(-3.95)
$EmpiricalDuration_{t-6m}$		-0.00215**							
		(-2.09)							
$EmpiricalDuration_{t-12m}$			-0.00735***						
			(-7.28)						
Constant	0.132***	0.122***	0.138***	0.131***	0.250***	0.0843***	0.0735***	0.132***	0.133***
	(42.82)	(40.21)	(44.87)	(23.52)	(38.80)	(34.32)	(32.14)	(31.98)	(31.40)
# Obs	119402	119402	119402	27741	49917	119402	119402	76988	78867
R-squared	0.154	0.154	0.155	0.0193	0.106	0.116	0.125	0.168	0.169
Years	All	All	All	All	All	All	All	2001-2007	2001-2007
Num Mo With TSY Trade	All	All	All	≤ 12	> 12	All	All	All	All
File	300	300	300	300	300	300	300	300	303/304
TSY Maturity	All	All	All	All	All	STN	LTD	LTD	All

t statistics in parentheses
** $p < 0.05$, *** $p < 0.01$

Regressions estimate a linear probability model of the probability of life insurance firms buying Treasuries due to lagged MBS duration. Life insurance holding data are from NAICS and are monthly from 2001 to 2012. MBS duration is empirical duration. Firm fixed effects are also included. Standard errors are clustered at the firm level. Models with 1, 6, and 12 month lags on the MBS duration variable are estimated. The regression model is: $\mathbb{1}_{t,i} = \alpha + \beta_1 \times Duration_{t-j} + \mathbb{X}_i + \epsilon_{t,i}$

Bibliography

Adao, Rodrigo, Michal Kolesár, and Eduardo Morales, 2019, Shift-share designs: Theory and inference, *The Quarterly Journal of Economics* 134, 1949–2010.

Adelino, M., A. Schoar, and F. Severino, 2015, House prices, collateral, and self-employment, *Journal of Financial Economics* 117, 288–306.

Akepanidtaworn, Klakow, Rick Di Mascio, Alex Imas, and Lawrence Schmidt, 2018, Selling fast and buying slow: Heuristics and trading performance of institutional investors, *Working Paper* .

Angrist, D, 1998, Estimating the labor market impact of voluntary military service using social security data on military applicants, *Econometrica* .

Bahaj, S.A., A. Foulis, and G. Pinter, 2017, Home values and firm behaviour, *Working Paper* .

Ballou, J., T. Barton, D. DesRoches, F. Potter, Z. Zhao, B. Santos, and J. Sebastian, 2007, Kauffman firm survey (kfs) baseline methodology report, *Working Paper* .

Barber, Brad M., and Terrance Odean, 2001, Boys will be boys: Gender, overconfidence, and common stock investment., *The quarterly journal of economics* 116, 261–292.

Barnatchez, K., L.D. Crane, and R. Decker, 2017, An assessment of the national establishment time series (nets) databases, *Working Paper* .

Bartik, T.J., 1991, Who benefits from state and local economic development policies?, *WE Upjohn Institute for Employment Research* .

Benmelech, Meisenzahl R.R., E., and R. Ramcharan, 2017, The real effects of liquidity during the financial crisis: Evidence from automobiles, *The Quarterly Journal of Economics* 132, 317–365.

Berends, Kyal R., Robert McMenamin, Thanases Plestis, and Richard J. Rosen, 2013, The sensitivity of life insurance firms to interest rate changes, *Economic Perspectives* 37.

Berger, A.N., and G.F. Udell, 1995, Relationship lending and lines of credit in small firm finance, *Journal of Business* 351–381.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Bernstein, S., T. McQuade, and R.R. Townsend, 2018, Do household wealth shocks affect productivity? evidence from innovative workers during the great recession, *NBER Working Paper (No. 24011)* .

Bhutta, N., and B.J. Keys, 2016, Interest rates and equity extraction during the housing boom, *American Economic Review* 106, 1742–74.

Blanchard, O.J., L.F. Katz, R.E. Hall, and B. Eichengreen, 1992, Regional evolutions, *Brookings Papers on Economic Activity* 1992, 1–75.

Blanchflower, D.G., and A.J. Oswald, 1998, What makes an entrepreneur?, *Journal of Labor Economics* 16, 26–60.

Bord, V.M., V. Ivashina, and R.D. Taliaferro, 2018, Large banks and small firm lending, *National Bureau of Economic Research (No. w25184)* .

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, 2018, Quasi-experimental shift-share research designs, *NBER Working Paper (No. w24997)* .

Boudoukh, Jacob, Matthew Richardson, Richard Stanton, and Robert F. Whitelaw, 1995, A new strategy for dynamically hedging mortgage-backed securities, *Journal of Derivatives* 2, 60–77.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru, 2018, Fintech, regulatory arbitrage, and the rise of shadow banks, *Journal of Financial Economics* 130, 453–483.

Busse, Jeffrey A., Tarun Chordia, Lei Jiang, and Yuehua Tang, 2016, Mutual fund transaction costs, *Working Paper* .

Campbell, J.Y., 2006, Household finance, *The Journal of Finance* 61, 1553–1604.

Card, D., and D.G. Sullivan, 1988, Measuring the effect of subsidized training programs on movements in and out of employment, *Econometrica* .

Chen, Qi, Itay Goldstein, and Wei Jiang, 2008, Directors' ownership in the us mutual fund industry, *The Journal of Finance* 63, 2629–2677.

Chernenko, Sergey, Samuel G. Hanson, and Adi Sunderam, 2016, Who neglects risk? investor experience and the credit boom, *Journal of Financial Economics* 122, 248–269.

Chodorow-Reich, G., 2013, The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis, *The Quarterly Journal of Economics* 129, 1–59.

Cochrane, John H., and Monika Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 138–160.

Corradin, S., and A. Popov, 2015, House prices, home equity borrowing, and entrepreneurship, *The Review of Financial Studies* 28, 2399–2428.

Coval, Joshua D., and Tobias J. Moskowitz, 1999, Home bias at home: Local equity preference in domestic portfolios, *The Journal of Finance* 54, 2045–2073.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.

Cremers, KJ Martijn, and Antti Petajisto, 2009, How active is your fund manager? a new measure that predicts performance, *The Review of Financial Studies* 22, 3329–3365.

Cremers, Martijn, Joost Driessen, Pascal Maenhout, and David Weinbaum, 2009, Does skin in the game matter? director incentives and governance in the mutual fund industry, *Journal of Financial and Quantitative Analysis* 44, 1345–1373.

Cremers, Martijn, and Ankur Pareek, 2016, Patient capital outperformance: The investment skill of high active share managers who trade infrequently, *Journal of Financial Economics* 122, 288–306.

Cronqvist, Henrik, Anil K. Makhija, and Scott E. Yonker, 2012, Behavioral consistency in corporate finance: Ceo personal and corporate leverage, *Journal of Financial Economics* 103, 20–40.

Culbertson, J. M., 1957, The term structure of interest rates, *Quarterly Journal of Economics* 71, 485–517.

Davis, S.J., J. Haltiwanger, and S. Schuh, 1996, Small business and job creation: Dissecting the myth and reassessing the facts, *Small Business Economics* 8, 297–315.

Davis, S.J., and J.C. Haltiwanger, 2019, Dynamism diminished: The role of housing markets and credit conditions, *National Bureau of Economic Research (No. w25466)* .

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda, 2014, The role of entrepreneurship in us job creation and economic dynamism, *Journal of Economic Perspectives* 28, 3–24.

DeFusco, A., S. Johnson, and J. Mondragon, 2017, Regulating household leverage, *Working Paper* .

DeFusco, A.A., 2018, Homeowner borrowing and housing collateral: New evidence from expiring price controls, *The Journal of Finance* 73, 523–573.

Evans, Allison L, 2008, Portfolio manager ownership and mutual fund performance, *Financial Management* 37, 513–534.

Evans, D.S., and B. Jovanovic, 1989, An estimated model of entrepreneurial choice under liquidity constraints, *Journal of Political Economy* 97, 808–827.

Fisher, Lawrence, and Roman. L. Weil, 1971, Coping with the risk of interest rate fluctuations: Returns to bondholders from naive and optimal strategies, *Journal of Business* 44, 111–118.

Gagnon, Joseph E., Matthew Raskin, Julie Ann Remache, and Brian Sack, 2011, The financial market effects of the Federal Reserve's large-scale asset purchases, *International Journal of Central Banking* 7, 3–43.

Gete, P., and M. Reher, 2018, Mortgage supply and housing rents, *The Review of Financial Studies* 31, 4884–4911.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, 2018, Bartik instruments: What, when, why, and how, *NBER Working Paper (No. w24408)* .

Goodman, L.S., 2017, Quantifying the tightness of mortgage credit and assessing policy actions, *BCJL and Soc. Just.* 37, 235.

Gourio, F., T. Messer, and M. Siemer, 2016, Firm entry and macroeconomic dynamics: a state-level analysis, *American Economic Review* 106, 214–18.

Greenspan, Alan, and James Kennedy, 2008, Sources and uses of equity extracted from homes, *Oxford Review of Economic Policy* 24, 120–144.

Greenstone, M., A. Mas, and H.L. Nguyen, 2014, Do credit market shocks affect the real economy? quasi-experimental evidence from the great recession and 'normal'economic times, *National Bureau of Economic Research (No. w20704)* .

Greenwood, Robin, and Stefan Nagel, 2009, Inexperienced investors and bubbles, *Journal of Financial Economics* 93, 239–258.

Greenwood, Robin, and Dimitri Vayanos, 2010, Price pressure in the government bond market, *American Economic Review* 100, 585–590.

Greenwood, Robin, and Dimitri Vayanos, 2014, Bond supply and excess bond returns, *Review of Financial Studies* 27, 663–713.

Grinblatt, Mark, and Matti Keloharju, 2009, Sensation seeking, overconfidence, and trading activity., *The Journal of Finance* 64, 549–578.

Gupta, Arpit, and Kunal Sachdeva, 2017, Skin or skim? inside investment and hedge fund performance, *Working Paper* .

Guren, A.M., A. McKay, E. Nakamura, and J. Steinsson, 2018, Housing wealth effects: The long view, *National Bureau of Economic Research (No. w24729)* .

Haltiwanger, J., R.S. Jarmin, R.B. Kulick, and J. Miranda, 2016, High growth young firms: Contribution to job, output and productivity growth, *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-49* .

Haltiwanger, J., R.S. Jarmin, and J. Miranda, 2013, Who creates jobs? small versus large versus young, *Review of Economics and Statistics* 95, 347–361.

Hanson, Samuel G., 2014, Mortgage convexity, *Journal of Financial Economics* 113, 270–299.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2016, . . . and the cross-section of expected returns, *The Review of Financial Studies* 29, 5–68.

Hathaway, I., and R.E. Litan, 2014, What's driving the decline in the firm formation rate? a partial explanation, *The Brookings Institution* .

Hayre, Lakhbir, ed., 2001, *Salomon Smith Barney Guide to Mortgage-Backed and Asset-Backed Securities* (John Wiley, New York).

Hicks, John R., 1939, *Value and Capital* (Clarendon Press, Oxford, UK).

Ho, Thomas S. Y., 1992, Key rate durations: Measures of interest rate risks, *Journal of Fixed Income* 2, 29–44.

Hong, Harrison, and Leonard Kostovetsky, 2012, Red and blue investing: Values and finance, *Journal of Financial Economics* 103, 1–19.

Hurst, E., G. Li, and B. Pugsley, 2014, Are household surveys like tax forms? evidence from income underreporting of the self-employed, *Review of Economics and Statistics* 96, 19–33.

Hurst, E., and A. Lusardi, 2004, Liquidity constraints, household wealth, and entrepreneurship, *Journal of Political Economy* 112, 319–347.

Iacus, S.M., G. King, and G. Porro, 2019, A theory of statistical inference for matching methods in causal research, *Political Analysis* 27, 46–68.

Ivković, Zoran, and Scott Weisbenner, 2005, Local does as local is: Information content of the geography of individual investors' common stock investments, *Journal of Finance* 60, 267–306.

Jarmin, R.S., and J. Miranda, 2002, The longitudinal business database, *Working Paper* .

Karahan, F., B. Pugsley, and A. Şahin, 2019, Demographic origins of the startup deficit, *National Bureau of Economic Research (No. w25874)* .

Kerr, S., W.R. Kerr, and R. Nanda, 2015, House money and entrepreneurship, *National Bureau of Economic Research (No. w21458)* .

Khorana, Ajay, Henri Servaes, and Lei Wedge, 2007, Portfolio manager ownership and fund performance, *Journal of Financial Economics* 85, 179–204.

Kiyotaki, N., and J. Moore, 1997, Credit cycles, *Journal of Political Economy* 105, 211–248.

Krishnamurthy, Arvind, and Annette Vissing-Jørgensen, 2011, The effects of quantitative easing on interest rates: Channels and implications for policy, *Brookings Papers on Economic Activity* 215–287.

Krishnan, Karthik, Debarshi K Nandy, and Manju Puri, 2014, Does financing spur small business productivity? evidence from a natural experiment, *The Review of Financial Studies* 28, 1768–1809.

Kumar, Alok, 2009, Who gambles in the stock market?, *Journal of Finance* 64, 1889–1933.

Laufer, S., and A. Paciorek, 2018, The effects of mortgage credit availability: Evidence from minimum credit score lending rules., *Working Paper* .

Lee, Mayer C.J., D., and J. Tracy, 2012, A new look at second liens, *National Bureau of Economic Research (No. w18269)* .

Liu, Crocker, and David Yermack, 2012, Where are the shareholders' mansions? ceos' home purchases, stock sales, and subsequent company performance, *Corporate Governance* 3–28.

Loutskina, E., and P.E. Strahan, 2015, Financial integration, housing, and economic volatility., *Journal of Financial Economics* 115, 25–41.

Lu, Yan, and Melvyn Teo, 2018, Do alpha males deliver alpha? testosterone and hedge funds, *Working Paper* .

Macaulay, Frederick R., 1938, Some theoretical problems suggested by the movements of interest rates, bond yields, and stock prices in the United States since 1856, Working paper, National Bureau of Economic Research.

Malkhozov, Aytek, Philippe Mueller, Andrea Vedolin, and Gyuri Venter, 2016, Mortgage risk and the yield curve, *Review of Financial Studies* 29, 1220–1253.

Mian, A., and A. Sufi, 2014, What explains the 2007–2009 drop in employment?, *Econometrica* 82, 2197–2223.

Modigliani, Franco, and Richard Sutch, 1966, Innovations in interest rate policy, *American Economic Review* 56, 178–197.

Mondragon, J., 2018, Household credit and employment in the great recession, *Kilts Center for Marketing at Chicago Booth–Nielsen Dataset Paper Series* 1–025.

Neumark, D., J. Zhang, and B. Wall, 2007, Employment dynamics and business relocation: New evidence from the national establishment time series, *Aspects of Worker Well-Being* 39–83.

Nguyen, H.L.Q., 2019, Are credit markets still local? evidence from bank branch closings, *American Economic Journal: Applied Economics* 11, 1–32.

Odean, T., and B. Barber, 1999, The courage of misguided convictions: the trading behavior of individual investors, *Financial Analyst Journal* 41–55.

Odean, Terrance, 1999, Do investors trade too much?, *American Economic Review* 89, 262.

Palmer, C., 2015, Why did so many subprime borrowers default during the crisis: Loose credit or plummeting prices?, *Working Paper* .

Patnaik, M., 2017, The impact of credit shocks: Micro versus small firms, *Working Paper* .

Petersen, M.A., and R.G. Rajan, 1994, The benefits of lending relationships: Evidence from small business data, *The Journal of Finance* 49, 3–37.

Pool, Veronika K., Noah Stoffman, and Scott E. Yonker, 2012, No place like home: Familiarity in mutual fund manager portfolio choice, *The Review of Financial Studies* 25, 2563–2599.

Pool, Veronika K., Noah Stoffman, and Scott E. Yonker, 2015, The people in your neighborhood: Social interactions and mutual fund portfolios, *The Journal of Finance* 70, 2679–2732.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2017, Do funds make more when they trade more?, *The Journal of Finance* 72, 1483–1528.

Pugsley, B.W., and A. Șahin, 2018, Grown-up business cycles, *The Review of Financial Studies* 32, 1102–1147.

Risa, Sefano, Rodrigo Ibanez-Meier, Min Fan, and Idriss Maoui, 2008, Prime fixed-rate prepayment model update, Research report, Barclays Capital.

Robb, A.M., and D.T. Robinson, 2014, The capital structure decisions of new firms, *The Review of Financial Studies* 27, 153–179.

Rosenbaum, P.R., and D.B. Rubin, 1983, The central role of the propensity score in observational studies for causal effects, *Biometrika* 70, 41–55.

Ross, Myron H., 1966, “Operation Twist”: A mistaken policy?, *Journal of Political Economy* 74, 195–199.

Rubin, D.B., 1973, Matching to remove bias in observational studies, *Biometrics* 159–183.

Rubin, D.B., 1974, Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of educational Psychology* 66, 688.

Rubin, D.B., 1977, Assignment to treatment group on the basis of a covariate, *Journal of Educational Statistics* 2, 1–26.

Sarsons, H., 2017, Interpreting signals in the labor market: evidence from medical referrals, *Job Market Paper* .

Schmalz, M.C., D.A. Sraer, and D. Thesmar, 2017, Housing collateral and entrepreneurship, *The Journal of Finance* 72, 99–132.

Siemer, M., 2016, Firm entry and employment dynamics in the great recession, *Available at SSRN 2172594* .

Song, Yang, 2019, The mismatch between mutual fund scale and skill, *Working Paper* .

Srinivasan, V. S., and Kumar Velayudham, 2010, Prepayment model update, Research report, Barclays Capital.

Stiglitz, J.E., and A. Weiss, 1981, Credit rationing in markets with imperfect information, *The American Economic Review* 71, 393–410.

Stoffman, Noah, Veronika K Pool, Scott E Yonker, and Hanjiang Zhang, 2018, Do shocks to personal wealth affect risk-taking in delegated portfolios?, *The Review of Financial Studies* 32, 1457–1493.

Törnqvist, L., P. Vartia, and Y.O. Vartia, 1985, How should relative changes be measured?, *The American Statistician* 39, 43–46.

Vasicek, Oldrich A., 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177–188.

Vayanos, Dimitri, and Jean-Luc Vila, 2009, A preferred-habitat model of the term structure of interest rates, Working Paper 15487, NBER.

Wallace, Neil, 1967, The term structure of interest rates and the maturity composition of the Federal debt, *Journal of Finance* 22, 301–312.

Appendix A

Appendix to Chapter 1

A.1 Merges

A.1.1 ATTOM - NETS (D&B) Merge

The micro-data utilized by this paper are a unique merge between four different datasets. It is not a trivial task to link home equity extraction activity of business owners to the tax records of their business, which necessitated a novel approach. ATTOM is a comprehensive property and transaction level dataset on residential (and commercial) real estate purchases and refinances. From this data, prior research has constructed the amount of home equity extracted for refinancing transactions (DeFusco, 2018)—this is described in Section 2 of the Online Appendix. However, linking ATTOM to information on businesses owned by homeowners is difficult, since ATTOM does not list employment information. Fortunately, ATTOM lists the name of the person/people who are the legal owners of the property at the time of each transaction. With this information, ATTOM can be merged to NETS, which for most businesses lists the names of business owners. This merge provides a dataset of the home equity extraction activity of homeowners linked to the businesses that they own. For confidentiality reasons, names of people and businesses are not utilized in the research and information on individual records are not reported.

To merge ATTOM to NETS, the name of the homeowners in ATTOM are merged to the name of firm owners in NETS. Roughly 60% of firms in NETS have the name of the firm owner listed. In both datasets, the names are first cleaned and standardized. Punctuation, prefixes, and suffixes are removed and 429 common names are mapped to a standard spelling (e.g. Anne and Annie are both mapped to Ann). ATTOM contains names for up to two people for each transaction. If two names are listed, both names are included for the merge to NETS. For some transactions, ATTOM lists the name of a trust or business (usually a bank for cases of foreclosure transactions, which are not included in the merge). In the case of a trust, the cleaning algorithm attempts to extract a name listed with the trust (i.e. John Smith for "The John Smith Family Trust").

The merge is attempted in six passes, with subsequent passes based on looser criteria.

For each pass, merged Dunsnumbers (identifier for establishments) in NETS are counted as successfully merged if only one record from ATTOM is merged. In cases where more than one record in ATTOM is merged to a business in NETS, the recorded merge is dropped. Any records merged in an earlier pass, even if the merge was not unique and consequently dropped, are not included in later passes. In the first pass, the records are merged on full first and last names. In cases where both ATTOM and NETS contain a middle initial or name, the middle initial is also included in the merge. For the second pass, records are matched on first and middle initials and full last name. Records in ATTOM and NETS must have a known middle initial to be included in this pass.

NETS allows for establishments of multi-unit firms to be linked to their headquarters. In some cases, the name of the firm owner reported for the headquarter is different from the name of the owner of an establishment. In the third pass, the name of the headquarter owner is used to merge records on full first name, middle initial (if both ATTOM and NETS list a middle initial/name), and full last name. The fourth pass also uses the headquarter owner name and merges on first name initial, middle name initial, and full last name. For the fourth pass, both ATTOM and NETS must contain information on middle name. These two passes replicate passes one and two with headquarter owner name instead of establishment owner name. Most establishments are also the headquarters, since most businesses are single-unit entities.

In the fifth pass, records are merged on full last name and the latitude and longitude of the home in ATTOM and the firm in NETS. This pass primarily includes home businesses. For the analysis of credit constraints, most of these records would be excluded as these firms are located in the same zip code of the firm owner's home. Latitude and longitude are rounded to the third decimal place, which represents 110 meters of accuracy.

For the sixth pass, records are merged on the first three letters of first name and full last name. Middle initials are also matched for cases where both NETS and ATTOM contain middle initial. Additional passes involving fuzzy merges were attempted but hand screening these merges showed that this resulted in an increase in the rate of false merges.

A.1.2 NETS (D&B) - SSEL Merge

The restricted use Longitudinal Business Database (LBD) contains a longitudinal panel of data on an establishment's survival, employment and payroll. The data include every employer establishment in the United States and are constructed from IRS tax records of the business. The LBDNUM variable tracks an establishment across time in the LBD. The LBD does not contain business name and address, which are necessary to merge the data to NETS. The Business Registrar (SSEL) is the raw underlying data that the LBD data are constructed from. The SSEL contains business name and address, which necessitates first having to merge NETS to SSEL prior to the merging NETS to LBD.

Both NETS and SSEL contain business name, address, and industry. NETS contains a firm's first and last year, while the SSEL are annual files that contain all employer establishments open in the year of the file (without the ability to track establishments across the different annual files). For each annual SSEL file, all open establishments in NETS, based on the firstyear (year business opens) and lastyear (year business closes) variables in NETS relative to the SSEL file year, are kept for the merge to SSEL. To allow for reporting errors and differences, the value for firstyear in NETS is allowed to be up to five years after the SSEL file year, and the value for lastyear in NETS is allowed to be up to five years prior to the SSEL file year. One reason for reporting differences is that the first year in NETS is the first year the business is in operation, even if it starts as a non-employer firm. In the SSEL, a firm only appears when the firm has at least one paid employee.

To merge NETS to SSEL, five passes are attempted. Prior to merging, company name and street address are cleaned to normalize the strings and COMPGED and SPEDIS functions in SAS are used to calculate distance scores on strings (this is the basis of the fuzzy matches). In the first pass, zip codes and street numbers are exact matched and street name, unit number, and company name are fuzzy merged. Within this first pass, five separate merges with varying degrees of tightness for the fuzzy part of the match are attempted. Matching on the first five characters of company name and the NAICS sector of the business (except for sector 54 [professional offices], which contain many businesses with the same initial string in the company name) is also attempted.

In the second through fifth passes, the maximum scores from the COMPGED and SPEDIS functions are lowered (the criteria are tightened) as criteria for other attributes of the merge are loosened. In the second pass, zip code and 2-digit NAICS sector are exact matched and fuzzy matches on street name, unit number, and company name are used. Street number and unit number are ignored in this pass. For the third pass, zip code and 4-digit NAICS category are exact matched and a fuzzy match on company name is attempted. Address information beyond the zip code is not used for this pass. In the fourth pass, county, state, and 6-digit NAICS category are exact matched and a fuzzy match on company name is attempted. In the fifth and final pass, county, state, and 4-digit NAICS are exact matched and a very tight fuzzy merge on company name is attempted. Most merges are from the first pass, with each subsequent pass producing a diminishing number of merges. Due to disclosure reasons, statistics on the merge rate from individual passes are not available.

A.1.3 LBD - SSEL Merge

The SSEL does not contain an identifier that links businesses through time. Once the SSEL is merged to NETS, each record in the SSEL that is merged will have the associated Dunsnumber from NETS. The Dunsnumber allows for longitudinal tracking of establishments in the SSEL. For each Dunsnumber, the modal LBDNUM in the LBD is selected. If there is more than one modal LBDNUM, one is randomly selected. If the modal LBDNUM is mapped to more than one Dunsnumber, the one that minimizes the difference in the establishment's last year between NETS and LBD is selected. If this a tie, the one with the closest SIC industry codes between NETS and LBD is used as a tiebreaker. If a third tiebreaker is needed, the one that minimizes the difference between the establishment's first year in NETS and LBD is selected. Due to disclosure reasons, statistics on the merge rate are not available.

A.2 Home Equity Extraction Data from ATTOM

ATTOM contains a record for every real estate purchase and refinance transaction, from which the amount of home equity extracted for each refinance transaction can be constructed. This paper borrows from the approach used by DeFusco (2018). The primary differences from DeFusco (2018)'s approach are that this paper accounts for re-suborindation of secondary liens when the first lien is being refinanced, and separately accounts for second-liens and HELOCs. These differences are discussed below. ATTOM tracks properties through time by the sr_property_id variable, while homeowners are not tracked via a similar id variable through time. To construct the amount of home equity extracted for refinancing transactions, debt histories are constructed for each property by homeowner tuple. For each purchase transaction, refinance transactions that occur before the next purchase transaction and that contain the same homeowner last name as the purchase transaction are used to update the debt histories.

There are three broad types of refinancing. First, a borrower may take out a home equity line of credit (HELOC). ATTOM identifies HELOC mortgages by the lndr_credit_line variable. A HELOC allows a borrower to draw upon a credit line, similar to a credit card. ATTOM reports the credit limit and not the amount drawn from this credit line. To determine the amount of home equity extracted, it is assumed that the entire credit line is extracted equity.

In a rate-refinance, the borrower does not extract home equity but originates a new mortgage of the same balance as their current outstanding mortgage(s) to obtain a lower interest rate. A borrower may also cash-out refinance, where they take out a new mortgage that increases their total loan balance. Within this type of refinance, the borrower either takes out a secondary lien, and generally the entire loan amount is home equity extracted[1], or takes out a new first lien and pays off their total outstanding mortgage debt. In the latter case, the new first lien balance is larger than the total outstanding mortgage balance (across all liens) and the difference between the two is the amount of home equity extracted. Complicating matters, ATTOM does not differentiate between rate refinance or cash-out refinance and within cash-out refinance does not state whether the refinance is a first or second lien.

Three different debt histories need to be tracked over time for each property by homeowner tuple. The first lien, second lien, and HELOC debt histories are constructed starting with each purchase origination. At the time of the purchase mortgage origination, ATTOM lists first and secondary liens together. During the mid-2000s, second liens at time of pur-

[1]The exception being if there already is a second lien, in which case the amount extracted is the difference between the new second lien amount and the outstanding balance on the existing second lien (if this is positive - a negative amount would imply a rate-refinance of the second lien).

chase were common and referred to as piggyback seconds. These were usually originated to circumvent private mortgage insurance (PMI) requirements for high loan to value (LTV) mortgages. ATTOM only records the origination of each mortgage and does not provide a date the mortgage is terminated (paid off) or the renaming balance due over time. For simplicity, it is assumed that the first and second liens are 30-year fixed rate mortgages with an interest rate equal to the market mortgage rate at the time of origination.[2]

The debt histories of the first and second liens pay down over time using the amortization schedule of 30-year fixed rate mortgages. The debt history for the HELOC balance does not decrease over time since HELOCs are generally interest-only for the first few years of payments. If additional HELOCs are originated, it is assumed that if the new credit limit is higher than the previous credit limit that the difference between the two credit limits is extracted home equity. If the new credit limit is less than the prior credit limit, no home equity is extracted. In both cases the HELOC debt history is updated to the new credit limit.

The challenge with non-HELOC refinances is determining whether the refinance is a rate refinance, a cash-out refinance with a new second lien, or a cash-out refinance with a new first lien. The difference between these last two is that for the former case, the cash-out amount is the entire lien amount (less the current second lien balance if there already is a second lien). In the latter case, the cash-out amount is the new mortgage amount less the balance of all outstanding liens. To help with the categorization, it is noted that with a new first lien, in most cases all existing liens (including second liens and HELOCs) are paid off. This is due to re-subrodination, where the secondary lien holders would have to agree to be re-suborindated when the first lien is being refinanced since they contractually become the first lien holder unless they waive this right (re-subordination). Lenders generally do not waive this right, which necessitates the need to pay off the second liens and HELOCs when originating a new first lien.

To categorize the mortgage, the absolute differences between the new loan amount and all outstanding liens (including HELOCs) and also the outstanding second liens (=0 if no outstanding second lien) are calculated. If the new loan amount is closer in balance to the total outstanding debt, it is assumed that the refinance is replacing all outstanding liens and becoming the new first lien. If the new mortgage amount is more (less) than the total outstanding debt balance, it is assumed to be a cash-out (rate-) refinance. In both cases, the debt histories for existing second liens and HELOCs (if any) are set to 0 and the first lien debt history is set to the new mortgage origination balance with a 30-year amortization schedule. The amount cashed out, if the mortgage is classified as a cash-out refinance, is set to the difference between the new balance less total outstanding mortgage debt.

If the new loan amount is closer in balance to the second lien (which is =0 if there is no second lien), it is assumed that the refinance is associated with a second lien. If there is no outstanding second lien, then the entire loan amount is extracted home equity. If there is an outstanding second lien then the amount cashed out is the the difference between the

[2]The market mortgage rate is obtained from Freddie Mac's PMMS survey.

new loan amount and the current balance of the second lien. If this amount is negative then no home equity is extracted (likely a rate-refinance of the existing second lien). In all three cases, the debt history for the second lien is set to the new loan amount with a 30-year amortization schedule.

A refinance mortgage origination costs approximately $5,000. To account for this, $5,000 is subtracted from the amount of home equity extracted calculated above. If the amount remains positive after this deduction, the refinance transaction is counted as home equity extraction.

A.3 Appendix Figures for Chapter 1

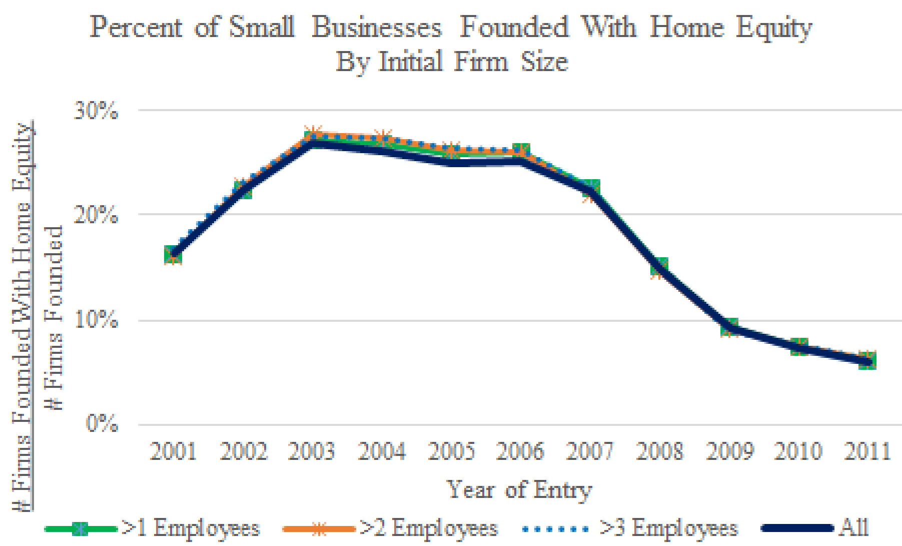


Figure A.1: Share of Small Businesses Founded with Home Equity Funding By Firm Size

The share of entrant small business funded by personal home equity by year of formation. The data are constructed from a merge of ATTOM and NETS. Small businesses are defined as having 10 or fewer employees in the year they are founded. A business is classified as being funded by home equity if the owner extracts over $5,000 of home equity in the year that the business is created or the prior year. Regions follow the Census region classification. The raw underlying data only includes business owners who own a home. To correct for this, the timeseries are adjusted by the home ownership rate of business owners based on the population of business owners by region and year from the American Community Survey micro-data. The data are further segmented by the initial size of the firm.

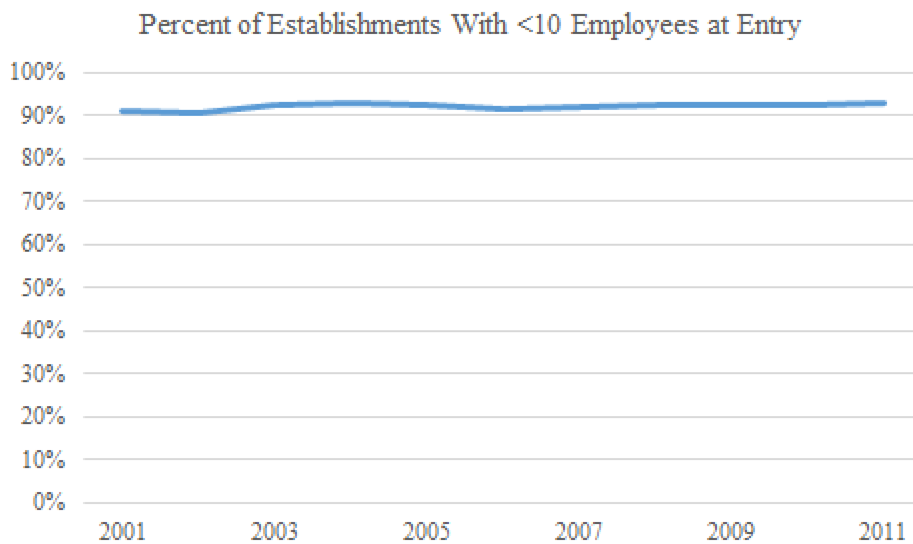


Figure A.2: Share of Entrant Small Business With Fewer Than 10 Employees

The share of entrant establishments with fewer than 10 employees at entry. Data are from the public use Census Business Dynamics Statistics (BDS) and are grouped by year of entry.

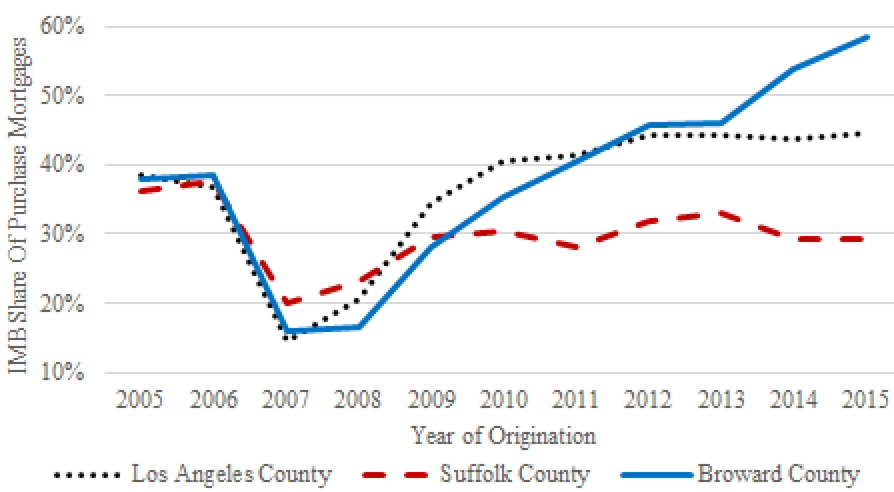


Figure A.3: Share of Mortgage Origination by Independent Mortgage Banks

The share of purchase mortgages originated by independent mortgage banks between 2005 and 2015 for Los Angeles, Suffolk, and Broward counties. The time series are calculated from HMDA and are weighted by loan balance.

Appendix B

Appendix to Chapter 2

B.1 Additional Robustness Tests

In this section, additional robustness tests are provided for the result of the effect of home price shocks on performance. In the baseline specification, fund (for each merged non-disjoint period of time) and time fixed effects are included. As a second specification, fund fixed effects are replaced with commuting zone by time fixed effects and controls to control for fund specific effects and for zip code preferences for their home are included. Including commuting zone by time fixed effects, partials out the common trend in home price growth within commuting zone, which will confirm that the result is driven by idiosyncratic home price growth within a narrow geographic area. After controlling for zip code location preferences it is unlikely that fund managers can forecast their realized idiosyncratic home price growth.

Under this framework, the regression model is:

$$\widehat{\alpha}^{FFC}_{i,t,cz} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \eta_{cz,t} + Controls_{it} + \epsilon_{i,t,cz} \tag{B.1}$$

where $\eta_{cz,t}$ are monthly time by commuting zone fixed effects and $Controls_{it}$ is a vector of control variables comprised of: Morningstar fund category fixed effects, number of manager fixed effects, an indicator for if the manager owns a second home, an indicator for if the manager has an advanced degree (above undergraduate), and non-parametric fixed effects for TNA, home value, combined LTV (CLTV), average zip code level income (from the 2010 IRS SOI), and the percent of non-white households at the zip code level (from the 2000 Census). Standard errors are clustered at the fund level (for each merged non-disjoint period of time). Table B.1 presents results using this second specification. The results are similar to the baseline specification. A one standard deviation positive shock induces a decline in annualized alpha of 38bps.

Next, other outcome variables are considered for the baseline regression model that includes fund and time fixed effects. Table B.2, column 1 estimates the regression on the 1-month ahead raw returns less the return on the fund's benchmark index. This is estimated over the entire 2001 to June 2018 date range without interactions on the home price growth variable. Overall, a marginally significant and negative coefficient is found, indicating that during both normal times and recessions there is a negative effect on returns from home price growth. During periods of economic expansion this is consistent with the results found in this paper and during periods of contraction the result is consistent with Stoffman et al. (2018). Stoffman et al. (2018) find a negative and marginally significant effect of home prices on raw returns during periods of negative home price growth, which they attribute to career concerns. While this paper finds that during periods of positive home price growth, overconfidence from positive home price growth also decreases performance. They found no effect of house price growth on risk-adjusted returns during the Great Recession, also consistent with the results in this paper.

In column 2 of Table B.2, the baseline model is estimated on monthly alpha for the subsequent 12-months. The subsequent one year ahead monthly alpha is estimated from a regression of Fama-French-Carhart on the subsequent 12 months using 12 monthly observations. The result is negative and significant at the 5% level. Column 3 replaces net alpha with gross alpha (alpha before fees) and the result is unaffected. As a fourth method for calculating alpha, Table B.3 replicates Table 2.2 with a 1-month forecast of alpha using ex-ante information. For each monthly fund observation, the prior 12 month period is used to estimate a Fama-French-Carhart regression (the monthly observation is omitted if the fund manager group has less than 12 months of historical data managing the fund). Using these factor loadings, the 1-month ahead alpha is estimated. The result becomes stronger with this procedure.

Another concern may be that for most funds, housing information is only observed for a subset of managers. If the sample is restricted to funds for which a 100% of the managers were merged, the result remains and slightly improves (Table B.2 column 4). The choice of clustering at the fund level does not allow for correlation across funds within a month. Regressions with returns typically cluster at the month level. However, in this paper the effect is being estimated off of changes in home price growth, which are correlated across time. To show that the choice of clustering is not affecting significance, in column 5 of Table B.2 standard errors are clustered by time and the result is unaffected. The remainder of the columns in Table 2.2 are also unaffected if clustered by month (available upon request).

An additional concern could be that the results are unique to the measure of 3-year lagged home price growth. To provide robustness around this, Table B.4 replicates Table 2.2 with the 3-year home price growth measure replaced with 2 and 4-year home price growth measures. The results are largely unchanged, showing that the result is not driven by the choice of the home price growth measure. Another concern is that the choice of restricting the population to fund manager groups that lasted for at least 24 months is causing a survivorship or other type of bias in the results. Table B.5 replicates Table 2.2 on samples where the fund manager groups survived for at least 12 months and survived for at least 36 months. Again, the results are largely unchanged. The above robustness checks rule out many potential concerns with the primary results of this paper.

B.2 Appendix Tables for Chapter 2

Table B.1: Effect Within Commuting Zone by Time of Home Price Growth on Fund Alpha

	(1) All Periods	(2) HPI≥0	(3) HPI<0	(4) All Periods
$ln(\Delta 3\text{yr HPI}_{i,t-1})$	-0.135*	-0.264**	-0.0153	
	(-1.88)	(-2.41)	(-0.10)	
$ln(\Delta 3\text{yr HPI}^{+}_{i,t-1})$				-0.284***
				(-2.60)
$ln(\Delta 3\text{yr HPI}^{-}_{i,t-1})$				-0.0508
				(-0.35)
# Obs	87247	59579	27668	87247
R-squared	0.190	0.202	0.216	0.206

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robustness of the baseline model for monthly forecasts of alpha, with commuting zone by time fixed effects and a battery of controls. Estimation of the regression model:

$$\widehat{\alpha}^{FFC}_{i,t,cz} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \eta_{cz,t} + Controls_{it} + \epsilon_{i,t,cz}$$

where $\eta_{cz,t}$ are monthly time by commuting zone fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative) and $Controls_{it}$ is comprised of: Morningstar fund category fixed effects, number of manager fixed effects, an indicator for if the manager owns a second home, an indicator for if the manager has an advanced degree (above undergraduate), and non-parametric fixed effects for TNA, home value, combined LTV (CLTV), average zip code level income (from the 2010 IRS SOI), and the percent of non-white households at the zip code level (from the 2000 Census). $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table B.2: Effect of Home Price Growth on Various Fund Return Measures

	(1)	(2)	(3)	(4)	(5)
	Y=$R^{\text{Gross}}_{i,t+1}$ $-R_{Benchmark,t+1}$	Y=$\alpha_{i,t+1\to t+12m}$	Y =Gross $\widehat{\alpha}^{FFC}$	Y =Net $\widehat{\alpha}^{FFC}$ Merged 100% of Managers	Cluster by Month
$ln(\Delta 3\text{yr HPI}_{i,t-1})$	-0.113*				
	(-1.67)				
$ln(\Delta 3\text{yr HPI}^{+}_{i,t-1})$		-0.189**	-0.263***	-0.302**	-0.258***
		(-2.10)	(-3.46)	(-2.43)	(-3.38)
$ln(\Delta 3\text{yr HPI}^{-}_{i,t-1})$		-0.154	-0.141	-0.00643	-0.148
		(-1.21)	(-1.27)	(-0.03)	(-1.02)
# Obs	84864	73367	84864	21730	87719
R-squared	0.782	0.264	0.0895	0.116	0.0901

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robustness of the baseline model for monthly forecasts of alpha, with fund and time fixed effects. Estimation of the regression model:

$$Y_{i,t,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. $Y_{i,t,a}$ is either: net $\widehat{\alpha}^{FFC}$, gross $\widehat{\alpha}^{FFC}$, $R_{i,t+1} - R_{Benchmark,t+1}$, or $\alpha_{i,t+1\to t+12m}$ (estimated from a 1-year regression of Fama-French-Carhart with 12 monthly observations). Column 4 is restricted to funds for which 100% of the managers were merged to ATTOM and column 5 repeats the primary regression with clustering by month (instead of by fund). $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table B.3: Effect of Home Price Growth on Fund Alpha, Using Ex-Ante Information

	(1) All Periods	(2) HPI≥0	(3) HPI<0	(4) All Periods
$ln(\Delta 3\text{yr HPI}_{i,t-1})$	-0.190***	-0.264**	0.131	
	(-2.70)	(-2.56)	(0.70)	
$ln(\Delta 3\text{yr HPI}^{+}_{i,t-1})$				-0.312***
				(-3.40)
$ln(\Delta 3\text{yr HPI}^{-}_{i,t-1})$				-0.187
				(-1.20)
# Obs	73087	49010	24077	73087
R-squared	0.0874	0.0841	0.125	0.0904

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robustness of the estimation of alpha in the baseline regression model. Alpha is estimated with factors estimated from ex-ante information. Estimation of the regression model:

$$\widehat{\alpha}^{FFC}_{i,t,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from estimating a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model on the prior 12 months of monthly data (from $t-12$ to t). The factors from these regressions are then used to calculate the alpha in month $t+1$. If the fund manager group has less than 12 months of data then the month is omitted. Standard errors are clustered at the fund level. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table B.4: Effect of 2 and 4-Year Home Price Growth on Fund Alpha

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		2-year HPI Growth				4-year HPI Growth		
	All HPI	HPI≥0	HPI<0	All HPI	All HPI	HPI≥0	HPI<0	All HPI
$ln(\Delta\text{n-year HPI}_{i,t-1})$	-0.179***	-0.270***	0.259		-0.165***	-0.203***	-0.121	
	(-2.67)	(-2.65)	(1.50)		(-3.37)	(-2.68)	(-1.09)	
$ln(\Delta\text{n-year HPI}^{+}_{i,t-1})$				-0.296***				-0.212***
				(-3.28)				(-3.26)
$ln(\Delta\text{n-year HPI}^{-}_{i,t-1})$				0.0751				-0.147*
				(0.53)				(-1.65)
# Obs	87719	58901	28818	87719	81841	55283	26558	81841
R-squared	0.0877	0.0893	0.117	0.0904	0.0880	0.0935	0.115	0.0905

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robustness of the baseline model for monthly forecasts of alpha, with fund and time fixed effects, on the choice of the 3-year home price growth measure. Estimation of the regression model utilizing either 2 year or 4 year home price growth (in lieu of 3 year home price growth):

$$\widehat{\alpha}^{FFC}_{i,t,a} = c + \beta ln(\Delta\text{n-year HPI}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using at least 24 months of data. Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 24 months are dropped and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.

Table B.5: Effect of Home Price Growth on Fund Alpha for Fund Managers Who Survived At Least 12 (36) Months

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Survived $\geq$ 12 months				Survived $\geq$ 36 months			
	All HPI	HPI≥0	HPI<0	All HPI	All HPI	HPI≥0	HPI<0	All HPI
$ln(\Delta 3\text{yr HPI}_{i,t-1})$	-0.163***	-0.219**	-0.0383		-0.159***	-0.229**	-0.00770	
	(-2.90)	(-2.57)	(-0.28)		(-2.75)	(-2.53)	(-0.05)	
$ln(\Delta 3\text{yr HPI}^{+}_{i,t-1})$				-0.242***				-0.256***
				(-3.28)				(-3.30)
$ln(\Delta 3\text{yr HPI}^{-}_{i,t-1})$				-0.149				-0.0920
				(-1.38)				(-0.82)
# Obs	94216	65016	29200	94216	78786	52797	25989	78786
R-squared	0.0915	0.0961	0.116	0.0937	0.0853	0.0877	0.114	0.0878

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robustness of the baseline model for monthly forecasts of alpha, with fund and time fixed effects, on the restriction of fund managers having to survive for at least 24 months. Estimation of the regression model:

$$\widehat{\alpha}^{FFC}_{i,t,a} = c + \beta ln(\Delta 3\text{yr HPI}_{i,t-1}) + \lambda_i + \sigma_t + \rho_a + \epsilon_{ita}$$

where λ_i are fund fixed effects (for each merged non-disjoint period of time), σ_t are monthly time fixed effects (interacted with indicators for positive and negative home price growth in regressions where home price growth is segmented on being positive or negative), and ρ_a are non-parametric fixed effects for TNA. $\widehat{\alpha}^{FFC}_{i,t,a}$ are monthly alphas obtained from backing out the monthly pricing errors from a fund level (for each merged non-disjoint period of time) regression of the Fama-French-Carhart 4-factor model. The pricing model is estimated on monthly data using between 12 and 36 months of data (depending on the columns). Standard errors are clustered at the fund level, for each merged non-disjoint period of time. Unbounded variables are winsorized at the 1% and 99% level. Funds with manager groups that lasted for fewer than 12 or 36 months are dropped (depending on the columns) and fund managers who owned their primary home for fewer than 36 months are dropped (the exception being if a fund manager moved from a previous home that was also merged). Fund months are dropped if there are greater than 10 managers or if TNA, measured in 2000 dollars, falls below $15million. Only actively managed US domestic equity funds are included.