# Mortgage Choices and Housing Speculation*

Gadi Barlevy
*Federal Reserve Bank of Chicago*
gbarlevy@frbchi.org

Jonas D.M. Fisher
*Federal Reserve Bank of Chicago*
jfisher@frbchi.org

June 15, 2011

**Abstract**

We describe a rational expectations model in which speculative bubbles in house prices can emerge. When a bubble emerges, both speculators and their lenders prefer interest-only (IO) mortgages to traditional mortgages. By contrast, absent a bubble there is no scope for mutual gains from using IOs. Using data compiled for over 200 US cities for the period 2000–2008, we find that IOs were used sparingly in cities where elastic housing supply kept house prices in check, but were common in cities with inelastic supply where house prices rose sharply and then crashed. We confirm that the use of IOs in these cities is not proxying for other mortgage market characteristics such as subprime, securitization, or high leverage. Moreover, the use of IOs does not appear to have been a response to houses becoming more expensive; if anything, their use anticipated future appreciation. We also confirm that, as implied by our model, IOs were more likely to be repaid early, and those that survived until prices fell were more likely to default. These findings suggest that the recent boom-bust in the housing market was associated with a speculative bubble.

*JEL* Classification Numbers: E0, O4, R0
*Keywords*: House prices, interest-only mortgages, speculation, bubble

*We thank Sumit Agarwal, Gene Amromin, David Andolfatto, Marco Bassetto, Huberto Ennis, Simon Gilchrist, Lars Hansen, Benoit Mojon, Monika Piazzesi, Ned Prescott, Amit Seru, and Amir Sufi for helpful comments. We are grateful to Albert Saiz for sharing his geographical data. We also thank Ryan Peters and Shani Schechter for superb research assistance. The views expressed herein are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

# 1 Introduction

The recent financial crisis has re-focused attention on the housing market and its apparent vulnerability to large swings in house prices. As evident from the U.S. experience, such cycles can severely disrupt financial markets and adversely affect real economic activity. Economists and policymakers have therefore sought to understand when and why boom-bust cycles can arise in the housing market. Are such price movements driven by fundamentals, or do they reflect speculation in which prices increasingly drift away from the expected value of the services the underlying assets provide? Are there any indicators that can predict when such boom-bust episodes might occur if policymakers wish to intervene before they develop?

This paper examines whether mortgage data can help address these questions. Our focus on the mortgage market is motivated by theoretical work that suggests credit markets can play a key role in allowing for speculative bubbles, *e.g.* Allen and Gorton (1993) and Allen and Gale (2000). These papers show that if traders finance asset purchases with borrowed funds, they would agree to pay more for a risky asset than its expected value. This is because they can always default should their gamble fail. Lenders would ordinarily refuse to finance such speculation that comes at their expense. But if lenders are unable to distinguish speculators from profitable borrowers, they may end up financing speculators after all.

If credit markets play a role in allowing for speculative bubbles, then credit market data may be relevant for predicting the occurrence of such episodes. For example, if borrowers bid up house prices above their underlying value because they can default if prices collapse, boom-bust cycles might be more likely to emerge when borrowers are able to use greater leverage. Indeed, previous work by Lamont and Stein (1999) has found that house prices tend to be more volatile in cities where a larger proportion of mortgages are highly leveraged.[1]

While previous work has focused on leverage, we consider other mortgage characteristics. Our motivation comes from Barlevy (2009), which argues that lenders have an incentive to offer particular types of contracts to the speculators they end up financing. We build on this insight and argue that when lenders know that some of those they lend to are speculating on overvalued assets, lenders and speculators can both be made better off using contracts with backloaded payments, *i.e.* contracts where the initial payments stipulated in the contract are low and later payments are onerously high. Lenders prefer these contracts because

[1] More precisely, Lamont and Stein (1999) show that house prices respond more to income shocks in cities with a larger share of mortgages whose loan-to-value ratio exceeds 80%. Their analysis is motivated by the work of Stein (1995) on down-payment constraints. In his model, house prices reflect fundamentals. Down-payment constraints impede the efficient allocation of houses and make the fundamentals more volatile, similarly to Kiyotaki and Moore (1997). This hypothesis is distinct but not mutually exclusive of the forces we study.

they preclude borrowers from gambling at their expense for too long; once payments rise, speculators must sell the asset (or else refinance with another lender, making the speculator someone else's problem). At the same time, borrowers prefer these contracts because they can defer building equity in what they know is a risky asset, allowing them to default on more principal should the prices collapse early. These contracts essentially compensate the borrower for committing to settle his debt earlier than he would under a traditional mortgage contract. We further show that this mutual preference for backloaded contracts is intimately related to the fact that the asset is a speculative bubble; if it were not, then absent any other frictions, backloaded contracts could no longer make both the lender and the borrower better off.

These results lead us to look at whether housing markets with boom-bust cycles also involved greater use of backloaded mortgages. We find that the use of backloaded payments, specifically interest-only (IO) mortgages, was highly concentrated in cities that experienced large boom-bust cycles.[2] In particular, we find that IOs were used only sparingly in areas with few restrictions on the supply of housing, where speculative bubbles couldn't emerge in principle. However, in cities with geographical and regulatory supply constraints, such contracts were quite prevalent, yet only in those cities that experienced boom-bust cycles.

To convey the spirit of our findings, consider two cities: Phoenix, AZ and Laredo, TX. Laredo is a low income border city in a state with little regulation and vast open spaces on which new homes can be built. By contrast, although Phoenix also has plenty of open space, it ranks relatively high on the Wharton Residential Land Use Index complied by Gyourko, Saiz, and Summers (2008). Figure 1 shows the Federal Housing Finance Agency (FHFA) house price index for these two cities, deflated by the Consumer Price Index. Real house prices in Laredo grew at roughly 2.5% per year between 2000 and 2008. In Phoenix, house prices grew much faster, averaging 9.5% per year between 2001 and 2006 and rising 36% in 2005 alone. Prices then fell sharply, reverting to their 2001 levels by 2010. The fact that cities with geographical or regulatory restrictions on housing supply have more volatile housing prices has been pointed out before; see, for example, Krugman (2005) and Glaeser, Gyourko, and Saiz (2008). However, Figure 1 also shows that home buyers in the two cities relied on different types of mortgage contracts to finance their purchases; at the peak, over 40% of all new mortgages for purchase in Phoenix were IO, but IOs accounted for at most 2% of mortgages for purchase in any given quarter in Laredo.

The association between extensive use of IOs and rapid house price appreciation evident

[2] The only other paper we know of that argues house price appreciation is associated with backloaded contracts is the testimony of Thompson (2006) at a congressional hearing on nontraditional mortgage products. She points out the statistical pattern without analyzing it.

in these two cities remains when we look at a cross-section of over 200 cities, and is robust to controlling for various city-level characteristics. We also show that IOs are not merely a proxy for some other mortgage characteristic such as subprime lending, securitization, high leverage, or borrowing against investment properties. Indeed, the frequency of these other types of mortgages appears to be only weakly correlated with the frequency of IO usage. Lastly, we argue that it is unlikely that borrowers simply flocked to IOs for reasons of affordability. In particular, we find no evidence that increasing house prices anticipate the use of these mortgages. In fact, we find the opposite pattern; the use of IOs anticipates price increases in cities with heavy IO usage. This pattern can be seen in Panel A of Figure 1. The use of IOs in Phoenix began in early 2004, while house prices only took off in late 2004 and early 2005. This pattern is consistent with our model, although we should emphasize that the model does not predict that the use of IOs will necessarily lead price growth.

We view the contribution of our paper as twofold. First, we document new facts about the recent boom-bust cycle in the housing market that any theory of house prices ought to explain, namely that rising house prices coincided with the use of certain types of mortgage contracts, and that the use of these contracts anticipated rather than followed the rise in prices. The latter finding may also be relevant for policymakers, since it suggests a potential warning indicator of boom-bust patterns. Second, we offer a new approach for identifying the existence of speculative bubbles that does not involve estimating the fundamental value of an asset and comparing it to its price. Rather, our approach looks at auxiliary behavior such as the choice of a particular type of mortgage that ideally should only be observed if there was a bubble. This is true in our specific model, although it obviously ignores various reasons that borrowers and lenders might prefer IO mortgages. Still, one can in principle test these alternative explanations, and indeed we consider and reject some of them. Ruling out alternative explanations for a particular behavior such as contract choice may be easier in practice than testing the restrictions on the stochastic process for future dividends that must be imposed in order to infer an asset's fundamental value.

The paper is organized as follows. In the next section, we discuss the theoretical environment that motivates us to focus on backloaded mortgages. In Section 3, we describe the data we use. Section 4 documents the cross-sectional relationship between house price appreciation and the types of mortgages agents choose. Section 5 shows that in cities where IO contracts were common, their use anticipated rather than followed house price appreciation. Section 6 confirms that, as implied by our model, IOs were more likely to be repaid early, and those that survived until prices fell were more likely to be defaulted on. Section 7 concludes.

# 2 Theory

In this section, we develop a model of speculation on housing due to risk-shifting, which builds on Allen and Gorton (1993), Allen and Gale (2000), and Barlevy (2009). We first show that booms and busts in house prices are possible, and that these may or may not be associated with a speculative bubble. We then show that backloaded mortgages will only be used in our model when booms and busts involve a bubble.

## 2.1 Economic Environment

A key feature of our setup is that agents differ in how much they value home ownership: some strongly prefer owning a home to renting, *e.g.* because they can customize the house to their personal tastes, while others derive no additional benefit from owning over renting. The former are natural owners who would want to hold on to their homes even if house prices fall. The latter are ordinarily indifferent between owning and renting, but may wish to buy houses for speculative reasons under certain conditions. Preferences towards home ownership are private information.

Formally, suppose agents are infinitely-lived, discount the future at rate $\beta$, and have preferences over a representative consumption good and housing services given by

$$\sum_{t=1}^{\infty} \beta^t \left(c_t + \nu \mathbb{I}_t\right),$$

where $c_t$ denotes consumption at date $t$, $\mathbb{I}_t$ is an indicator that is equal to 1 if the individual occupies a house at date $t$, and $\nu$ is a taste parameter that varies across types. We set consumption as the numeraire. Low valuation types are assumed to value occupying a house for a period at $\nu = \left(\beta^{-1} - 1\right) d$ consumption units, whether or not they own the house they occupy. Hence, their valuation of owning a house indefinitely is $d$. High valuation types receive the same flow utility of $\nu = \left(\beta^{-1} - 1\right) d$ from renting. However, if they own the house they occupy and can therefore customize it, they receive a higher service flow $\nu = \left(\beta^{-1} - 1\right) D$ where $D > d$, so that their valuation of owning a house indefinitely is $D$. Individuals can occupy only one house. We assume that once high types customize a house, they cannot derive the same high flow elsewhere. This ensures that if house prices fall, high types who borrowed to buy their house will not default and move to a cheaper house.[3]

We analyze the equilibrium for a single city. The initial stock of houses in this city at

[3] In practice, high types who value owning may be reluctant to default and move for other reasons, most notably because default would hurt their credit score and make it difficult for them to buy a new home.

date 0 is normalized to 1, but can grow over time. This stock can be purchased by a number of potential homeowners, which can also grow over time. Agents who do not buy a house can either rent in the city or move on to another city. The way houses are allocated before trade starts at date 0 is irrelevant for our analysis, so we do not specify it. We also assume an unlimited number of agents who are willing to rent a house in this city for $(\beta^{-1} - 1)\, d$ per period. That is, not all renters need to be potential homeowners, and homeowners can always count on finding someone to whom they can rent.

As a preliminary step, consider the case where the number of potential buyers and houses remain constant over time. Since houses can always be rented out at $(\beta^{-1} - 1)\, d$ per period, there would be excess demand for houses if the house price were below $d$. The price that clears the market depends on how the number of houses compares with the number of high type buyers. If there are more houses than high types, so some houses must be occupied by low-type owners or renters, house prices will equal $d$: Prices must be at least $d$, and by a standard argument cannot exceed $d$ without violating some agent's transversality condition.[4] If there are more high types than houses, so all houses will be occupied by high types in equilibrium, the price of houses could not fall below $D$, and the transversality condition imples it cannot exceed $D$. The equilibrium house price thus corresponds to the value of an additional house if one became available and could be auctioned off to potential buyers.

In what follows, we assume that there are initially more houses than high types. Let $\phi_0 \in (0, 1]$ denote the mass of houses not occupied by high types at date 0. We then let a random number of new potential buyers arrive, starting at date 0. New potential buyers include high and low types, so the number of high types may exceed the number of houses. We think of this shock as reflecting financial innovation that allows agents shut out of credit markets to buy a home.[5] However, the shock can be equally viewed as a migration wave.

We further assume that new potential buyers arrive gradually, so agents might remain unsure for some time as to whether there ultimately will be more high types than houses. Specifically, we assume new potential buyers arrive sequentially in cohorts of constant size $n$ until some random date $\mathcal{T}$ that is distributed geometrically, *i.e.* $\Pr(\mathcal{T} = t) = (1 - q)^{t-1}\, q$ for $0 < q < 1$ and $t = 1, 2, \ldots$. Agents do not know when the flow of new arrivals will stop until date $\mathcal{T}+1$, the first date in which no new buyers arrive. Up until $\mathcal{T}+1$, agents assign probability $q$ that new buyers will cease to arrive next period. To simplify the analysis, we

[4] In particular, if the price exceeded $d$, no one will hold houses not occupied by high types unless they expect to sell these houses at even higher prices in the future. This implies there cannot be any finite date beyond which house prices always grow by less than $\beta^{-1} - 1$. But this violates the transversality condition, which holds that the value of any asset at date $t$ discounted to the present tends to 0 as $t \to \infty$.

[5] For evidence on the expansion of credit in the period we consider, see Mian and Sufi (2009).

assume that between dates 1 and $\mathcal{T}$, only those who arrive each period can buy houses, *i.e.* agents cannot time or delay their purchases.

Since one way to at least partly accommodate new buyers is to build new homes, we need to specify the cost of building additional houses. We assume these costs are constant and equal to $d$, *i.e.* absent constraints on supply, new homes can be added at a cost that is no more than the valuation of low types.

Lastly, we assume new buyers own no initial resources and must borrow the full price of the house to purchase one. This is in line with interpreting new arrivals as agents with limited credit access. However, as we shall see, leverage plays an essential role in allowing for a bubble. In what follows, we allow lenders to offer only *non-recourse* mortgage contracts. Agents who borrow must make a sequence of payments $\{m_\tau\}_{\tau=1}^{T}$ where $\tau$ indexes time since the loan was taken out and $T$ represents the term of the mortgage. If an agent fails to make a payment, he is found in default and ownership of the house transfers to the lender. Non-recourse means that once a lender takes possession of a house, he cannot go after the borrower's other income sources.[6] To ensure borrowers can potentially repay their debt, we assume they receive an income flow $\{\omega_\tau\}_{\tau=1}^{\infty}$ where $\omega_\tau$ are such that there exists a sequence of mortgage payments $m_\tau \leq \omega_\tau$ that allow the borrower to pay off his debt in finite time. At the same time, $\omega_\tau$ cannot be so high that the borrower can repay his debt obligation too rapidly, in a sense we make precise below.

## 2.2 Equilibrium House Prices

We now characterize house prices when new potential buyers arrive over time. We argue that prices can exhibit a boom-bust pattern, which may or may not involve a speculative bubble.[7]

The path of house prices naturally depends on what happens with housing supply. Absent restrictions on supply, the equilibrium house price will equal building costs $d$, since otherwise supply of houses would be infinite. If house prices are constant over time, loans are fully collateralized and the equilibrium interest rate equals the risk-free rate $\beta^{-1} - 1$. At price $d$ and a mortgage rate equal to $\beta^{-1} - 1$, newly arriving high types will strictly prefer to buy a home and low types will be indifferent about buying. Therefore, absent supply constraints, the equilibrium price is identical to the benchmark case with constant numbers of houses and buyers where there are more houses than high types.

[6] For a discussion on recourse in the U.S., see Ghent and Kudlyak (2009). They argue that even in states that allow for recourse, lenders often find it unprofitable to go after other sources of income.

[7] For an alternative theory of boom-bust cycles in housing, see Burnside, Eichenbaum, and Rebelo (2010).

The more interesting case is when new construction is constrained. It is in this case that boom-bust cycles and speculative bubbles can emerge. For simplicity, consider the extreme case where no new construction is possible, so the supply of houses if fixed.[8] Let $\phi$ denote the fraction of new arrivals each period who are low types. We assume $\phi$ is small, which ensures that equilibrium interest rates will be close to the risk-free rate. Since the mass of homes not initially owned by high types is $\phi_0$, and each period $(1-\phi)\,n$ of these houses would have to be reallocated to newly arriving high types to ensure all high types occupy a house, buyers would need to keep arriving for $\phi_0/\left[n\left(1-\phi\right)\right]$ periods before the number of high types surpasses the stock of houses. Let $t^*$ denote the smallest integer strictly greater than $\phi_0/\left[n\left(1-\phi\right)\right]$. Then by date $t^*$, all uncertainty about the housing market will be resolved. If buyers stop arriving before $t^*$, *i.e.* if $\mathcal{T} < t^*$, then there will be fewer high types than houses, and house prices will equal $d$ from date $\mathcal{T}$ on. But if buyers arrive through $t^*$, *i.e.* $\mathcal{T} \geq t^*$, then there will be more high types than houses, and the price at $t^*$ will be $D$.[9]

When the size of arriving cohorts $n$ is either very small or very large, there will be no uncertainty on whether there will be more high types or houses. If $n = 0$, then $t^* = \infty$ and $\Pr\left(\mathcal{T} < t^*\right) = 1$, so house prices equal $d$ at all dates. Conversely, if $n > \phi_0/\left(1-\phi\right)$, the mass of high types that arrive in the first period is enough to buy out any houses not yet occupied by high types. In this case, $t^* = 1$ and $\Pr\left(\mathcal{T} \geq t^*\right) = 1$. The price of housing would immediately jump to $D$. House prices rise because land becomes genuinely more valuable given the marginal unit of land can now deliver more housing services.

For intermediate cohort sizes, *i.e.* $0 < n < \phi_0/\left(1-\phi\right)$, $t^*$ will be finite but larger than 1. As long as new potential buyers keep arriving, the price at date $t^*$, $p_{t^*}$, will remain uncertain, since it is uncertain whether the number of high types will eventually outnumber the stock of houses. We now show that this uncertainty will lead to boom-bust pattern that may or may not involve a speculative bubble.

As a first step, we need to define the fundamental value of a house. We define it to be the shadow value of an additional house, *i.e.* the value of relaxing the constraint on the total number of available houses in the city.[10] Since we assume that there are many potential

[8] We can always reinterpret the fixed supply model as a model with some construction by redefining the flow of new buyers $n$ as the flow of new arrivals *net of new construction*.

[9] We could allow supply constraints to be temporary, so new construction would eventually drive house prices down to the cost of construction $d$. A temporary shortage would still lead the price to exceed $d$ at $t^*$ if $\mathcal{T} \geq t^*$, even if they remain below $D$. Our results only require that house prices be uncertain at date $t^*$.

[10] Another common definition for the fundamental value of an asset is the value of holding it indefinitely, taking into account the issues in Allen, Morris, and Postlewaite (1993) when different agents value the asset differently. As long as the set of rental contracts is rich enough that home owners can rent out houses to high types in a way that lets them derive a high service flow, *e.g.* a perpetual lease where the renter cannot be evicted as long as he meets his payments, this definition yields the same value as ours.

renters willing to live in the city, an additional house can always be used to deliver at least $(\beta^{-1}-1)\,d$ in housing services per period. But in any period with more high types than houses, an additional house can generate $(\beta^{-1}-1)\,D$ housing services by allocating it to a high type. Since there are more houses than high types before $t^*$ by definition, the flow value in these periods is $(\beta^{-1}-1)\,d$. After $t^*$, the flow value will be either permanently high or low, depending on whether enough high types arrived to exhaust the stock of housing. From our earlier analysis with constant numbers of both houses and people, it follows that $p_{t^*}$ will equal the value of housing to the marginal buyer at date $t^*$. Hence, the fundamental value of housing is given by

$$f_t = \left[\sum_{s=t+1}^{t^*} \beta^{s-t}\left(\beta^{-1}-1\right) d\right] + \beta^{t^*-t} E_t\left[p_{t^*}\right]. \tag{1}$$

These fundamentals evolve over time as follows. Until date $t^*$, the fundamentals keep growing as long as buyers arrive and collapse if arrivals stop; if buyers keep arriving until $t^*$, the fundamentals remain high forever. The intuition for this boom-bust pattern is as follows. The relevant uncertainty in the model is whether the number of high types will exceed the number of houses at date $t^*$. If that happens, a marginal house becomes more valuable, since it can be allocated to a high type. When new agents arrive, the scenario where the number of high types exceeds the stock of housing at $t^*$ remains possible. Since date $t^*$ is now one period closer, the fundamentals rise because of discounting. In addition, each arrival makes the event that high types ultimately outnumber houses more likely, so $E_t\left[p_{t^*}\right]$ rises. If new buyers stop arriving, agents know the flow value will be $d$ forever, so $E_t\left[p_{t^*}\right]$ and $f_t$ both fall to $d$. The higher the ratio $D/d$ and the smaller is $q$, the larger the incremental changes in fundamentals will be each period.[11]

If new buyers were to buy houses with their own funds rather than borrowed funds, or if they repaid their loans sufficiently quickly, similar arguments as in the benchmark case with constant numbers of buyers and houses establish that $p_t = f_t$ is the unique market clearing price. Thus, a boom-bust pattern can be compatible with asset prices reflecting fundamentals. However, we now show that house prices exceed $f_t$ when buyers are sufficiently leveraged, so a boom-bust pattern can also be associated with a bubble. Since house prices would then no longer reflect the true value of additional homes, in a richer model it may be possible that in this case new homes will be built even when the cost of construction exceeds the value of these homes. A bubble would also involve a larger price collapse if buyers stop

[11]This process for fundamentals is similar to Zeira (1999). He assumes dividends keep growing until a random date, while we assume dividends can jump up at a known date with some probability. In both cases, positive news – dividends will keep growing, or the jump remains possible – will raise the fundamental value, while negative news – dividends will stop growing, or the jump will not occur – will lower it.

arriving, and hence larger losses to lenders. Therefore, while fundamental and bubble paths can lead to similar price paths, the two can have different economic implications.

To show that bubbles are possible, let us try to construct an equilibrium where $p_t = f_t$. If this path were an equilibrium and the interest rate on mortgages was close to the risk-free rate $\beta^{-1} - 1$, which it will be when $\phi$ is small, both high and low types would prefer to buy houses upon arrival. High types value a house at $D > f_t$ and would prefer to buy than to move on, while low types can assure themselves positive expected profits – and hence more consumption – by buying a house, waiting one period, then selling if new buyers arrive and defaulting otherwise. More generally, if $p_t = f_t$, low types should hold on to a house for as long as their outstanding debt obligation at the end of the period exceeds $d$, since the option to default increases the value of waiting to sell the asset.

A problem with sustaining $p_t = f_t$ as an equilibrium emerges if the only way to meet the demand of new arrivals is for low types who previously bought homes and still owe at least $d$ agree to sell them. In particular, if all $\phi_0$ houses not originally occupied by high types sell before date $t^*$, new buyers could only buy from low types who bought houses after date 0. If the latter owe more than $d$ on their house, they would require more than $f_t$ to sell.

Formally, observe that lenders will only lend against one house per person, since any property not occupied by its owner will be used for speculation. Hence, for the first $\phi_0/n$ periods, some of the houses not originally occupied by high types will not sell. Let $t^{**}$ denote the smallest integer strictly greater than $\phi_0/n$. By date $t^{**}$, we will be assured that if all new arrivals buy a house, some of them must purchase houses from previous (low type) buyers who bought their houses after date 0. Recall that it takes $\phi_0/\left[(1-\phi)\,n\right]$ periods for the number of high types to exceed the stock of housing, so $t^{**} \leq t^*$. If either $t^{**} = t^*$ or if mortgages are such that the outstanding debt obligation after $t^{**}$ periods is less than $d$, there will be no need to induce any low types who owe at least $d$ to sell their homes. In this case, $p_t = f_t$ will indeed be an equilibrium. But if $t^{**} < t^*$ and all agents who arrived between dates 0 and $t^{**}$ owe at least $d$ to their respective lenders, then if buyers arrive past $t^{**}$, someone other than the original owners of the $\phi_0$ houses would have to sell for the market to clear, and in that period the price must exceed $f_t$ to induce them to sell. In Appendix A, we show that this implies $p_t$ will exceed $f_t$ in all periods until date $t^*$ or the date buyers stop arriving. That is, for $t < \min\left(\mathcal{T}+1, t^*\right)$, the equilibrium price is given by

$$p_t = f_t + b_t$$

where $b_t > 0$. The equilibrium price path thus involves a bubble that bursts with constant probability $q$ until date $t^*$, at which point it bursts for sure.

We can further characterize the evolution of the bubble conditional on its not bursting. Since some original owners must hold on to the house until $t^{**}$, the original owners must be indifferent between selling at date $t$ and waiting until $t+1$, entitling them to a service flow $(\beta^{-1}-1)\,d$ for one more period before selling. This indifference implies

$$p_t = \beta\left[\left(\beta^{-1}-1\right)d + (1-q)\,p_{t+1} + qd\right].$$

Substituting in $p_t = f_t + b_t$ and using the recursion $f_t = \beta\left[(\beta^{-1}-1)\,d + qd + (1-q)\,f_{t+1}\right]$ yields the condition $b_{t+1} = (\beta\,(1-q))^{-1}\,b_t$. Intuitively, owners who wait to sell risk losing the chance to sell the asset at an overvalued price. As such, they need to be compensated for waiting, and this compensation accrues as capital gains if the bubble does not burst. We thus have a stochastically bursting bubble as in Blanchard and Watson (1982):

$$b_t = \begin{cases} (1+g)\,b_{t-1} & \text{with probability } 1-q \\ 0 & \text{with probability } q \end{cases}$$

where $g = (\beta\,(1-q))^{-1} - 1$. Beyond date $t^{**}$, the bubble will still grow, but at a lower rate than $g$. This is because agents who value default do not need as much compensation for holding on to the asset, since they do not lose as much if the bubble bursts before they sell.

Note that after date $t^{**}$, the buyer who is just indifferent between holding and selling a house will have some debt obligation against it. This implies that an agent with no debt obligation would strictly prefer to sell the house. Hence, when there is a bubble, there will be a finite date – $t^{**}$, to be precise – beyond which unleveraged agents would prefer to sell a house if they owned one. This result will prove to be important below.

In sum, we have shown that when housing supply is constrained and there is some uncertainty about future house values, booms and busts in house prices can occur with or without bubbles. A speculative bubble only emerges if financial contracts leave agents sufficiently indebted $t^{**}$ periods after taking out their loans. Such a bubble arises because low types who are otherwise indifferent between owning and renting strictly prefer to buy because they can profit if house prices go up and default if not. Lenders fund these speculators because they cannot distinguish them from high types, to whom it is profitable to lend.

## 2.3 Mortgage Contract Choice

We now turn to the question of how speculative bubbles affect the types of mortgages agents use. We show that in our model, backloaded contracts will be used if booms and busts are associated with a speculative bubble, but not if they are driven by fundamentals.

As we noted earlier, we restrict lenders to simple mortgage contracts with a fixed repayment schedule $\{m_\tau\}_{\tau=1}^{T}$. More generally, repayments can be contingent on some underlying state. In our model, the payoff-relevant state space can be summarized by $\mathcal{T}$, which captures the state of the housing market, and what the borrower does with the house he bought. We ignore contracts that condition on these variables. Our results would go through if lenders offered such contracts in addition to the contracts we study. However, we do need to assume lenders do not offer contingent contracts only, to the exclusion of other contracts. As our model is written, lenders have an incentive to do just that. For example, they can benefit from only offering contracts where payments rise with $\mathcal{T}$ and penalize a borrower for selling his house. These stipulations would not matter to high types who plan to occupy the house forever and only care about their expected repayment, but they will deter low types, who only profit from selling their house when new buyers arrive. In practice, of course, lenders offer non-contingent mortgages, presumably for reasons not captured in the model.[12]

Given these restrictions on contracts, all lenders can do is structure the sequence of payments $\{m_\tau\}_{\tau=1}^{T}$. We allow lenders to offer two types of repayment schedules. The first is a traditional fixed-rate mortgage in which payments are constant over time:

$$m_\tau = \frac{r\,(1+r)^T}{(1+r)^T - 1} L \quad \text{for } \tau = 1, ..., T \tag{2}$$

These payments imply the lender earns a return $r$ on any outstanding principal. The other type of mortgage we allow sets payments that rise with $\tau$. While there are different ways to backload payments, we focus on one particular way: interest only (IO) mortgages. These are by far the most popular backloaded mortgage product in the period we study. An IO mortgage requires the borrower to only pay interest for the first $T_0$ periods, and then repay as under a fixed-rate mortgage with term $T - T_0$:

$$\widehat{m}_\tau = \begin{cases} \widehat{r}L & \text{if } \tau = 1, .., T_0 \\ \\ \dfrac{\widehat{r}\,(1+\widehat{r})^{T-T_0}}{(1+\widehat{r})^{T-T_0} - 1} L & \text{if } \tau = T_0 + 1, .., T \end{cases}$$

To justify assuming lenders only offer these two schedules, recall that lenders would like to avoid lending to low types. Since low types prefer to keep their outstanding debt as

[12] For example, in practice the housing market depends on many factors and not a single statistic $\mathcal{T}$, making such contracts costly to implement. While there are mortgages that penalize early sale and repayment, the vast majority of mortgages do not, most likely because good borrowers value the option to move and there are gains to catering to such borrowers. Speculators might also be able to circumvent such penalties by permanently leasing rather than selling their homes where the renter has the right to keep the house if he makes his payments.

large as possible, in any equilibrium where both types borrow, high types will be offered the contract with the fastest repayment path, *i.e.* which fully exhausts the borrowers' income until the debt is repaid:

$$m_\tau = \omega_\tau \quad \text{for } \tau = 1, ..., T \tag{3}$$

where $T$ is set so the lender earns a return $r$ that ensures zero expected profits. Otherwise, a lender could further frontload payments in a way that makes high types better off and low types worse off, allowing him to cherry-pick high types and earn positive profits. Assuming lenders offer fixed-rate mortgages is thus equivalent to assuming $\omega_\tau$ is constant.

There is no similar justification for why lenders should offer an IO contract in particular, but we can give some sense of why lenders may want to offer a backloaded contract in addition to (2). Since (3) implies the traditional mortgage exhausts the borrowers income, if a borrower were to take out an IO mortgage at the same interest rate $r$ as on fixed-rate mortgages where $r > \beta^{-1} - 1$, he could not not afford to meet all of the payments on a more backloaded contract if he could only save at the risk-free rate. Thus, a borrower who chose a backloaded contract would at some point have to sell his house, default, or refinance. A backloaded contract thus forces the borrower to repay his debt earlier than he would have to under a fixed-rate mortgage, leaving the lender exposed to the risk of the bubble bursting for a shorter period. While the lender will lose more if a bubble bursts early, it may be desirable to shorten the number of periods he remains exposed to this risk.

Hence, the key feature of a backloaded contract is that it forces early repayment. To make things simple, we assume an agent will be forced to repay as soon as the payment is reset, so after $T_0$ periods, even though borrowers might be able to meet payments for some time afterwards with their savings. We also assume that at this point, borrowers must default or sell the asset, and cannot refinance. The reason is that borrowers would not be able to refinance if lenders observe the contract they chose before. In principle, we can complicate the model by adding types who choose IOs for other reasons. We prefer to simply ignore refinancing, but this is not consequential. If borrowers could refinance with another lender, a backloaded contract would achieve the same goal of forcing the borrower to repay early, and would thus lead to the same results.

In sum, lenders can offer two types of repayment schedules, and borrowers choose among the contracts offered. Once a borrower chooses a contract, he must subsequently choose each period whether to repay his debt, make the stipulated payment, or default. Using backwards induction, both parties can figure out the expected payoffs from any given contract, and choose accordingly. We show how to construct these payoffs in Appendix A.

We now sketch the argument for why, when there is a bubble, we can always find some

IO mortgage that both the borrower and the lender prefer to a fixed-rate mortgage at the same interest rate. Recall that in equilibrium, if a bubble exists, then beyond date $t^{**}$ all unleveraged agents would strictly to prefer to sell a house if they owned one. Since the lender and borrower collectively have no debt obligations against the house, their joint interests are maximized by selling the house. But recall that for a bubble to exist, equilibrium contracts must leave agents owing at least $d$ after $t^{**}$ periods, including the maximally frontloaded contract. So under the fixed-rate contract, the borrower would hold on to the house. Hence, after date $t^{**}$, borrowers and lenders can make themselves better off by agreeing to sell the house and splitting the proceeds appropriately. That is, the amount the lender would agree to pay the borrower to sell the asset exceeds the amount the borrower requires to give up the option to keep the house one more period. Since a backloaded mortgage forces early repayment, it can capture these gains. If the reset date $T_0 = 0$, only the lender will be made better off relative to the fixed-rate mortgage. If instead $T_0$ is pushed as far as possible to a date at which the borrower would have sold anyway, only the borrower be made better off, since the forced sell date is not binding and he avoids building equity in a risky asset. Given the borrower and lender are collectively better off selling early, then ignoring integer constraints, by continuity there should exist some intermediate value of $T_0$ which leaves both parties better off. Note they are benefitting at the expense of a third party, a burden that ultimately falls on high types who arrive later and pay higher prices and interest rates to buy a home. An IO is thus optimal for the borrower and lender, but not for society as a whole.

Consider the following numerical example. Set $T = 30$ and $T_0 = 5$, in line with the terms on the modal IO mortgage for the period we study. Let $\beta = 0.97$, implying a discount rate of 3% per year, and set the real interest rate $r = 0.04$, slightly above the risk-free rate. We normalize $d = 1$ and set $D = 30$ so assets can appreciate at empirically plausible rates. We set $q$ to 0.2, implying the average duration of a bubble is 5 years. We assume the bubble starts at $b_0 = 0.1$ and then grows at a constant rate $g = 0.05$ to keep things simple. This implies house prices $p_t$ will grow at a rate of between 10 and 15% in the first five years. This in on par with the average annual house price appreciation of the cities in our sample with the fastest appreciation. Finally, we set $t^*$ to 15, implying any uncertainty about housing prices will be resolved after 15 years and that the odds that $\mathcal{T} > t^*$ are small, $(1-q)^{15} = 0.03$. For these parameter values, we can use the calculations in Appendix A to show that both borrower and lender prefer the IO to the fixed-term mortgage when they first take out the loan. If we increased $q$ to 0.5, the lender will no longer agree to the interest-only contract with $T_0 = 5$, since the odds that the bubble bursts within the first five years are now higher. However, both parties would prefer a contract with an interest-only period of $T_0 = 4$ to the fixed-rate mortgage.

In Appendix A, we show that if both lenders and borrowers prefer some IO contract to a fixed-rate mortgage with the same interest rate, then in equilibrium lenders will offer both contracts; high types choose the fixed-rate mortgage and low types choose the interest-only mortgage. Furthermore, interest-only loans carry a higher interest rate in equilibrium. For example, for our numerical example above, given $r = 0.04$ on the interest-only contract, the equilibrium interest rate on traditional mortgages will be 0.035, or 50 basis points lower. This is on par with the empirical penalty for the interest-only option.[13]

The fact that the equilibrium is separating may seem surprising: Why don't lenders offer only those mortgages that high type borrowers take? The answer is that in equilibrium, low types are indifferent between the two mortgages. What lenders will choose to do depends on what they believe low types would do if only offered fixed-rate contracts. In equilibrium, lenders must believe speculators will accept fixed-rate contracts with positive probability if only offered those. Otherwise, all lenders would offer only the fixed-rate contract. But low types would take the fixed-rate contract if that was the only contract available. In equilibrium, then, lenders must expect not to benefit from only offering fixed-rate mortgages.

Of course, perfect sorting as implied by the model is unrealistic; non-speculators may prefer backloaded mortgages for reasons not captured by our model, such as liquidity constraints, and some speculators may for whatever reason choose a traditional mortgage. Indeed, if mortgages were perfectly separating, backloaded mortgages would be revealed to be unprofitable, but in practice these mortgages were bought and sold at positive prices. Thus, traders must have believed some IO mortgages were profitable.[14] However, our key result is not what contracts are taken by the agents who cross-subsidize speculators, but the fact that a speculative bubble encourages greater use of backloaded contracts.

Next, consider the case where there is no bubble, so $p_\tau = f_\tau$. The next proposition shows that absent other frictions, when there is no bubble, there is no way to make both parties better off by moving to an IO contract. The intuition is the opposite of what happens when there is a bubble. Since the borrower and lender are no longer collectively better off selling the asset early, there is no scope for gains from trade.

**Proposition** *Suppose $p_\tau = f_\tau$ for all dates $\tau$. Then the sum of the expected utility of the*

[13] Lacour-Little and Yang (2008) cite a spread of 25 basis points on lender pricing sheets. Guttentag looks at wholesale mortgage prices in 2006 and finds a spread of 37.5–100 bp. See www.mtgprofessor.com/A%20-%20Interest%20Only/how_much_more_does_interest-only_cost.htm.

[14] Another reason is that the agents who purchase securities are themselves risk-shifting, and thus willing to buy assets with negative returns. Landier, Sraer, and Thesmar (2011) argue that lenders gambling for resurrection would have had incentives to hold backloaded mortgages, including IOs.

*borrower and lender is constant, so if a mortgage contract makes one party better off relative to some benchmark contract, it must make the other party worse off.*

Note that the proposition does not tell us what mortgages we should observe when there is no bubble. When housing supply is unconstrained, so $p_\tau = f_\tau = d$ for all $\tau$, which contracts we observe is indeterminate. Recall that in this case, loans are fully collateralized and the equilibrium interest is equal to the risk free rate $\beta^{-1} - 1$. Hence, borrowers who take out an IO mortgage can save during the interest-only phase to meet all higher payments from date $T_0 + 1$ on. Agents are thus indifferent between IOs and fixed rate contracts.

When housing supply is constrained and there is no bubble, the model predicts all borrowers will take the fixed-rate mortgage. The reason is that high types will prefer these contracts to an IO contract they would force them to sell their house. Lenders will therefore offer fixed-rate contracts in equilibrium. As for low types, either lenders will prefer to offer them fixed rates contracts, so only fixed rates contracts will be offered, or they prefer to offer them IOs, but by the proposition low types must prefer fixed-rate contracts and would choose those. Either way, only fixed rate contracts will ever be observed. Of course, our model ignores other frictions that may make IOs preferable to both borrower and lenders, such as liquidity constraints combined with fast income growth. But we can try to control for these alternatives in our empirical implementation.

To summarize, to the extent that we can control for city characteristics that encourage the use of backloaded mortgages for reasons not in our model, we should observe these mortgages in cities with booms and busts in house prices if the latter are associated with a speculative bubble, but not otherwise. Since our model suggests lenders offer backloaded mortgages because they encourage borrowers to repay their debt more quickly, we should observe that IO mortgages are more likely to be paid off early, and more likely to default if prices collapse. The remainder of the paper investigates these predictions.

# 3 Data

This section describes the data on house prices, mortgages, and controls for city characteristics that we use to explore the implications of our model. More details are in Appendix B.

## 3.1 House Prices

For house prices, we use house price indices for all cities reported by the Federal Housing Finance Agency (FHFA), previously known as OFHEO. Since we are interested in real house prices, we deflated the FHFA index for each city by the national Consumer Price Index. The FHFA house price index is compiled quarterly from house prices for mortgages purchased or securitized by Fannie Mae and Freddie Mac. The FHFA house price index has several advantages. First, it is a repeat-sales index based on the change in price for the same home over time. This makes it robust to changes in the composition of houses sold over time. Second, the index tracks a large number of cities over a long time period. Third, the FHFA index is publicly available and widely used. However, the FHFA index has some well-known shortcomings. For example, it excludes homes that were financed with non-conforming mortgages such as jumbo or subprime loans. Other price indices, such as the Case-Shiller index, the Zillow Home Value Index, and the CoreLogic House Price Index, do include such homes. The Case-Shiller index is only publicly available for 20 cities, and we confirmed that for these cities, the rate of price appreciation during the boom phase was similar to the FHFA index. We also confirmed that our results hold for both the Zillow Home Value Index and the CoreLogic House Price Index where the samples are larger.[15]

To capture house price appreciation in each city with a single statistic, we first identified the peak real price between 2000q1 and 2008q4 for each city. We then computed the maximum 4-quarter log real price growth between 2000q1 and the city-specific peak. That is, we summarize the rate of price appreciation during the boom for each city as the fastest rate at which real house prices grew within a 4-quarter window. We also considered the average price appreciation between 2000q1 and the peak, and below we discuss the implications for our results of using this measure. The reason we prefer the maximum 4-quarter growth rate is that it emphasizes especially rapid house price growth concentrated over a short time period. That is, given two cities with the same average growth rate, this measure ranks a city higher if house prices grow slowly at first but then surge. Maximal 4-quarter price growth seems to better identify those cities that are often singled out for the boom-bust cycle they experienced. For example, the two cities with the highest maximal 4-quarter price growth

[15] We explored two other issues concerning the FHFA index. First, the index relies on both market transactions and appraised home values from refinances. The FHFA also reports a purchase-only house price index for 25 cities based solely on transaction prices. Our price growth measures based on the two indices are almost perfectly correlated for the cities where both are reported. Second, the FHFA index uses a simple average of house price appreciation rather than weighting by house value. We therefore looked at the Conventional Mortgage House Price Index, which is essentially a value-weighted version of the FHFA index. Again, we found that our price growth measures were nearly identical to those based on the FHFA index. This is consistent with the fact that we obtain similar results for the CoreLogic house price index, which is only based on transactions prices and is value-weighted.

are Las Vegas and Phoenix, respectively, yet these cities rank only 53rd and 57th in terms of their average price appreciation from 2000 to their peak, respectively.

## 3.2 Mortgages

For mortgage data, we use the Lender Processing Services (LPS) Applied Analytics dataset, previously known as the McDash dataset. The data consists of information on mortgages from the servicers who process mortgage payments. LPS includes data from 9 out of the 10 top mortgage servicers, and covers 60% of the mortgage market by value.[16] However, the dataset is not meant to be a representative sample. Indeed, it underrepresents mortgages held by banks on their portfolios, since smaller and mid-size banks often service their own loans and do not report to LPS. The dataset also appears to undersample subprime mortgages, which again tend not to be serviced by those firms that report to LPS. However, since our analysis uses variation across cities, this should only compromise our analysis if selection varies systematically across cities. Still, we verify that our results are robust to measures of subprime shares compiled from completely different sources that we discuss below.

Given that our model suggests speculation would tend to favor backloaded contracts, we need a measure of how pervasive such contracts were in each city. For this, we first need to take a stand on what mortgages should count as backloaded. There are several mortgage products that involve unambiguously backloaded payments. One such mortgage is the graduated payment mortgage, first introduced in the 1970s. As suggested by its name, this mortgage offered payments that gradually increased over the duration of the loan, often during the first five years. However, these mortgages were rarely used in the period we look at. Similarly, balloon mortgages, where the stipulated payment jumps up at some preset date, were not very common during this period.[17] Another backloaded contract is the IO mortgage we emphasize in our model, in which the borrower only pays interest for some specified period and then repays both principal and interest. These mortgages were used much more extensively during the period, at least in certain cities. Finally, a popular but less widely used product is the option-ARM mortgage, which for some initial period gives the borrower the option to pay both principal and interest, interest only, or, to a limited extent, less than the required interest. Table 1 reports some characteristics for IO and option-ARM mortgages, as well as traditional fixed-rate and adjustable-rate mortgages for comparison.

[16] This estimate of the coverage is reported in Foote, Gerardi, Goette, and Willen (2010).

[17] Note that while the balloon option forces the borrower to make a large payment, our analysis suggests the borrower needs to be rewarded for restricting himself this way. While balloon mortgages charge slightly lower rates, in practice this may not have been enough to draw speculators to such contracts.

Although both IO and option-ARM use backloaded payments, we focus on IOs in our empirical work. We do this for two reasons. First, LPS only began to identify mortgages as IO or option-ARM from 2005 on. Mortgages that originated and terminated before January 2005 are not classified. However, we can still detect IO mortgages using the scheduled payment, since the scheduled payment exactly equals the interest rate times the loan amount for IOs. Unfortunately, there is no analogous way to identify option-ARMs, since the scheduled payment can reflect any of the options available to the borrower. Thus, we trust the time series on IOs more than we do the series on option-ARMs.[18] Second, as suggested by Table 1, the two types of mortgages apparently served different purposes. In particular, option-ARMs were associated with high rates of prepayment penalties. This is inconsistent with the notion that backloaded contracts were designed to induce early repayment of the loan.[19] By contrast, the fraction of IOs with prepayment penalties is only a little higher than the fraction for all mortgages. Thus, IOs seem closer to the type of backloaded contract that would be mutually preferred according to the model.[20] However, we tried using option-ARM mortgages and both types of mortgages combined, and the results were similar.

To capture the use of backloaded mortgages, we use the share of IOs in all first-lien mortgages for purchase (as opposed to refinancing). We also considered the share of IOs weighted by loan size, but this ratio proved similar to the unweighted share.[21] When we need to summarize the use of IOs with a single statistic in our cross-sectional analysis, we use the highest share of IOs in each city in any quarter over the sample period.

We constructed analogous statistics for other relevant mortgage characteristics. Specifically, the shares of first-lien for-purchase mortgages with a 30-year hybrid repayment (so-

[18] In private correspondence, Paul Willen pointed out that option-ARM mortgages before 2003 were largely held in portfolio because lenders liked that these mortgages were readjusted monthly, providing a better hedge against interest-rate movements than conventional ARM mortgages. Since loans held in portfolio are underrepresented in LPS, the data is likely to misrepresent the time series pattern for these mortgages. Willen's points are mirrored in press releases from Golden West Financial Corp, one of the leading issuers of option-ARM mortgages prior to 2003. See, for example, the note "History of the Option ARM" at http://www.goldenwestworld.com/wp-content/uploads/history-of-the-option-arm-and-structural-features-of-the-gw-option-arm3.pdf

[19] Prepayment penalties come in two varieties; hard penalties, which penalize any early repayment, and soft penalties, which waive the penalty if the house is sold. Anecdotal evidence suggests that penalties on option-ARMs were increasingly shifted towards the soft variety, *i.e.* lenders were allowing more borrowers to sell the asset without penalty. Still, our model suggests lenders would want speculators to refinance.

[20] Amromin, Huang, and Sialm (2010) explore the choice between IOs, option-ARMs, and other mortgages, and find that option-ARMs also appealed to a different population than IOs. They also provide additional details on what type of borrowers selected IOs that are consistent with what we find, *i.e.* these products appealed mostly to prime, high-income borrowers.

[21] This might seem to contradict the fact that average IO loan in Table 1 is larger than the average loan across all mortgages. But recall that IOs were more common in relatively expensive cities. Within cities, IOs do not appear to systematically involve either larger or smaller loans.

called 2/28 and 3/27 mortgages with a fixed rate for 2-3 years followed by an adjustable rate), the share of mortgages with a term of 30 years or more, the share of subprime mortgages, the share of mortgages reported as privately securitized one year after origination, and the share of mortgages by non-occupant investors.

Finally, we constructed a measure of the degree of leverage of all mortgages in a city. For this we consider the combined loan-to-value (CLTV) of *all* loans against a given property. Unfortunately, LPS does not match liens taken against the same property. We therefore turned to the LoanPerformance ABS database on non-prime privately securitized mortgages. This data is reported by trustees of privately securitized mortgage pools rather than servicers, and does report the CLTV for each loan. Thus, we have data on total leverage for each city, but only for non-prime mortgages, which includes a mix of Alt-A mortgages for borrowers with high credit scores and subprime mortgages for low-quality borrowers. Following Lamont and Stein (1999), we look at the share of mortgages in each city with a CLTV exceeding 80%. Our summary measure for CLTVs is the average share rather than the maximum share we use for the other mortgage characteristics.[22] Table 2 reports descriptive statistics for how all these summary statistics are distributed across the cities in our sample.

### 3.3 Other Data

Lastly, we compiled data used in previous studies to explain house price appreciation across cities, *e.g.* Case and Shiller (2003), Himmelberg, Mayer, and Sinai (2005), and Glaeser et al. (2008). Our variables include real per capita income, unemployment, population, property tax rates, the share of undevelopable land due to water or steep land terrain as compiled by Saiz (2010), and the Wharton Residential Land Use Index from Gyourko et al. (2008).

## 4 Cross-Sectional Evidence

In this section, we lay out the evidence on the relationship between the use of IOs and house price changes in the cross-section of cities. Based on our model, we expect that if rapid price appreciation were associated with a speculative bubble, it should coincide with the use of IO mortgages, and these should only appear in cities with inelastic housing supply. Our aim is to establish whether IO usage is correlated with large house price appreciation and depreciation. We lay no claims to estimating any causal relationships.

[22]The reason is that this series is based on nonprime mortgages, a relatively small market in the beginning of our sample. The maximum share is thus particularly prone to outliers.

## 4.1 House Prices and Housing Supply

To distinguish cities that are and are not vulnerable to speculation, we identify cities that rank in the bottom half of all cities in terms of the share of undevelopable land and the Wharton regulation index as cities with relatively elastic supply. Similarly, we identify cities in the top half of both measures as cities with relatively inelastic supply. Figure 2 plots our preferred measure of price appreciation against the maximal share of IOs in each city. As evident from the top panel, cities with few restrictions on supply exhibit low rates of house price appreciation. This was already demonstrated in Glaeser et al. (2008). However, Figure 2 also shows that these cities tended to forgo IOs. For cities where supply is restricted, there is wide variation in both house price appreciation and the use of IOs, as evident from the bottom panel of the figure, and house price appreciation is strongly correlated with the use of IO contracts. Essentially, IOs are only used in cities where housing supply is inelastic, and then only where house prices grew especially fast.

## 4.2 Baseline Estimates

For a more rigorous analysis, we regressed the maximal rate of house price appreciation on the maximum share of IOs, using data for all of the cities in our sample, with controls for various city characteristics, including those related to supply elasticity. We report our results in Table 3. The first column shows that the positive relationship between house price appreciation and IO contracts remains when we expand our sample to all cities. The coefficient on the share of IOs is statistically significant at the 1% level. To help interpret the coefficient of .416, note that the difference between the largest and smallest maximal IO share in our data is equal to $.609 - .017 = .592$. Multiplying this by .416 implies that the maximum 4-quarter growth rate in the city with the largest share of IO mortgages should exceed the rate with the smallest share by exp(.246)=27.9%. This is comparable to the difference in peak growth rates between Phoenix (36%) and Laredo (7.8%) in Figure 1.

Of course, some of the variation in house price appreciation across cities might be due to differences in other factors that help determine the value of housing services in a city. The second column in Table 3 omits the maximal IO share and includes these factors, both in levels and in annualized changes, from the beginning of our sample and the peak date in each city. The change in population growth, unemployment, and property tax rates all enter significantly with the expected signs, as do the two supply variables we use. Interestingly, the $R^2$ for these variables combined is not much larger than for the share of IOs by itself. In the third column of Table 3, we use these variables as controls when looking at the relationship between house price appreciation and the share of IOs. The coefficient on the share of IOs is smaller than in the first column, although we cannot reject that the two coefficients are equal at the 5% level. More importantly, the coefficient remains tightly estimated and significantly different from zero. Hence, the share of IOs is significantly related to the residual variation in house price appreciation that cannot be explained by the set of covariates typically used to explain house price appreciation. Note that accounting for the share of IOs renders the two supply variables we use statistically insignificant. That is, once we know which cities relied on IOs, additional information on the elasticity of housing supply does not help to better explain which cities experienced house price booms. In the last column, we add state fixed effects so that our identification relies on variation from cities in the same state. The coefficient falls to .225, but remains highly statistically significant.[23]

## 4.3 Controlling for Other Mortgage Characteristics

We now argue that the predictive power of IO mortgages for house price appreciation does not reflect some other characteristic of these mortgages. We consider several alternative mortgage characteristics that the IO feature may be a proxy for. The first two characteristics, the share of hybrid mortgages and the share of mortgages with a term of at least 30 years, are alternative affordability products. The next three correspond to various explanations that have been proposed for the boom-bust cycle in the housing market, namely the share of subprime borrowers, the share of mortgages that were securitized soon after origination, and the share of highly leveraged mortgages (with a CLTV of at least 80%). Lastly, we consider the share of mortgages taken out against investment properties, following Robinson and Todd (2010) who argue that the share of mortgages taken by non-occupants may have affected house price appreciation.

The effect of including additional mortgage variables can be seen in Table 4. As evident from the first row of the table, adding any one variable by itself has little impact on the coefficient on the IO share. This confirms that IOs are not merely a proxy for one of these other mortgage characteristics. When we add all of these variables, the coefficient on the share of IOs falls, but remains statistically significant at the 1% level. Moreover, since the standard error on IO share doubles when we combine all of these alternative measures, we cannot reject that the coefficient is the same as in our benchmark specification.

[23] We also added as a control the city-specific house price growth between 1985 and 1989, the last time there was a boom-bust cycle in national house prices. This variable may capture omitted variables that indicate a propensity towards boom-bust patterns. When we omit the share of IOs, this variable is indeed significant. But once we include the share of IOs, price growth during this last period turns insignificant, and the coefficient on the maximal IO share is nearly identical to that in Table 3.

Neither share of the two types of affordable mortgages we consider - hybrids and long term mortgages - is significant at the 5% level. However, both variables are highly significant when we include them individually but omit the share of IOs. That is, borrowers were more likely to take out hybrid and long-term mortgages in cities where house prices grew rapidly. But the use of IOs is a better indicator of whether a city experienced particularly high price appreciation, and once we control for it, the extent to which these additional types of contracts were used provides no additional information on the rate of house price growth.

By contrast, the share of subprime mortgages is statistically significant when we add it as an explanatory variable, but its sign is the opposite of what we might expect: cities with more subprime mortgages have *lower* house price appreciation.[24] However, this result is not robust; the share is not statistically significant when we include all of our mortgage characteristics in the final column of Table 4. The negative coefficient reflects the prevalence of subprime mortgages in lower income cities, while rapid house price appreciation was mostly concentrated in medium and high income cities. However, this does not mean that the expansion of the subprime market was unimportant, either for house prices or for the housing market more generally. Mian and Sufi (2009) find that within cities house price appreciation tended to be concentrated in poorer areas where more low quality borrowers resided in 2000. In addition, the rise in subprime lending appears to have played an important role in the rise in home ownership and the subsequent foreclosure crisis, which is just as important of a concern for policymakers as the boom-bust cycle in housing prices.[25]

The share of mortgages privately securitized one year after origination is not statistically significant once we control for the share of IOs. Again, this variable is significant when we omit the IO share, suggesting cities with higher appreciation relied more on securitization. IOs are simply a better indicator of whether a city experienced high price appreciation.

The average share of mortgages with a CLTV of over 80% enters significantly, but has a surprising sign: cities where borrowers were more leveraged experienced *less* house price appreciation. One potential explanation is that in cities with house price speculation, lenders knew to protect themselves by insisting borrowers own a greater equity stake. Regardless, controlling for CLTV has little effect on the coefficient on IOs.

[24] Since subprime mortgages are underrepresented in LPS, we also considered alternative measures of subprime mortgages. First, we used the ratio of all mortgages included in subprime mortgage pools as reported by LoanPeformance to the total number of mortgages for each city as reported in the Home Mortgage Disclosure Act (HMDA) data. Since the vast majority of subprime mortgages were securitized, this should be a better measure of the true share of subprime mortgages in each market. We also looked at the share of loans issued by known subprime lenders as classified by HMDA. Both measures imply a negative coefficient on the share of subprime mortgages, although in neither case is it significant.

[25] The role of subprime mortgages for these patterns are explored in Chambers, Garriga, and Schlagenhauf (2009) and Corbae and Quintin (2009), respectively.

Lastly, the share of mortgages taken out by investors does enter significantly, both by itself and when we control for all other mortgage characteristics. However, the share of investor mortgages is essentially orthogonal to the share of IOs, as evidenced by the fact that adding it has no effect on the coefficient on IO share.

All of the regressions in Table 4 use the maximum 4-quarter price appreciation as the dependent variable. As we noted above, we also considered average house price appreciation between 2000q1 and each city's peak price. Most of the results are similar. The share of IOs are statistically significant at the 1% level when we use this as our left-hand variable. Adding various controls did not knock this variable out, although when we include all of the alternative mortgage characteristics that we use in the last column of Table 4, the share of IOs only comes in significant at the 5% level rather than the 1% level. Still, the statistic we use to summarize prices matters. For example, the share of privately securitized mortgages is significant at the 5% level for this measure, both by itself and when we control for all other characteristics, which is different from what we find in Table 4.

## 4.4 The Affordability Hypothesis

Although we argue in the previous section that IOs do not seem to be a proxy for some other mortgage characteristic, one might still worry that the correlation we identify arises for reasons that have nothing to do with speculation and that we ignore in our model. For example, cities with faster house price appreciation may have faster income growth, and individuals who expect their income to grow may prefer backloaded mortgages for liquidity reasons even when assets are priced at their fundamental value. The fact that the correlation survives when we control for growth in per capita income in Table 3 should discount this particular explanation.[26] But an even more mundane explanation is that as houses become expensive, borrowers resort to more affordable mortgage products. IOs have the advantage that they offer very low payments during the IO period, and so they offer at least temporary affordability. By this view, the use of IOs simply mirrors the rapid appreciation of housing prices rather than offering evidence of speculation. We check this by testing whether adding the level of house prices at the peak drives out the share of IOs as a predictor of appreciation. To do this, we took data on the median price of single family homes for each city from the National Association of Realtors in 2000q1, and then used the rate of real price appreciation

[26] We also considered two related variables - the variance of the change in real log per-capita income between 1969 and 2000, and a vector of employment shares in 2000 for 8 industry categories (agriculture and mining, construction, manufacturing, transportation and utilities, trade, finance and real-estate, services, and government). In both cases the coefficient on IOs remains significant.

in the FHFA to compute the implied price of the same home at the peak.

The first two columns in Table 5 show the effect of adding the log of the price level at the peak when we continue to control for the same explanatory variables as in Table 4. When we include the log peak price by itself, this variable has a positive and statistically significant effect, confirming that places with greater appreciation were also more expensive. However, when we add the share of IOs, the coefficient on log peak price becomes statistically insignificant, while the coefficient on the share of IOs remains highly significant and not statistically different from the estimate we report in Table 3. We also considered the ratio of the peak price to per-capita income in the year of the peak. These estimates are reported in the third and fourth columns of Table 5, and lead to the same conclusion.

## 4.5 Price Declines

So far, our analysis has focused on house price appreciation. However, the model suggests that the use of IOs should similarly be concentrated in the cities with big house price declines if the bust phase was associated with the bursting of a bubble. To investigate this, we constructed a measure of house price declines analogously to the way we measure price appreciation before the peak. That is, for each city we measure the largest 4-quarter decline between the city-specific peak and the end of our sample in 2008q4. In 43 cities, the highest price recorded occurred in 2008, so the period of decline was not long enough to compute a 4-quarter growth rate. In 31 of these cities, the peak price was in 2008q4, implying there was never a bust phase in house prices. So our sample is necessarily smaller. Among the cities in which there was a bust, in 85% of the cases the largest 4-quarter price decline occurred between 2007q3 and 2008q3, even though these cities peaked at different dates, some as early as 2003. The collapse in house prices was thus highly synchronized. In what follows, we adopt the convention of using the negative of the price change.

We report our key results in Table 6. Again, IOs appear to be concentrated in cities where prices fell by a larger amount. Moreover, this result is robust to controlling for the level and changes in population, unemployment, per capita income, and property tax rates in the period *after* the peak, as well as to including other mortgage characteristics.

The main difference between the results for price declines and price increases does not involve the results for IOs, but for other mortgage characteristics; variables that were not significant for explaining house price appreciation do appear to be important in explaining house price declines, although they do not knock out or even significantly change the coefficient on the IO share. We suspect this is because these other variables help to predict the excess foreclosure rate in each city, *i.e.* the excess of foreclosures beyond what one would predict based on the share of IO contracts alone. As documented in Campbell, Giglio, and Pathak (2009), foreclosures are likely to drive house prices down, both because foreclosed properties sell at a steep discount and because they drive down the value of neighboring properties. Some of the differences in house price declines between otherwise similar cities do seem related to differences in foreclosures. For example, the IO share in Washington DC was 47%, compared to 49% in Stockton, and both experienced relatively similar rates of house price appreciation, with a maximum 4-quarter growth of 22.5% and 28.2%, respectively. Yet house prices fell only 17.3% in Washington DC but 42.3% in Stockton in one 4-quarter period after price peak. The difference in foreclosure rates between these two cities is equally striking: Stockton had a foreclosure rate of 9.5% in 2008 according to RealtyTrac, a firm which tracks foreclosure rates from public records and court notices, while Washington DC had a foreclosure rate of only 3.0%.

To explore this conjecture, we constructed a measure of unanticipated foreclosures in each city by regressing the maximum share of the mortgages in LPS that enter foreclosure for each city on the maximum share of IOs. The residual from this regression represents the rate of foreclosure that cannot be predicted by the use of IOs. When we add this residual to the list of controls, it comes in highly significant and knocks out all of the other mortgage characteristics we consider. This suggests that the reason various mortgage characteristics explain price declines but not price increases is that they can help to predict which cities subsequently experienced unusually high foreclosure rates.[27]

# 5 Time Series Evidence

The cross-sectional evidence in the previous section suggests the use of backloaded mortgages was concentrated in cities with large house price swings. However, a lingering concern is that this result occurs mechanically because borrowers require backloaded mortgages for affordability reasons when houses become expensive. We have already provided some cross-sectional evidence against this interpretation. In this section, we offer evidence based on time series information for our panel of cities. In particular, we examine whether the rise in use of IOs appears to be a response to houses becoming more expensive. Formally, we look at whether house price appreciation Granger-causes the increased use of IOs. We find it does not, and that if anything, house price appreciation leads to a decline in the use of

[27]As a robustness check, we looked whether the residual foreclosure rate predicts price growth *before* the peak, which it wouldn't necessarily under our argument. When we add it as a variable, it does come in significant, but only at the 5% level, and turns insignificant when we control for other mortgage characteristics.

IOs. Indeed, this can be seen in the case of Phoenix in Figure 1. Our findings are thus at odds with affordability as the reason for why IOs were common in cities where house prices grew especially quickly.

We also look at whether the use of IOs Granger-causes house price appreciation. This is consistent with our model, although not an inherent implication of the model. More precisely, while speculators take out IO loans in anticipation of possible future house price appreciation, the share of IOs depends on the fraction of speculators among all borrowers, or $\phi$ using the notation of the model. In general, our model imposes no restriction on how $\phi$ evolves over time, and so there is nothing that requires it must anticipate price appreciation.

## 5.1 Construction of Panel Data

In testing whether price appreciation Granger causes the use of IOs, we do not want to look at all cities. Recall that in cities where prices did not appreciate much during this period, the share of IOs was negligible. We would not want these cities to drive our finding that price appreciation does not anticipate the use of IOs. We therefore restrict attention to a subset of cities in which the peak share of IO mortgages was large. Figure 3 plots the contemporaneous correlation between the change in IO use and the change in price, the autocorrelation of quarterly price changes, and the autocorrelation of quarterly changes in IO shares, for each city, in each case plotted against the maximal IO share recorded in the city. The figure suggests that cities with a maximal IO share of 40% or more are relatively homogeneous, and certainly different from cities below this cutoff in terms of both the dynamics of prices and IO usage and the way in which the two move together. For this reason, we consider cities with a maximum IO share that is at least 40%. This leaves us with a sample of 29 cities. We confirm that our results remain unchanged for lower cutoffs up to 30%.

We also focus attention on the period of rising house prices, since we are interested in whether borrowers shifted to IOs when houses became more expensive. Evidence of a shift away from IOs when prices fell may be unrelated to affordability, especially since the market for non-traditional mortgages froze once house prices fell. We therefore only consider data up to 2006q4, when national FHFA real house prices peaked. We confirmed that our results are robust to ending the sample in any of the other quarters of 2006.

We first offer some simple correlations which show the dynamics of IO use and house price appreciation across the 29 cities in our sample. Figure 4 displays dynamic correlations between changes in the share of IOs at date $t + j$, $\Delta io_{t+j}$, with log changes in real house prices at date $t$, $\Delta p_t$, for $j = -4, -3, \ldots, 4$. The correlations are positive for $j \leq 2$, peaking at $j = -2$ and negative for $j > 2$. In the language of time-series analysis, this pattern of correlations indicates that increased use of IOs *leads* house price appreciation.

## 5.2 Granger-Causality

We now present our formal Granger-causality tests. We estimate the two equations

$$\Delta p_{it} = \alpha_p + \beta_p(L)\Delta p_{it-1} + \gamma_p(L)\Delta io_{it-1} + \varepsilon_{it}^{p} \tag{4}$$

$$\Delta io_{it} = \alpha_{io} + \beta_{io}(L)\Delta io_{it-1} + \gamma_{io}(L)\Delta p_{it-1} + \varepsilon_{it}^{io} \tag{5}$$

where $\Delta io_{it}$ and $\Delta p_{it}$ refer to city $i$ at date $t$ and the $\beta(L)$ and $\gamma(L)$ functions are lag polynomials. For simplicity, we focus on models where the number of included lags is the same for both $\beta(L)$ and $\gamma(L)$. Table 7 and Table 8 report our results based on two different ways of estimating this model, depending on whether we assume (4) and (5) are homogeneous across the different cities in our sample or not.

### 5.2.1 OLS Estimates

If the coefficients $\alpha$, $\beta(L)$, and $\gamma(L)$ in (4) and (5) are the same for all cities, we can estimate these equations using ordinarily least squares. These estimates are reported in Table 7. The first four columns show that for all lag specifications, an increase in the share of IOs does Granger-cause house price appreciation in the period we consider. More precisely, as evident from the row labeled with summations, we reject the hypothesis that the sum of the coefficients on the share of IOs is zero. In addition, the null hypothesis that all the coefficients on IOs are zero is rejected for all four specifications. The F-statistics associated with this hypothesis are in the row labeled "F-stat" with the associated p-values below.

In the opposite direction, with only one lag, we cannot reject the hypothesis that price growth does not Granger-cause increased IO use. When we allow for two and three lags, respectively, house price appreciation does appear to Granger-cause future use of IOs. However, the sum of the coefficients is negative for these specifications, implying a rise in prices causes a decline in the use of IOs. When we allow for four lags, the sum of the coefficients remains negative but is no longer significant at the 5% level. In none of the specifications is there any support for the notion that an increase in house prices leads the subsequent use of IOs.

The OLS estimates are valid only if the estimated residuals are serially uncorrelated. Otherwise the estimates are inconsistent. Table 7 reports p-values for Arellano and Bond

(1991) tests of the null hypothesis that the residuals exhibit no serial correlation of order one through four, $AR(j), j = 1, 2, 3, 4$. Using conventional significance levels, these tests indicate that it is not possible to reject serial correlation for some $j \geq 2$ for all lag specifications. That is, there is always some value at which the p-value on the hypothesis that the errors are serially correlated falls below 5%. Serial correlation remains even when we increase the number of lags to six, calling into question the use of OLS.

### 5.2.2 GMM Estimates

One potential source of serial correlation is if the intercept terms $\alpha$ in (4) and (5) vary across cities. To accommodate city-specific constants, that is "fixed effects," with homogeneous slope coefficients, we use the System-GMM estimator developed by Arellano and Bond (1991), Arellano and Bover (1995), Blundell and Bond (1998) and Holtz-Eakin, Newey, and Rosen (1988). System-GMM involves estimating (4) or (5) with a system of two equations. The first equation of the system involves differencing the original equation to remove the fixed effects, and using lagged values of the variables in the original estimation equation as instruments. The second is the original estimation equation using differences of the variables in the original estimation equation as instruments. In the latter case we assume that the differences are orthogonal to the fixed effects. Our GMM estimates, reported in Table 8, are based on using the third and fourth lags as instruments.[28] The validity of our instruments with GMM depends on the lack of serial correlation in the estimation errors for the differenced equation of order three and higher. The Arellano and Bond tests of serial correlation now indicate that four lags are necessary for this condition to be satisfied when forecasting house price appreciation, but two or more lags satisfy the serial correlation criterion when forecasting changes in the use of IOs.[29] Allowing for fixed effects thus appears to resolve the problem of serial correlation.

Turning to the results, the conclusions are similar to what we found using OLS. The

[28] These estimates were computed using `xtabond2` for `STATA`, described in Roodman (2009). The standard errors are robust to arbitrary patterns of autocorrelation within cities (clustering) and include the small sample correction developed in Windmeijer (2005). We follow the convention of including orthogonality conditions that are valid at each date. That is, the expectations are evaluated over cities at each date. Consequently, even with just two lags as instruments the number of orthogonality conditions is quite large. While we fail to reject the $J$-test of the over-identifying restrictions with a $p$-value of unity in all cases, the actual values of the test statistic are relatively small. This indicates that the non-rejection of the over-identifying restrictions is not driven by noise in the data. Evaluating expectations over cities and dates rather than over cities at each date separately dramatically reduces the number of orthogonality conditions. Our findings are robust to using a smaller number of orthogonality conditions based on this latter approach. Our findings are also robust to using a single lag and including a third lag as an instrument.

[29] While we report results for various lag lengths, we explored using the lag selection criterion for dynamic panel data models in Andrews and Lu (2001). For both (4 and (5), this criterion favored a one lag model.

first four columns of Table 8 continue to imply that for all lag specifications, an increase in IOs still Granger-causes house price appreciation in the period we consider. The next four columns show that there is little evidence that that house price appreciation Granger-causes the use of IOs, for all lags. With only one lag, the coefficient is essentially zero. With more lags, the sum of the coefficients on lagged prices is negative as with OLS, but is no longer significant at the 5% level. The use of IOs thus appears to arise in anticipation of future appreciation and not in response to past appreciation.[30]

# 6 Mortgage Pre-payment and Default

Our theory implies that lenders prefer IOs over traditional mortgages because they encourage speculators to re-pay their mortgages sooner rather than later. A natural check on our theory, then, is whether IOs indeed pre-paid at a faster rate than other kinds of mortgages. Of course, there may be other reasons for this pattern, *e.g.* borrowers who know they intend to move within a few years may select into IOs. Still, if we were to find that IOs are pre-paid at a slower rate than other mortgages, it would be evidence against our theory. Our model further predicts that when prices collapse, holders of IO mortgages will default rather than pre-pay. This suggests looking at foreclosure rates for different types of mortgages.

We focus on the same 29 cities with high IO shares that we considered in our time-series analysis in Section 5. Once again, we want to make sure that our results are not driven by other cities, since the prediction really applies to these cities. However, we verified that we find similar results when we look at all cities. To minimize the impact of composition and time effects, we separate mortgages depending on when they originated. In what follows, we report results for mortgages originating in 2005q1, although mortgages originating at any quarter of 2004 and 2005 look similar. The results are different for 2003, but this may be because few IOs originated that year and so the estimates are likely to be noisy.

We estimate pre-payment and foreclosure rates using the Kaplan-Meier estimator. More precisely, denote the times we can track each mortgage by $t_i$, $i = 1, 2, \ldots, N$ and $N$ is the total number of mortgages. Corresponding to each value for $t_i$ is the number of mortgages that we could have observed pre-paying (foreclosing) at this date, namely all the mortgages that survived to date $t_{i-1}$ and for which data is still available at date $t_i$. Denote this number by $n_i$. Let $d_i$ denote the number of mortgages that pre-paid (foreclosed) at date $t_i$. The

[30]As a robustness check, we re-ran our estimates including lags of log GDP growth and the Federal Funds interest rate, and our findings remain unchanged.

Kaplan-Meier estimator for pre-payment (foreclosure) is given by

$$\hat{S}(t) = \prod_{t_i \leq t} \frac{n_i - d_i}{n_i}$$

That is, $\hat{S}(t)$ denotes the fraction of mortgages that have not prepaid (foreclosed) after $t$ periods. We report the statistic $1 - \hat{S}(t)$, which corresponds to the fraction of mortgages that pre-paid (foreclosed) by date $t$. Note that the Kaplan-Meier estimator addresses right-censoring, which is important since pre-payment right censors the propensity for mortgages to default, and default right censors the propensity for mortgages to pre-pay. Given the large number of either type mortgages in the 29 cities that originate in 2005q1, we omit standard errors in our reporting.

Figure 5 displays estimated pre-payment and foreclosure rates for two types of mortgage, IO and non-backloaded. By "non-backloaded" we mean all mortgages that are neither IOs nor Option-ARMs. We confirm that IOs were indeed more likely to pre-pay than non-backloaded mortgages. To get some sense of magnitudes, we estimate that after 8 quarters, 40 percent of IOs have pre-paid while less than 30 percent of non-backloaded mortgages are pre-paid.

Turning to defaults, the fraction of mortgages that enter foreclosure is the same for both IOs and non-backloaded mortgages until 2007q1, the quarter after the aggregate FHFA house price index begins to fall. At this date, the propensity of IOs to enter foreclosure begins to rise faster than for non-backloaded mortgages, in line with the model's prediction that speculators who would have chosen IOs default when prices begin to fall. Comparing pre-payment and foreclosures reveals that when the gap in foreclosure rates between IOs and non-backloaded mortgages opens up, the tendency for IOs to pre-pay early reverses. Although we do not report the results, we did confirm that the survival probabilities for the two types of mortgages are statistically different using the log rank test of equality.

# 7 Conclusion

In this paper we argue that when there is a speculative bubble and lenders cannot distinguish speculators from profitable borrowers *a priori*, both lenders and speculators will prefer mortgages with backloaded payments over traditional mortgages. This insight motivated our analysis of house prices and mortgages for a sample of US cities over the period 2000-2008. Our main findings are that IO usage is a strong indicator of which cities experienced both rapid price appreciation and subsequent depreciation, that IOs do not seem to be used in response to house prices increasing, and that IOs lead price appreciation rather than vice versa. We also confirm that IOs were more likely to be pre-paid than traditional mortgages early on, but once prices fell were more likely to enter foreclosure.

At a minimum, these facts can inform the search for an explanation for the recent boom-bust pattern in house prices. That is, any persuasive theory should also explain the tendency for home buyers to rely on certain types of mortgages to finance these purchases, and why the use of such mortgages anticipated the rise in prices rather than responded to it. In addition, we offer a model that suggests our findings support the view that house price changes were at least partly driven by a speculative bubble, since in the model IOs should be observed if the boom bust was associated with a speculative bubble but not if the boom and bust reflect fundamentals.[31] To be sure, there are other potential explanations for the concentration of IOs in cities with boom-bust episodes that our model ignores, and one would have to reject these to argue that the pattern we find is indicative of a speculative bubble. We have attempted to control for some of these explanations, such as differences in income growth across cities. There may be additional explanations we have ignored. But we view exploring these alternatives as a more productive way to settle this question than trying to estimate the true fundamental value of housing and comparing it to prevailing prices.

Finally, we should emphasize that our findings do not imply that backloaded mortgages caused a bubble in housing, nor do they imply that regulators should have disallowed these contracts. In fact, our model predicts a speculative bubble would occur even if lenders could only offer traditional mortgages, and that backloaded mortgages actually keep overvaluation in check by encouraging speculators to unload the houses they bought. Our analysis also ignores positive aspects of backloaded mortgages such as their benefits for liquidity-constrained households, and these must be taken into account in formulating policy. Finally, a potential argument for allowing backloaded contracts is that they may be the "canary in the coal mine" for anticipating price movements. That said, there is nothing in our analysis that tells us the exact form of backloaded contract one should look for, and once policymakers condition their actions on the choice of contracts, this may affect the incentives for lenders and borrowers to choose these contracts in the first place.

[31] Interestingly, Minsky (1982) also argued that a telltale sign of speculation was that borrowers only cover the interest obligation on their loans. However, his argument relied on a rather different intuition that had nothing to do with the backloading of payments. Nor did Minsky explain why lenders should not be alarmed by the rise in such borrowing, other than arguing that they too might be swept up in some general euphoria.

# Appendix A Proofs of Theoretical Results

**Proof that if a bubble exists at some date $t$, then it exists at other dates.** Formally, we show that if $p_t > f_t$ for some $t$, then $p_\tau > f_\tau$ for $\tau \leq \min(t^*, \mathcal{T})$ and $p_\tau = f_\tau$ otherwise.

First, suppose $\tau > \min(t^*, \mathcal{T})$. This case is equivalent to having a constant number of buyers: Either $t > \mathcal{T}$ and no new buyers arrive, or $t > t^*$ and no houses trade, which is equivalent to no new buyers arriving. In the text, we argued that with constant buyers, either $p_t = d$ or $p_t = D$. It is easy to check that $f_t$ coincides with $p_t$ in either case.

Next, consider the case where $\tau \leq \min(t^*, \mathcal{T})$. First, consider any date $\tau < t$. Suppose $p_\tau = f_\tau$. In this case, no agent would agree to sell the asset at date $\tau$, since it is better to wait to sell the asset at date $t$, where $p_t \geq f_t$ with strict inequality in some states of the world. That is, $p_t > f_t$ implies

$$E_\tau \left[ \sum_{s=\tau}^{t} \beta^s \left(\beta^{-1} - 1\right) d + p_t \right] > E_\tau \left[ \sum_{s=\tau}^{t} \beta^s \left(\beta^{-1} - 1\right) d + f_t \right] = f_\tau = p_\tau$$

where the second to last inequality comes from (1). Hence, it is better to collect dividends until $t$ and then sell the asset than to sell it at $\tau$. So no agents sell. But if $p_\tau = f_\tau$, all new arrivals at date $\tau$ would want to to buy the asset, which cannot be an equilibrium.

Next, consider $\tau > t$. The argument is by induction. First, at date $\tau = t+1$, if $p_{t+1} = f_{t+1}$, then any agent who bought a house before date $t$ would prefer to sell the asset at date $t$ than to wait and sell at date $t+1$ if $\mathcal{T} > t+1$. So none of these agents plan to sell at date $t+1$ if $\mathcal{T} > t+1$. But if $p_{t+1} = f_{t+1}$, new buyers will wish to buy a house if $\mathcal{T} > t+1$. Supply must therefore come from low types who bought the house at date $t$. But they would demand a price above $f_{t+1}$ to sell since they are giving up the option to wait and default, so $p_{t+1} > f_{t+1}$. The same argument can be applied to later dates.

**Determining the value of a contract to both borrowers and lenders.** Let $V_\tau$ denote the expected value to a low type who still owns the house $\tau$ periods after buying it and before knowing whether new buyers will arrive that period. Under the fixed-rate mortgage, the speculator can either sell the house and pay back $(1+r) L_{\tau-1}$; pay $m_\tau$ and retain ownership of the house; or default. The payoffs to the three options are $(\beta^{-1} - 1) d + p_\tau - (1+r) L_\tau$, $(\beta^{-1} - 1) d + \beta V_{\tau+1} - m_\tau$, and 0, respectively. Let $\tau^*$ denote the number of periods between when the contract originated and $t^*$. Assuming $L_{t^*+1} > d/\beta$ so agents owe more $d$ for at least $t^*$ periods, the optimal strategy at date $\tau^*$ is to sell the asset if new buyers arrive at $\tau^*$ and default if they don't. Hence,

$$V_{\tau^*} = (1-q) \left[ \left(\beta^{-1} - 1\right) d + D - (1+r) L_{\tau^*-1} \right] \tag{6}$$

For $\tau < \tau^*$, $V_\tau$ is defined recursively as

$$V_\tau = (1-q) \max \left[ \left(\beta^{-1} - 1\right) d + p_\tau - (1+r) L_{\tau-1}, \left(\beta^{-1} - 1\right) d + \beta V_{\tau+1} - m_\tau, 0 \right] \tag{7}$$

Under the IO contract, the speculator would have to either sell or default at date $T_0 + 1$. If $T_0 + 1 < \tau^*$, the boundary condition (6) will be replaced with

$$V_{T_0+1} = (1-q) \max \left[ \left(\beta^{-1} - 1\right) d + p_{T_0+1} - (1+\widehat{r}) L_{T_0}, 0 \right] \tag{8}$$

At earlier dates, $V_\tau$ can be again defined recursively using (7) with $r = \widehat{r}$. Computing back to date $\tau = 1$ reveals which contract borrowers would prefer when they take out the loan.

We can similarly compute the revenue $\Pi_\tau$ a lender expects to earn $\tau$ periods into the loan, before knowing if buyers will arrive at date $\tau$. Since at $\tau^*$ the borrower would default if no buyers arrive and will repay his loan otherwise, we have

$$\Pi_{\tau^*} = qd/\beta + (1-q)(1+r) L_{\tau^*-1}$$

For $\tau < \tau^*$, the value $\Pi_\tau$ is given by

$$\Pi_\tau = qd/\beta + (1-q)\pi_\tau$$

where $\pi_\tau$ depends on what the borrower does. If he sells the house, $\pi_\tau = (1+r) L_{\tau-1}$. If he remains current on his payments, $\pi_\tau = m_\tau + \beta\Pi_{\tau+1}$. If he defaults, $\pi_\tau = [(\beta^{-1}-1) d + p_\tau]$. Expected profits when the loan is originated are thus $\beta\Pi_1 - L$.

**Equilibrium contracting when there is a bubble.** We now argue that if at the equilibrium interest rate on the traditional mortgage, there exists an IO mortgage that low type borrowers and lenders both prefer, then (1) both traditional and IO mortgages will be offered in equilibrium, with low-types choosing the IO contract and high types choosing the traditional mortgage; (2) IO contracts will carry a higher interest charge in equilibrium.

First, we argue that low-types must receive IO contracts in equilibrium. For suppose not, i.e. they receive a traditional mortgage contract with interest rate $r^*$. We first argue that $r^*$ must exceed $\beta^{-1} - 1$ to ensure non-negative profits. In particular, the interest rate on any loan must be at least $\beta^{-1} - 1$, or else the lender would never offer it. If the traditional mortgage involved a rate $\beta^{-1} - 1$, high types would at most repay at the risk free interest rate, so lenders will earn no profits from high types. But given expected profits from lending to low types at a rate $r^* = \beta^{-1} - 1$ are negative since they will default with positive probability, lenders will not be able to earn positive profits. Hence, $r^* > \beta^{-1} - 1$.

Now, consider a lender who offers the same set of contracts as those offered in equilibrium, but also offers an IO contract with the same interest rate $r^*$ as charged on the traditional rate in equilibrium. High types will not choose this contract, since when $r^* > \beta^{-1} - 1$, the present discounted value is lower under the traditional contract, and the IO contract would force them to give up the house at some date. Hence, these types will stick with whatever contract they were originally choosing in equilibrium. However, by assumption, both borrowers and lenders are better off under the IO contract. Since lenders earn zero profits in equilibrium, this implies a lender can earn strictly positive profits by offering a traditional mortgage together with an IO mortgage both at the same rate $r^*$ as in equilibrium. Hence, the original contracting arrangement could not have been an equilibrium.

Next, we argue that high types will end up with a traditional mortgage in equilibrium. For suppose in equilibrium all borrowers took the IO mortgage with interest rate $\widehat{r}^*$. Consider a lender who offers a traditional mortgage at rate $\widehat{r}^*$. The rate on traditional mortgages in equilibrium must have been higher than $\widehat{r}^*$, or else high types would have already chosen it, since high types prefer a traditional mortgage to an IO mortgage when at an interest rate above the risk-free rate, which must be the case in equilibrium. Since by assumption low types prefer IO over the traditional mortgage at the equilibrium rate, and since low types will find IO contracts even more attractive at lower rates, they must also prefer the IO mortgage at rate $r^*$. Thus, a lender offering a traditional mortgage with rate $\widehat{r}^*$ will not attract low types, but will attract high types. Since lending to high types at an interest rate that exceeds the discount rate yields positive profits, such a lender will earn a strictly positive profit. But then the original contracts could not have been an equilibrium.

Both contracts will therefore be offered in equilibrium. We now argue that in equilibrium, low types must be indifferent between the types of mortgages contracts in equilibrium. For suppose not, i.e. low-types strictly prefer the IO contract. Consider a lender who offers only the traditional mortgage contract offered in equilibrium, but lowers the interest rate by $\varepsilon$. Given low-types strictly prefer the IO contract, there exists an $\varepsilon$ such that they would still prefer the IO contract. High types will prefer this contract. But there exists an $\varepsilon$ small enough that the lender offering this contract and attracting only the safe borrowers will earn a strict profit.

Finally, since low types prefer the IO contract with rate $r^*$ to the traditional mortgage contract with rate $r^*$, the only way to ensure they are indifferent between the two contracts is to charge a lower rate on the traditional mortgage in order to make it more attractive. Hence, the equilibrium rate on the traditional mortgage contract will be lower than on the IO contract.

**Proof of Proposition.** We now show that if $p_t = f_t$ for all $t$, then $V_t + \Pi_t = (\beta^{-1} - 1)\, d + E_{t-1}[f_t]$ for any contract $\{m_\tau\}_{\tau=0}^{T}$.

Using the definition of $V_\tau + \Pi_\tau$ shows that if the borrower either sells the asset or defaults, this sum will equal $(\beta^{-1} - 1)\, d + E_{\tau-1}[f_\tau]$. We therefore only need to check what happens if the borrower keeps making payments. In that case, the sum of the two terms is given by

$$V_\tau + \Pi_\tau = \left(\beta^{-1} - 1\right) d + \beta\left(V_{\tau+1} + \Pi_{\tau+1}\right)$$

We will consider two separate cases, $f_\tau = d$ for all dates and $f_\tau$ given by (1). In the first case, note that at $\tau = T+1$, all debt would have been retired. Hence, $V_{T+1} = (\beta^{-1} - 1)\, d + d$ and $\Pi_{T+1} = 0$. It follows that

$$\begin{aligned} V_{T+1} + \Pi_{T+1} &= \left(\beta^{-1} - 1\right) d + d \\ &= \left(\beta^{-1} - 1\right) d + E_{T-1}[f_T] \end{aligned}$$

Next, suppose $V_s + \Pi_s = (\beta^{-1} - 1)\, d + E_{s-1}[f_s]$ for $s = \tau + 1, ..., T + 1$. Then at date $\tau$,

$$\begin{aligned} V_\tau + \Pi_\tau &= \left(\beta^{-1} - 1\right) d + \beta\left(\left(\beta^{-1} - 1\right) d + E_\tau[f_{\tau+1}]\right) \\ &= \left(\beta^{-1} - 1\right) d + \beta\left(\left(\beta^{-1} - 1\right) d + d\right) \\ &= \left(\beta^{-1} - 1\right) d + E_{\tau-1}[f_\tau] \end{aligned}$$

This establishes the claim for $f_\tau = d$.

Next, let $f_\tau = \sum_{s=\tau+1}^{\tau^*} \beta^{s-t} (\beta^{-1} - 1)\, d + \beta^{\tau^* - \tau} E_\tau[p_{\tau^*}]$, where $p_{\tau^*}$ is equal to $D$ with probability $1 - q$ and $d$ with probability $q$. At date $\tau^*$, since all uncertainty is resolved, the borrower will weakly prefer to sell the asset and strictly prefer to sell the asset if $r > \beta^{-1} - 1$. Regardless of whether the borrower defaults or sells the asset, we have

$$\begin{aligned} V_{\tau^*} + \Pi_{\tau^*} &= \left(\beta^{-1} - 1\right) d + E_{\tau^*-1}\left[p_{\tau^*}\right] \\ &= \left(\beta^{-1} - 1\right) d + E_{\tau^*-1}\left[f_{\tau^*}\right] \end{aligned}$$

Finally, suppose $V_s + \Pi_s = \left(\beta^{-1} - 1\right) d + E_{s-1}\left[f_s\right]$ for $s = \tau + 1, ..., T + 1$. Then at date $\tau$, we have

$$V_\tau + \Pi_\tau = \left(\beta^{-1} - 1\right) d + \beta E_{\tau-1}\left(\left(\beta^{-1} - 1\right) d + E_\tau\left[f_{\tau+1}\right]\right)$$

However, from the definition of $f_\tau$, we have

$$\begin{aligned} f_\tau &= \sum_{s=\tau+1}^{\tau^*} \beta^{s-\tau}\left(\beta^{-1} - 1\right) d + \beta^{\tau^*-\tau} E_\tau\left[p_{\tau^*}\right] \\ &= \beta\left(\beta^{-1} - 1\right) d + \beta\left[\sum_{s=\tau+2}^{\tau^*} \beta^{s-\tau-1}\left(\beta^{-1} - 1\right) d + \beta^{\tau^*-\tau-1} E_\tau\left[p_{\tau^*}\right]\right] \\ &= \beta E_{\tau-1}\left[\left(\beta^{-1} - 1\right) d + \sum_{s=\tau+2}^{\tau^*} \beta^{s-\tau-1}\left(\beta^{-1} - 1\right) d + \beta^{t^*-\tau-1} E_{\tau+1}\left[p_{\tau^*}\right]\right] \\ &= \beta E_{\tau-1}\left[\left(\beta^{-1} - 1\right) d + E_\tau\left[f_{\tau+1}\right]\right] \end{aligned}$$

This allows us to rewrite the sum $V_\tau + \Pi_\tau$ as

$$V_\tau + \Pi_\tau = \left(\beta^{-1} - 1\right) d + E_{\tau-1}\left[f_\tau\right]$$

which establishes the claim. ■

# Appendix B Data

This appendix provides a detailed description of our data construction.

## B1 House Price Data

Our primary data source for house prices is the Federal Housing Finance Agency (FHFA) house price index for Core-based Statistical Areas (CBSAs) as defined by the Office of Management and Budget. If a CBSA has a population greater than 2.5 million, the CBSA is divided into Metropolitan Divisions.[32] For these CBSAs, FHFA reports data for each

[32]The Metropolitan Divisions are: Boston-Cambridge-Quincy, MA-NH; Chicago-Naperville-Joliet, IL-IN-WI; Dallas-Fort Worth- Arlington, TX; Detroit-Warren-Livonia, MI; Los Angeles-Long Beach-Santa Ana, CA; Miami-Fort Lauderdale-Miami Beach, FL; New York-Northern New Jersey-Long Island, NY-NJ-PA; Philadelphia-Camden-Wilmington, PA-NJ-DE-MD; San Francisco-Oakland-Fremont, CA; Seattle-Tacoma-Bellevue, WA; and Washington-Arlington-Alexandria, DC-VA-MD-WV.

Division rather than the CBSA as a whole. We follow this convention throughout, using Metropolitan Divisions in lieu of the CBSA where applicable.

For robustness, we also used CBSA-level prices from the CoreLogic House Price Index and the Zillow Home Value Index. CoreLogic reports prices at the same CBSA level as FHFA. Zillow was available to us at the CBSA level as well, but for fewer cities. For each series, we construct our price variables as follows.

First, we convert house prices into a real index by dividing each price series by the Consumer Price Index for urban consumers as reported by the Bureau of Labor Statistics.

Due to limitations on mortgage data, we restrict our attention to the period between 2000q1 and 2008q4. For each city, we identify the quarter during this period in which the real price reaches its peak. Let $p_t$ denote the price in a given city at date $t$. The quarter in which price peaks is given by

$$t^* = \arg\max_t \{p_t\}_{t=2000q1}^{2008q4}$$

We measure real house price appreciation in each city as the highest 4-quarter growth in real house prices between $t = 1$ and $t^*$, i.e.

$$\text{Max4QGrowth} = \max \{\ln (p_t/p_{t-4})\}_{t=2001q1}^{t^*}$$

Since $t^* \geq 2001q1$ in all cities in our sample, i.e. the peak occurs at least a year after the start of our sample, our measure is defined for all cities.

In addition to the maximum rate of appreciation, we also calculate average annualized real house price appreciation prior to peak for each city as follows:

$$\text{MeanGrowth} = \frac{4}{t^* - 2000q1} \ln \frac{p_{t^*}}{p_{2000q1}}$$

Similarly, we calculate the largest 4-quarter decline following the peak analogously for the period between $t^*$ and $T$:

$$\text{Max4QDecline} = \max \{\ln (p_{t-4}/p_t)\}_{t=t^*+4}^{2008q4}$$

Since in some cities $t^* \geq T - 4$, we can only define this measure for a subset of cities.

For data on price levels, which we use in Table 5, we took the median price of single family homes sales in each CBSA in 2000q1 as reported by the National Association of Realtors. Denote this price as $P_{2000q1}^{NAR}$. We then use each CBSA's appreciation rate in the FHFA index to arrive at the price of the same median home at the peak date $t^*$. That is, the peak price in each city is given by $P_{t^*} = P_1^{NAR} \times (p_{t^*}/p_1)$.

## B2 Mortgage Data

Our mortgage data is primarily drawn from Lender Processing Services (LPS) mortgage performance data, formerly known McDash. However, we also constructed some of our variables from other datasets, and we report our construction of these variables as well.

### B2.1 LPS Mortgage Data

The LPS data is reported at a monthly frequency. From these monthly reports we construct a single "static" file that includes a single record on each loan ever observed.

An issue with LPS is that different servicers begin to report data to LPS at different dates. When a servicer joins, they report all of their outstanding loans. Mortgages that originated before the servicer reported to LPS are thus only disclosed if they survive into the reporting period. One way to avoid this survivorship bias is to only include mortgages that are reported in LPS shortly after their origination date. This approach is used, for example, by Foote et al. (2010). A problem with this resolution is that if servicers are heterogeneous and the servicers that join LPS later tend to issue different types of mortgages, the data will suffer from composition biases. We find that servicers that join later tended to issue a disproportionate share of backloaded mortgages (IOs and Option-ARMs). Ignoring the mortgages that originated before entering the sample will severely undercount the share of backloaded mortgages in early periods. For this reason, we chose to include all mortgages, regardless of whether they originate before the servicer reports to LPS. To mitigate survivorship bias, we only look at mortgages starting in 2000q1. Since backloaded mortgages are more likely to prepay (see Figure 5), survivorship bias will cause our series on the use of such mortgage to lag true usage earlier in our sample.

As our first step, we obtained counts of the total number of first-lien, for-purchase mortgages originating in each CBSA each quarter. One issue is that if a mortgage was transferred between servicers in LPS, we may end up counting the same loan twice. To avoid this, we matched all loans on zip code, origination amount, appraisal amount, interest type, subprime status, level of documentation, the identity of the private mortgage insurance provider if relevant, payment frequency, indexed interest rate, balloon payment indicator, term, indicators for VA or FHA loans, margin rate, and an indicator of whether the loan was for purchase or refinance – treating missing and unknown values as wildcards. Loans that matched were treated as duplicates, and we kept only one record in such cases. Mortgages are reported by zip code. We aggregate these codes to the CBSA level, where for zip codes that do not fall entirely within a single CBSA we assigned all mortgages for that zip code to the CBSA with the largest share of houses for that zip code.

We then obtained counts by CBSA of mortgages originating each quarter that meet various criteria, allowing us to compute shares. To identify IO mortgages, we use the IO flag (IO_FLG) reported by LPS, which in turn is based on payment frequency type (PMT_FREQ_TYPE). Since LPS only started classifying loans as IO in 2005, for mortgages that originated and terminated before 2005, we looked at whether the initial scheduled payment in the first month (MTH_PI_PAY_AMT) was equal to the interest rate on the mortgage that month (CUR_INT_RATE) times the initial amount of the loan (ORIG_AMT). Using mortgages that survive past 2005 revealed that in a small but non-negligible number of IOs, the scheduled payment was not equal to but exactly twice the monthly interest rate times the initial loan amount, perhaps because of a quirk in the reporting convention of some servicers. Experimenting with the post-2005 data led us to classify as IOs those mortgages where the ratio of the scheduled payment to the interest rate times original loan amount was in either [0.985, 1.0006] or [1.97, 2.0012]. This approach correctly identified 98.5% of IOs while falsely identifying about 1.5% of non-IOs as IO in the post-2005 period.

Other mortgage shares are constructed as follows. Hybrid mortgages are all 30-year adjustable rate mortgages whose first adjustment in rates (FIRST_RATE_NMON) is scheduled either 24 or 36 months after origination. The share of long-term mortgages is the share of mortgages with an amortization term (TERM_NMON) of 360 months or longer. Subprime mortgages are mortgages whose mortgage type (MORT_TYPE) is coded as Grade 'B' or 'C', following Foote et al. (2010). Privately securitized mortgages are mortgages who remain active for at least 12 months and whose investor status (INVESTOR_TYPE), which is reported each month, corresponds to a privately securitized mortgage pool exactly 12 months after origination. For mortgages the do not last 12 months, we use the last investor reported. Mortgages purchases as an investment property are those for which the occupancy status (OCCUPANCY_TYPE) is given by "Non-owner/Investment."

In addition, we compute a foreclosure rate for each CBSA as the ratio of all mortgages that report being in foreclosure for the first time each quarter to the total stock of first-lien for purchase mortgages that are reported by LPS in that quarter.

For all of these variables, we measure of the propensity to use a particular mortgage product (or to enter foreclosure) by taking the maximum share of that mortgage type among all first-lien for-purchase mortgages in each CBSA between 2000q1 and 2008q4.

### B2.2 Other Mortgage Data

We used two other sources of mortgage data to supplement the LPS. The first is the LoanPerformance (LP) data on mortgages in private-label mortgages pools of nonprime mortgages, meaning Alt-A and subprime mortgages. The second is Home Mortgage Disclosure Act (HMDA) data on mortgage applications.

Unlike the LPS dataset, the LP data matches all liens against a property and reports the combined loan-to-value (CLTV) ratio for each property. We computed the share of first-lien for-purchase mortgages reported in LP in each CBSA in each quarter with a CLTV greater than 80%. Our measure for the propensity towards leverage in each city is the average of this share between 2000q1 and 2008q4 rather than the maximum, since both in the beginning and the end of the period the number of mortgages in LP is small.

We also used LP to construct an alternative measure of subprime mortgage shares. In particular, we counted the total number of first-lien for-purchase mortgages in all private-label nonprime mortgage pools and subprime mortgage pools, respectively. We then aggregated this measure up to the CBSA level using translation tables from Geocorr2K. To convert this into a share, we divide by the total number of first-lien for purchase mortgages reported for each county and quarter under HMDA. In particular, we generated these counts by county, and then aggregated to the CBSA level using translation tables from Geocorr2K.

In addition, we tabulated the number of mortgages issued by known subprime lenders as identified in the HUD Subprime and Manufactured Home Lender List, divided by the total number of mortgages reported under HMDA in each CBSA.

## B3 Other Data

Lastly, we compiled additional control variables for CBSAs from various sources. Where necessary, we used translation tables from MABLE/Geocorr2K, the Geographic Correspondence Engine based on the 2000 Census from the Missouri Census Data Center, to convert data to the CBSA level. For each variable, we calculated both the average level and the average change between 2000q1 and the quarter in which real houses peak in that CBSA (or year in which the peak occurs for annual variables). We used these as controls for regressions involving price appreciation between 2000q1 and the city-specific price peak. For regressions involving price depreciation between the city-specific price peak and 2008q4, we calculate the same two averages for these periods.

Population for each CBSA comes from the Census Bureau's Current Population Reports, P-60, at an annual frequency. All of our averages use log average annual population.

Real per capita personal income for each CBSA comes from the Bureau of Economic Analysis at an annual frequency. All of our averages use log real per capita income.

Unemployment rates for each CBSA come from the Bureau of Labor Statistics (BLS), and are available at a monthly frequency. We aggregate up to a quarterly frequency by averaging the months in each quarter and then compute quarterly averages.

Property tax rates for each CBSA are constructed from data in the American Community Survey (ACS) from the US Census Bureau, using an extract request from IPUMS USA (available at http://usa.ipums.org). In particular, we took data on annual property taxes paid (PROPTX99) and house value (VALUEH). Since PROPTX99 is a categorical variable, we set the tax amount to the midpoint of each respective range. Thus, a tax in the range of $7,001-$8,000 is coded as $7,500. Anything above $10,000 is coded as $10,000. For each household, we estimate the tax rate as the ratio of taxes paid divided by the value of the house. We then compute the median tax rate across all households in the survey in each CBSA in each year. Focusing on the median mitigates the top-coding in taxes paid. Since the ACS has its own definition of metro areas, we need to use the IPUMS metro area-to-MSA/PMSA translation table and then use a MSA/PMSA-to-CBSA table from GEOCORR2K. We also weight households by household weight (HHWT).

Our measure of regulation is the Wharton Residential Land Use Regulation Index from Gyourko, Saiz, and Summers (2008), who summarize the stringency of the local regulatory environment in each community.

Our share of the undevelopeable area in each CBSA comes from Saiz (2010).

# References

Allen, F. and D. Gale (2000). Bubbles and crises. *The Economic Journal 110*(460), 236–255.

Allen, F. and G. Gorton (1993). Churning bubbles. *Review of Economic Dynamics 60*(4), 813–836.

Allen, F., S. Morris, and A. Postlewaite (1993). Finite bubbles with short sale constraints and asymmetric information. *Journal of Economic Theory 61*, 206–229.

Amromin, G., Huang, and Sialm (2010). Mortgage complexity and house price dynamics. Chicago Fed manuscript.

Andrews, D. W. and B. Lu (2001). Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models. *Journal of Econometrics 101*, 123–164.

Arellano, M. and S. Bond (1991). Some tests of specifiation for panel data: Monte carlo evidence andan application to employment equations. *Review of Economic Studies 58*, 277–297.

Arellano, M. and O. Bover (1995). Another look at the instrumental variables estimation of error-components models. *Journal of Econometrics 68*, 29–51.

Barlevy, G. (2009). A leveraged based model of speculative bubbles. Chicago Fed manuscript (revised version of Working Paper 2008-01).

Blanchard, O. and M. Watson (1982). Bubbles, rational expectations, and financial markets. In P. Wachtel (Ed.), *Crisis in the Economic and Financial Structure: Bubbles, Bursts, and Shocks*. Lexington Books.

Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics 87*, 11–143.

Burnside, C., M. Eichenbaum, and S. Rebelo (2010). Understanding booms and busts in housing markets. Northwestern University manuscript.

Campbell, J., S. Giglio, and P. Pathak (2009). Forced sales and house prices. NBER Working Paper 14866.

Case, K. E. and R. J. Shiller (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity 2003*(2), 299–342.

Chambers, M., C. Garriga, and D. E. Schlagenhauf (2009). Accounting for changes in the homeownership rate. *International Economic Review 50*(3), 677–726.

Corbae, D. and E. Quintin (2009). Mortgage innovation and the foreclosure boom. University of Texas at Austin manuscript.

Foote, C., K. Gerardi, L. Goette, and P. Willen (2010). Reducing foreclosures: No easy answers. In D. Acemoglu, K. Rogoff, and M. Woodford (Eds.), *NBER Macroeconomics Annual 2010*, Volume 204, pp. 89–138.

Ghent, A. C. and M. Kudlyak (2009). Recourse and residential mortgage default: Theory and evidence from US states. Federal Reserve Bank of Richmond Working Paper No. 09-10.

Glaeser, E. L., J. Gyourko, and A. Saiz (2008). Housing supply and housing bubbles. *Journal of Urban Economics 64*(2), 198–217.

Gyourko, J., A. Saiz, and A. Summers (2008). A new measure of the local regulatory environment for housing markets. *Urban Studies 45*(3).

Himmelberg, C., C. Mayer, and T. Sinai (2005). Assessing high house prices: Bubbles, fundamentals and misperceptions. *The Journal of Economic Perspectives 19*(4), 67–92.

Holtz-Eakin, D., W. Newey, and H. Rosen (1988). Estimating vector autoregressions with panel data. *Econometrica 56*, 1371–1395.

Kiyotaki, N. and J. Moore (1997). Credit cycles. *Journal of Political Economy 105*(2).

Krugman, P. (2005, August 8). That hissing sound. *New York Times*.

Lacour-Little, M. and J. Yang (2008). Pay me now or pay me later: Alternative mortgage products and housing consumption. California State University at Fullerton manuscript.

Lamont, O. and J. C. Stein (1999). Leverage and house-price dynamics in us cities. *RAND Journal of Economics 30*(3), 498–514.

Landier, A., D. Sraer, and D. Thesmar (2011). The risk-shifting hypothesis: Evidence from subprime mortgage originations. Princeton University manuscript.

Mian, A. and A. Sufi (2009). The consequences of mortgage credit expansion: evidence from the US mortgage default crisis. *Quarterly Journal of Economics 124*(4), 1449–1496.

Minsky, H. P. (1982). *Can "It" Happen Again? Essays on Instability and Finance*. Armonk, New York: M. E. Sharpe, Inc.

Robinson, B. and R. Todd (2010). The role of non-owner-occupied homes in the current housing and foreclosure. Federal Reserve Bank of Richmond No. 10-11.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system gmm in stata. Center for Global Development Working Paper No. 103.

Saiz, A. (2010). The geographic determinants of housing supply. *Quarterly Journal of Economics 125*(3), 198–217.

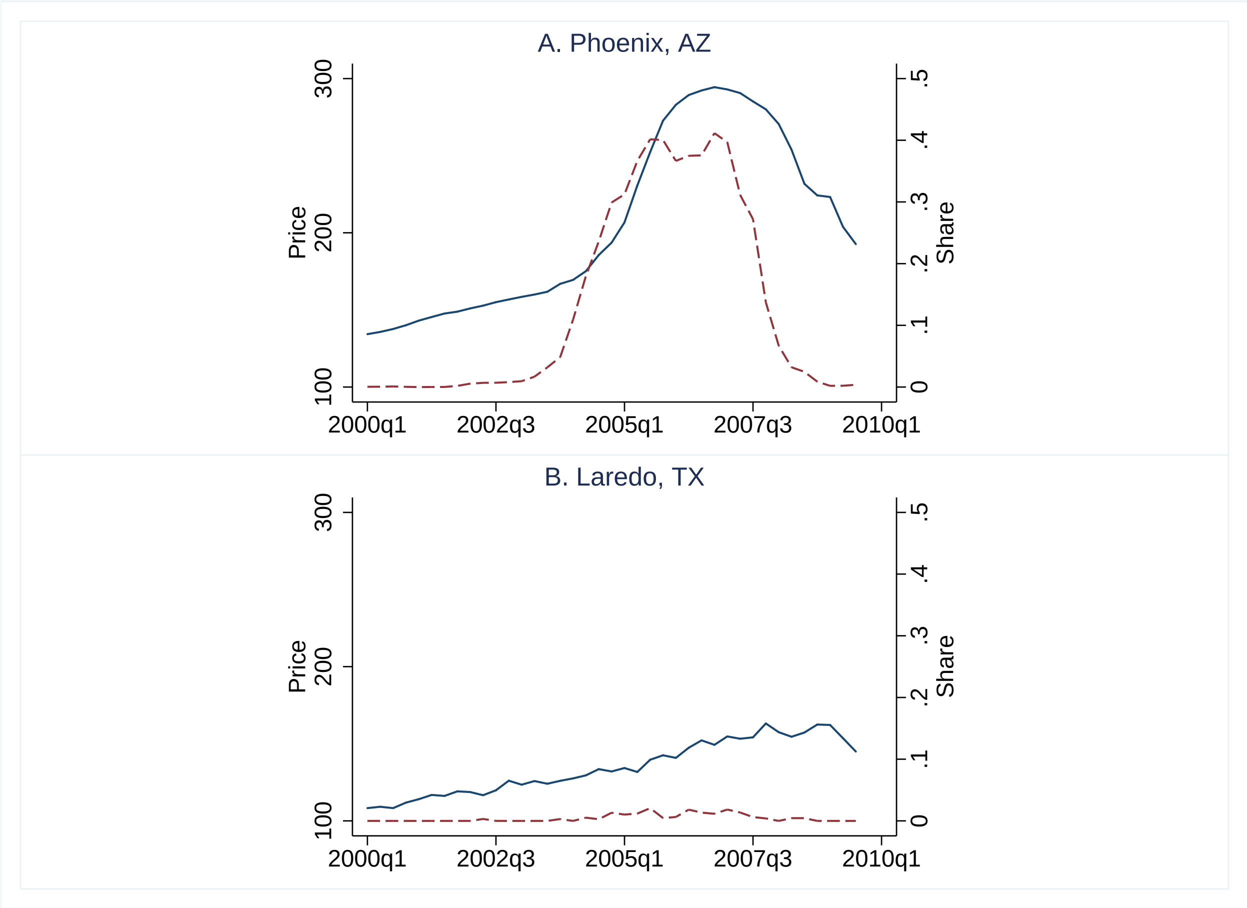
Stein, J. C. (1995). Prices and trading volume in the housing market: A model with down-payment effects. *Quarterly Journal of Economics 110*(2).

Thompson, S. L. (2006). Nontraditional mortgage products. Statement before the Senate Subcommittee on Economic Policy, Subcommittee on Housing and Transportation and Committee on Banking, Housing and Urban Affairs.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics 126*, 25–51.

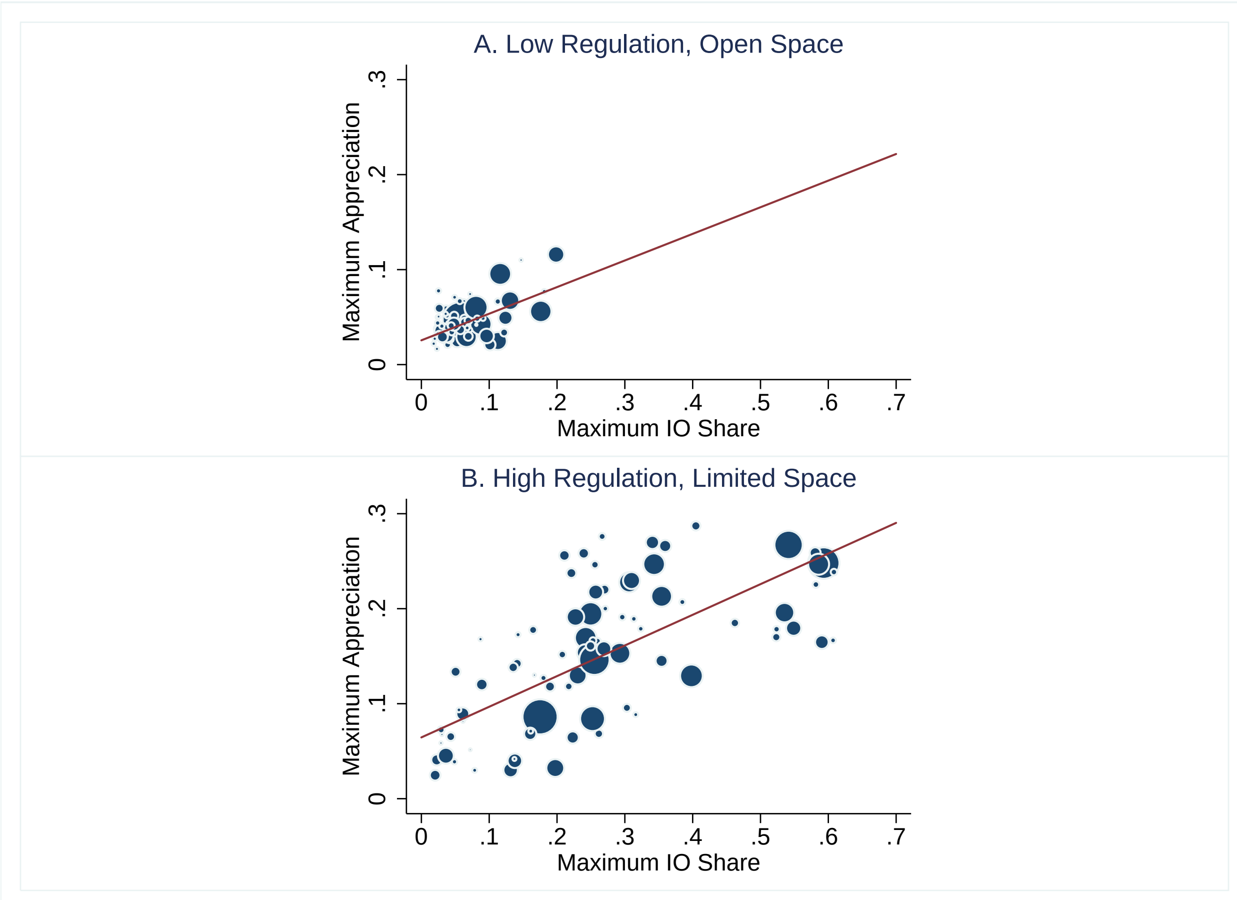
Zeira, J. (1999). Informational overshooting, booms, and crashes. *Journal of Monetary Economics 43*, 237–257.

Figure 1: House Prices and Mortgage Use in Phoenix, AZ and Laredo, TX



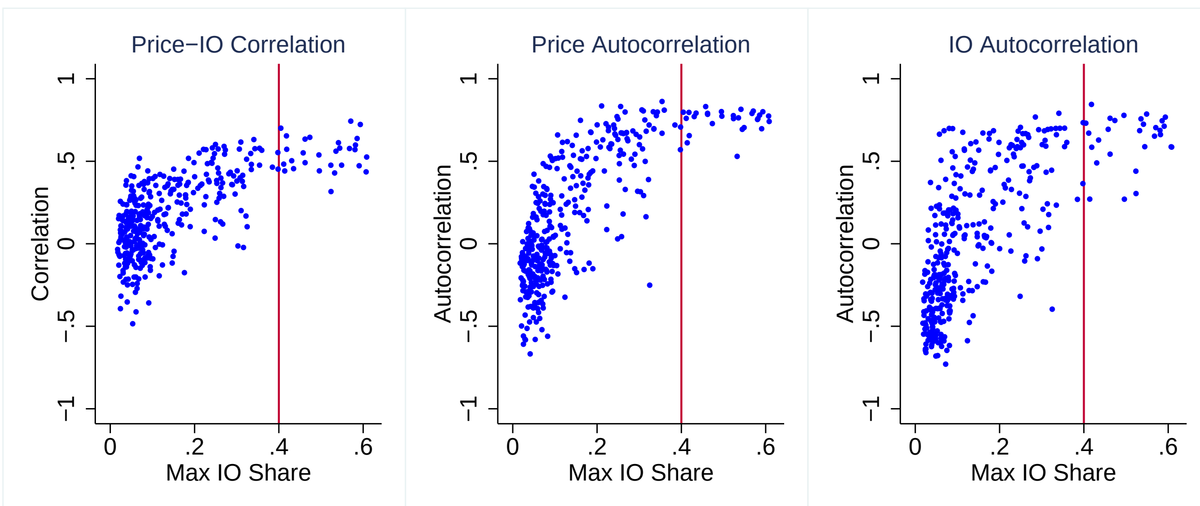

Note: Blue solid lines – Real Price, Red dashed lines – IO Share.

Figure 2: Maximum 4 Quarter Appreciation versus Maximum IO Share



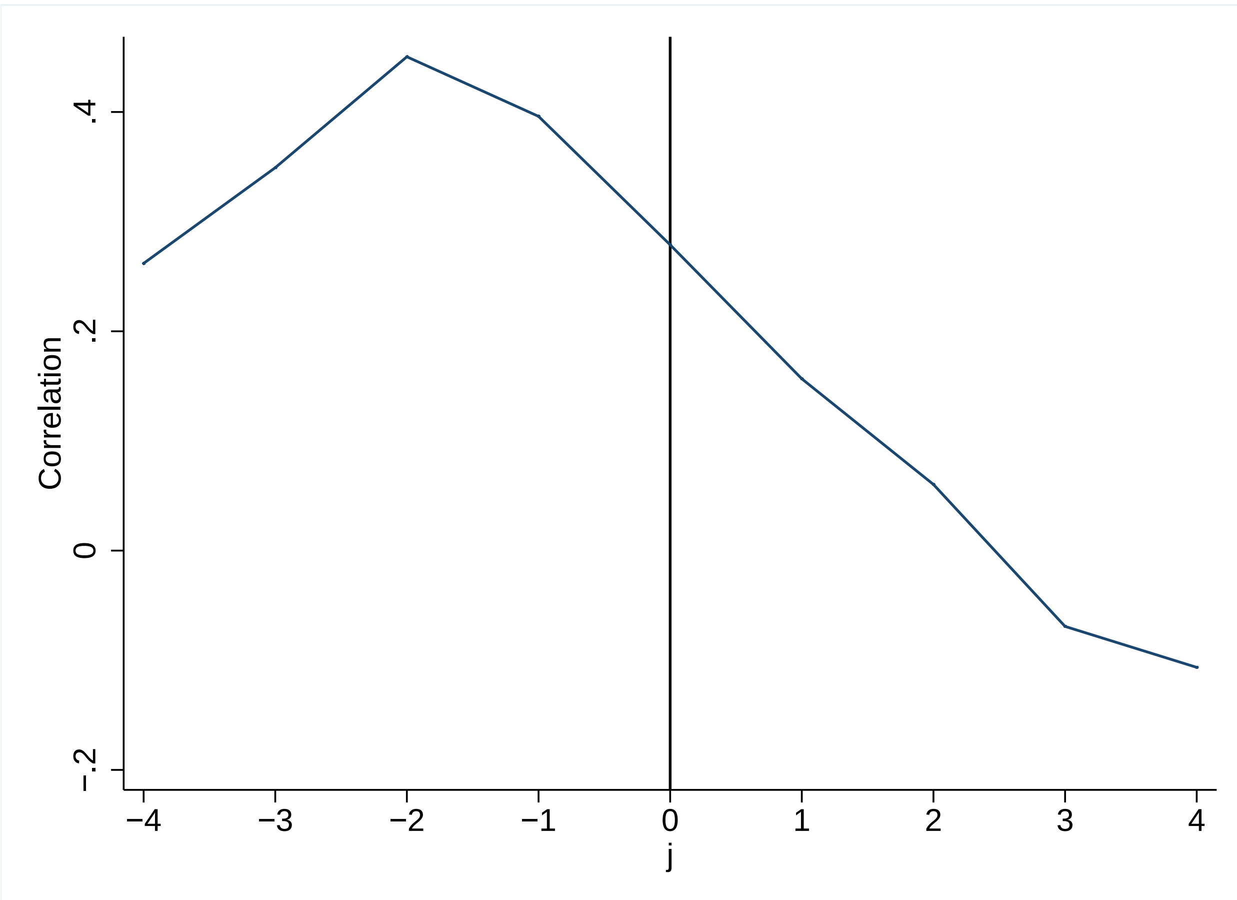

Note: Circles indicate relative size of city. Red lines correspond to OLS regression lines.

Figure 3: Distinct Dynamics Among Cities with Large IO Shares



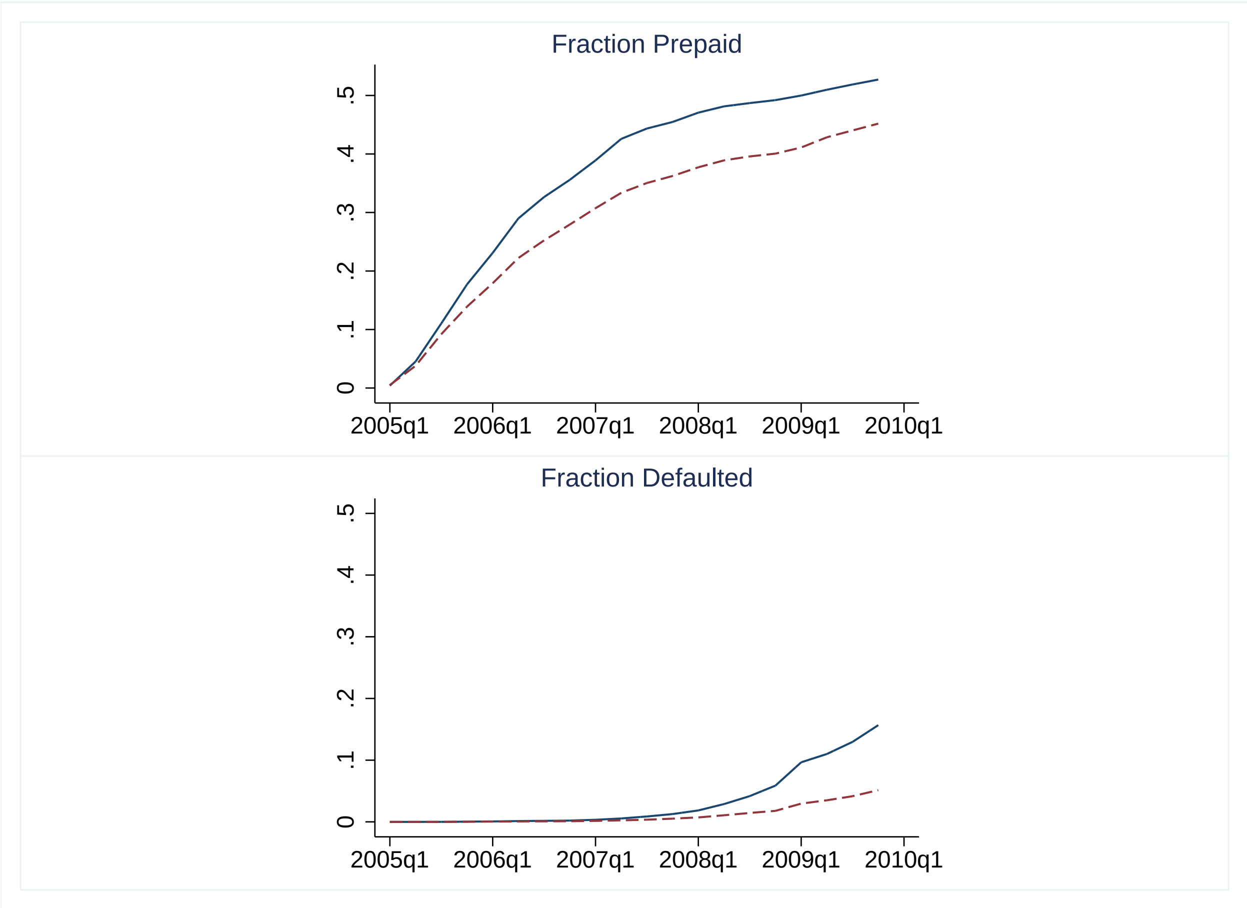

Note: Displayed are correlations and autocorrelations for $\Delta io_{it}$ and $\Delta p_{it}$ for the full sample of cities over the period 2000q1 to 2006q4.

Figure 4: Correlations between $\Delta io_{t+j}$ and $\Delta p_t$




Note: Sample restricted to those cities with maximum IO share of .4 and higher. Red lines correspond to the IO share cut-off of .4.

Figure 5: Pre-Payment and Default of Mortgages Originated in 2005q1




Note: Blue solid lines – IO Mortgages, Red dashed lines – Non-backloaded Mortgages.

Table 1: Mortgage Characteristics

| Type | Year | Mean Amount | All | Owner | Investor | Long Term | Priv. Sec. | Sub-prime | Pre Pay Penalty | ARM |
|---|---|---|---|---|---|---|---|---|---|---|
| IO | 2000 | 451.0 | 0.1 | 97.1 | 2.9 | 58.9 | 58.7 | 0.2 | 0.3 | 92.8 |
| | 2005 | 322.3 | 19.7 | 88.6 | 11.4 | 99.4 | 52.5 | 9.3 | 17.8 | 91.8 |
| Option ARM | 2000 | 192.4 | 1.1 | 97.3 | 2.7 | 98.6 | N/A* | 1.2 | 31.9 | 100.0 |
| | 2005 | 342.2 | 8.2 | 84.7 | 15.3 | 99.9 | 58.4 | 24.3 | 64.2 | 100.0 |
| Fixed | 2000 | 120.9 | 85.2 | 95.8 | 4.2 | 90.9 | 9.4 | 0.6 | 4.5 | 0.0 |
| (Not IO) | 2005 | 185.5 | 57.3 | 92.6 | 7.4 | 91.3 | 14.3 | 2.1 | 2.0 | 0.0 |
| ARM | 2000 | 217.9 | 10.6 | 98.1 | 1.9 | 94.1 | 17.6 | 7.6 | 14.5 | 100.0 |
| (Not backloaded) | 2005 | 215.7 | 14.2 | 90.4 | 9.6 | 99.8 | 45.0 | 29.8 | 28.0 | 100.0 |
| Other | 2000 | 138.7 | 3.0 | 99.9 | 0.1 | 99.1 | 23.7 | 0.0 | 0.0 | 0.0 |
| | 2005 | 195.6 | 0.5 | 96.4 | 3.6 | 98.4 | 4.2 | 9.8 | 0.3 | 0.0 |
| All | 2000 | 132.6 | 100.0 | 96.1 | 3.9 | 91.5 | 10.7 | 1.3 | 5.7 | 11.7 |
| | 2005 | 225.6 | 100.0 | 90.9 | 9.1 | 94.9 | 29.8 | 9.3 | 13.9 | 40.5 |

Note: Entries are percent of indicated type of mortgage except for “Mean Amount” which is in units of thousands of current dollars.

Table 2: Summary Statistics for Price and Mortgage Variables

| | Mean | St. Dev. | Min | 25p | Median | 75p | Max |
|---|---|---|---|---|---|---|---|
| Prices | | | | | | | |
| Max 4Q Growth | 0.10 | 0.07 | 0.02 | 0.04 | 0.07 | 0.15 | 0.33 |
| Mean Growth | 0.05 | 0.04 | 0.01 | 0.02 | 0.03 | 0.07 | 0.14 |
| Max 4Q Decline | 0.10 | 0.10 | 0.00 | 0.05 | 0.07 | 0.11 | 0.65 |
| Mean Decline | 0.05 | 0.06 | 0.00 | 0.02 | 0.03 | 0.06 | 0.35 |
| Mortgages | | | | | | | |
| Max IO Share | 0.15 | 0.14 | 0.02 | 0.05 | 0.09 | 0.22 | 0.61 |
| Max Hybrid Share | 0.14 | 0.06 | 0.04 | 0.10 | 0.13 | 0.16 | 0.45 |
| Max Long Term Share | 0.97 | 0.02 | 0.90 | 0.96 | 0.97 | 0.98 | 1.00 |
| Max Sub-prime Share | 0.11 | 0.05 | 0.03 | 0.08 | 0.10 | 0.14 | 0.32 |
| Max Securitiezed Share | 0.28 | 0.10 | 0.12 | 0.21 | 0.25 | 0.32 | 0.64 |
| Mean 80+ Share | 0.68 | 0.08 | 0.38 | 0.63 | 0.69 | 0.73 | 0.85 |
| Max Investor Share | 0.11 | 0.04 | 0.04 | 0.08 | 0.10 | 0.14 | 0.38 |

Table 3: Baseline Models of Maximum 4 Quarter Price Appreciation in Boom Phase

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Max IO Share | .42 | | .35 | .22 |
| | (.03)*** | | (.03)*** | (.04)*** |
| Population Growth | | 2.76 | 1.35 | .86 |
| | | (.64)*** | (.36)*** | (.27)*** |
| Population | | .008 | -.002 | .001 |
| | | (.006) | (.004) | (.002) |
| Per Capita Income Growth | | .5 | .87 | .21 |
| | | (.44) | (.32)*** | (.15) |
| Per Capita Income | | -.04 | -.1 | -.06 |
| | | (.04) | (.03)*** | (.01)*** |
| Property Tax | | -.04 | .02 | -.006 |
| | | (.01)*** | (.01)* | (.01) |
| Property Tax Change | | -.85 | -.58 | -.64 |
| | | (.2)*** | (.1)*** | (.1)*** |
| Unemployment | | .001 | -.005 | .001 |
| | | (.004) | (.003) | (.002) |
| Unemployment Growth | | -.42 | -.3 | -.06 |
| | | (.11)*** | (.08)*** | (.08) |
| Regulation | | .02 | .004 | -.004 |
| | | (.01)** | (.008) | (.005) |
| Undevelopable Land | | .09 | .02 | .02 |
| | | (.03)*** | (.02) | (.01) |
| State Fixed Effects | No | No | No | Yes |
| Observations | 237 | 237 | 237 | 237 |
| $R^2$ | .65 | .74 | .87 | .97 |

Note: OLS regressions of Maximum 4 Quarter Price Appreciation on indicated variables, weighted by number of mortgages. Most variables are mean values from 2000q1 to quarter of peak real house price. Property taxes are for 2000 and the change between 2000 and the year of the peak price. The Regulation variable is the Wharton Regulation Index and the Undevelopable Land variable is from Saiz (2010). ***, **, and * denote significance at the 1, 5 and 10 percent levels respectively.

Table 4: Controlling for Additional Mortgage Characteristics

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Max IO Share | .35 | .4 | .33 | .36 | .35 | .31 | .34 | .25 |
| | (.03)*** | (.04)*** | (.04)*** | (.03)*** | (.05)*** | (.04)*** | (.02)*** | (.07)*** |
| Max Hybrid Share | | -.14 | | | | | | .01 |
| | | (.09) | | | | | | (.17) |
| Max Long Term Share | | | .43 | | | | | .43 |
| | | | (.24)* | | | | | (.24)* |
| Max Sub-prime Share | | | | -.19 | | | | -.2 |
| | | | | (.09)** | | | | (.21) |
| Max Securitized Share | | | | | .008 | | | .1 |
| | | | | | (.06) | | | (.06) |
| Mean 80+ Share | | | | | | -.15 | | -.06 |
| | | | | | | (.07)** | | (.07) |
| Max Investor Share | | | | | | | .62 | .58 |
| | | | | | | | (.13)*** | (.12)*** |
| Observations | 237 | 237 | 237 | 237 | 237 | 237 | 237 | 237 |
| $R^2$ | .87 | .88 | .87 | .88 | .87 | .88 | .9 | .91 |

Note: OLS regressions of Maximum 4 Quarter Price Appreciation on indicated variables plus the variables in Table 3 excluding Regulation, Undevelopable Land and State Fixed Effects, weighted by number of mortgages. ***, **, and * denote significance at the 1, 5 and 10 percent levels respectively.

Table 5: Controlling for Affordability

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Max IO Share | | .28 | | .29 |
| | | (.07)*** | | (.07)*** |
| Price at Peak | .13 | .04 | | |
| | (.02)*** | (.04) | | |
| Price/Income at Peak | | | .13 | .03 |
| | | | (.02)*** | (.03) |
| Observations | 105 | 105 | 105 | 105 |
| $R^2$ | .86 | .9 | .86 | .9 |

Note: OLS regressions of Maximum 4 Quarter Price Appreciation on indicated variables plus the variables in Table 3 excluding Regulation, Undevelopable Land and State Fixed Effects, weighted by number of mortgages. ***, **, and * denote significance at the 1, 5 and 10 percent levels respectively.

Table 6: Controlling for Additional Mortgage Characteristics With Price Declines

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Max IO Share | .59 | .43 | .6 | .59 | .39 | .7 | .58 | .31 | .35 |
| | (.08)*** | (.08)*** | (.08)*** | (.07)*** | (.09)*** | (.08)*** | (.07)*** | (.12)*** | (.08)*** |
| Max Hybrid Share | | .51 | | | | | | .6 | .26 |
| | | (.15)*** | | | | | | (.3)** | (.24) |
| Max Long Term Share | | | -.33 | | | | | -.64 | -.17 |
| | | | (.49) | | | | | (.53) | (.36) |
| Max Sub-prime Share | | | | .47 | | | | -.39 | -.27 |
| | | | | (.15)*** | | | | (.3) | (.23) |
| Max Securitized Share | | | | | .32 | | | .27 | .12 |
| | | | | | (.09)*** | | | (.1)*** | (.08) |
| Mean 80+ Share | | | | | | .39 | | .27 | -.13 |
| | | | | | | (.15)** | | (.14)* | (.11) |
| Max Investor Share | | | | | | | .37 | .56 | -.04 |
| | | | | | | | (.2)* | (.19)*** | (.15) |
| Foreclosure | | | | | | | | | 12.3 |
| | | | | | | | | | (1.9)*** |
| N | 188 | 188 | 188 | 188 | 188 | 188 | 188 | 188 | 188 |
| $R^2$ | .84 | .85 | .84 | .85 | .85 | .85 | .84 | .88 | .93 |

Note: OLS regressions of Maximum 4 Quarter Price Declines on indicated variables plus the variables in Table 3 excluding Regulation, Undevelopable Land and State Fixed Effects, weighted by number of mortgages. ***, **, and * denote significance at the 1, 5 and 10 percent levels respectively.

Table 7: Granger-Causality Based on OLS Regressions

| | $\Delta p_t$ | | | | $\Delta io_t$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) |
| $\Delta p_{t-1}$ | .50 | .37 | .37 | .42 | .04 | -.02 | .05 | .04 |
| | (.04)*** | (.05)*** | (.05)*** | (.04)*** | (.06) | (.07) | (.07) | (.07) |
| $\Delta p_{t-2}$ | | .14 | .10 | .12 | | -.14 | -.03 | .008 |
| | | (.04)*** | (.04)** | (.04)*** | | (.07)* | (.08) | (.08) |
| $\Delta p_{t-3}$ | | | .15 | .27 | | | -.21 | -.18 |
| | | | (.04)*** | (.05)*** | | | (.08)*** | (.08)** |
| $\Delta p_{t-4}$ | | | | -.31 | | | | -.04 |
| | | | | (.04)*** | | | | (.07) |
| $\Delta io_{t-1}$ | .17 | .13 | .15 | .13 | .43 | .32 | .33 | .31 |
| | (.03)*** | (.03)*** | (.03)*** | (.03)*** | (.05)*** | (.05)*** | (.05)*** | (.05)*** |
| $\Delta io_{t-2}$ | | .12 | .13 | .12 | | .29 | .32 | .34 |
| | | (.03)*** | (.03)*** | (.03)*** | | (.05)*** | (.05)*** | (.05)*** |
| $\Delta io_{t-3}$ | | | -.07 | -.07 | | | -.10 | -.07 |
| | | | (.04)** | (.04)** | | | (.06) | (.07) |
| $\Delta io_{t-4}$ | | | | .02 | | | | -.11 |
| | | | | (.03) | | | | (.06)* |
| $\sum \Delta p$ | 0.50 | 0.51 | 0.63 | 0.50 | 0.04 | -0.16 | -0.19 | -0.17 |
| | (.04)*** | (.04)*** | (.05)*** | (.06)*** | (.06) | (.07)** | (.09)** | (.10)* |
| $\sum \Delta io$ | 0.17 | 0.25 | 0.21 | 0.20 | 0.43 | 0.61 | 0.54 | 0.46 |
| | (.03)*** | (.03)*** | (.03)*** | (.04)*** | (.05)*** | (.05)*** | (.06)*** | (.06)*** |
| AR(1) | 0.00 | 0.58 | 0.00 | 0.01 | 0.00 | 0.36 | 0.02 | 0.07 |
| AR(2) | 0.20 | 0.60 | 0.00 | 0.00 | 0.00 | 0.01 | 0.65 | 0.00 |
| AR(3) | 0.00 | 0.00 | 0.00 | 0.09 | 0.40 | 0.62 | 0.87 | 0.04 |
| AR(4) | 0.00 | 0.00 | 0.00 | 0.71 | 0.60 | 0.09 | 0.42 | 0.14 |
| F Statistic | 36.2 | 29.6 | 23.0 | 15.4 | 0.47 | 3.21 | 3.12 | 1.42 |
| P-value | 0.00 | 0.00 | 0.00 | 0.00 | 0.49 | 0.04 | 0.03 | 0.23 |
| Observations | 754 | 725 | 696 | 667 | 754 | 725 | 696 | 667 |

Note: OLS regressions of log price growth or change in IO share on indicated variables (without weights). ***, **, and * denote significance at the 1, 5 and 10 percent levels respectively. "$\sum x$" denotes sum of coefficients associated with variable $x$. "$AR(j)$" indicates the p-value of the Arellano and Bond (1991) test for serial correlation in the residuals of order $j$. "F Statistic" is the test statistic for the null that the non-regressor lag coefficients are all zero with the p-value below.

Table 8: Granger-Causality Based on System-GMM

| | $\Delta p_t$ | | | | $\Delta io_t$ | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (1) | (2) | (3) | (4) |
| $\Delta p_{t-1}$ | .45 | .33 | .35 | .40 | .02 | -.003 | .10 | .16 |
| | (.05)*** | (.05)*** | (.05)*** | (.06)*** | (.08) | (.11) | (.09) | (.11) |
| $\Delta p_{t-2}$ | | .12 | .12 | .09 | | -.22 | -.10 | -.17 |
| | | (.03)*** | (.04)*** | (.06) | | (.10)** | (.11) | (.17) |
| $\Delta p_{t-3}$ | | | .20 | .26 | | | -.24 | -.37 |
| | | | (.06)*** | (.07)*** | | | (.14)* | (.22)* |
| $\Delta p_{t-4}$ | | | | -.32 | | | | -.19 |
| | | | | (.09)*** | | | | (.20) |
| $\Delta io_{t-1}$ | .22 | .16 | .21 | .17 | .60 | .50 | .52 | .45 |
| | (.02)*** | (.04)*** | (.03)*** | (.04)*** | (.05)*** | (.08)*** | (.09)*** | (.10)*** |
| $\Delta io_{t-2}$ | | .12 | .12 | .11 | | .22 | .25 | .24 |
| | | (.03)*** | (.03)*** | (.04)*** | | (.08)*** | (.08)*** | (.09)*** |
| $\Delta io_{t-3}$ | | | -.10 | -.07 | | | -.15 | -.13 |
| | | | (.05)** | (.06) | | | (.07)** | (.07)* |
| $\Delta io_{t-4}$ | | | | .05 | | | | -.02 |
| | | | | (.05) | | | | (.08) |
| $\sum \Delta p$ | 0.45 | 0.45 | 0.67 | 0.43 | 0.02 | -0.22 | -0.24 | -0.58 |
| | (.05)*** | (.06)*** | (.08)*** | (.17)** | (.08) | (.12)* | (.18) | (.44) |
| $\sum \Delta io$ | 0.22 | 0.28 | 0.23 | 0.26 | 0.60 | 0.72 | 0.62 | 0.55 |
| | (.02)*** | (.03)*** | (.04)*** | (.06)*** | (.05)*** | (.06)*** | (.05)*** | (.14)*** |
| J-stat | 28.9 | 28.7 | 28.2 | 28.4 | 29.0 | 28.9 | 28.7 | 26.6 |
| dof | 140 | 138 | 134 | 126 | 140 | 138 | 134 | 126 |
| AR(1) | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| AR(2) | 0.99 | 0.01 | 0.37 | 0.68 | 0.001 | 0.14 | 0.39 | 0.54 |
| AR(3) | 0.00 | 0.00 | 0.01 | 0.37 | 0.05 | 0.43 | 0.85 | 0.77 |
| AR(4) | 0.00 | 0.00 | 0.00 | 0.52 | 0.78 | 0.95 | 0.89 | 0.79 |
| F Statistic | 83.8 | 62.6 | 35.4 | 16.8 | 0.07 | 2.87 | 1.29 | 0.89 |
| P-value | 0.00 | 0.00 | 0.00 | 0.00 | 0.80 | 0.07 | 0.30 | 0.48 |
| Observations | 754 | 725 | 696 | 667 | 754 | 725 | 696 | 667 |

Note: System-GMM estimates of log price growth or change in IO share on indicated variables (without weights). ***, **, and * denote significance at the 1, 5 and 10 percent levels respectively. See Table 7 for descriptions of "$\sum x$," "$AR(j)$" and "F-statistic". "J-stat" indicates Hansen-Sargan test statistic for the over-identifying restrictions, where "dof" is the degrees of freedom of the test. In all cases the p-value is 1.