Cash is King? Understanding Financing Risk in Housing Markets*

Lu Han
Wisconsin School of Business
University of Wisconsin – Madison
lu.han@wisc.edu

Seung-Hyun Hong
Department of Economics
University of Illinois
hyunhong@illinois.edu

January 2024
First Version: July 2018

Abstract

In Los Angeles, all-cash home purchases quintupled during the last decade. Compared with an else-equal mortgage offer, a cash offer is associated with 29% shorter time-to-close and a 2-3.9% price discount, indicating a substantial amount of financing risk – the risk to a seller that a transaction may not close on time and may fail to occur again because a mortgage contingency fails. The estimated cash discount aligns well with a canonical model calibrated to the sample market. Our findings reveal that closing risk alone is insufficient to explain the cash discount. Rather, it turns on the possibility that a property back on the market may fail to sell, requiring a substantial risk compensation. The estimated cash discount is smaller during booms and in larger markets, highlighting the inseparability between financial frictions in the mortgage market and search frictions in the housing market.

Keywords: cash, time-to-close, financing risk, mortgage contingency, liquidity

JEL classification: D12, D83, G21, L85, R30

*Lu Han gratefully acknowledges financial support from the Social Sciences and Humanities Research Council of Canada. This paper is also based on work supported by (while Seung-Hyun Hong serving at) the National Science Foundation. We are grateful to the Editor, Christopher Parsons, and two anonymous referees for their valuable comments and suggestions that significant improved the manuscript. All errors are our own. Any opinion, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

1. Introduction

The recent decade has witnessed a steady increase in all-cash home purchases, which skyrocketed to make up nearly one-third of all home purchases in 2021.[1] While cash purchases are dominant for institutional buyers (Lambie-Hanson, Li, and Slonkosky 2022), their fraction among individual buyers has quintupled in recent years, reaching nearly 20% of non-foreclosure transactions in Los Angeles (Figure 1). This rising trend has been mirrored in many metropolitan areas in the U.S. (Figure 2) and is likely to continue.[2] Despite the rapid growth in cash transactions that contrasts the housing market experience during the pre-crisis period, much less attention has been devoted to understanding cash transactions, contrary to the substantial literature that examines the implications of mortgage payments and related policies (Amromin et al. 2018). This paper aims to fill this gap in the literature by estimating how mortgage and cash transactions differ in sales price and time-to-close, and by understanding how these differences are rationalized by the relative degree of *financing risk* — the risk to a seller that a transaction may not close on time and may fail to occur again because a mortgage contingency fails.

Real estate asset is unique in that a home sale is subject to frictions from two interrelated markets: the mortgage market and the housing market. As a result, financing risk involves an interaction of two risk components: closing risk that a mortgage offer may fall through before the closing date, and re-listing risk that a property back on the market may fail to sell had the mortgage transaction failed to close. Relative closing risk is an unavoidable aspect of the mortgage approval process due to the time it takes to underwrite or process the loan qualification, financial documents, and all of the legal paperwork that needs to be ready for closing a mortgage offer. One of the most common reasons a pending sale falls through is that the buyer cannot qualify for financing.[3] On the other hand, re-listing risk stems from

[1] "Share of homes bought with all cash hits 30% for first time since 2014" (*Redfin News*, July 15, 2021).

[2] In recent years, companies such as FlyHomes, Ribbon, and Orchard started to help homebuyers make cash offers. They buy the house with cash on behalf of the buyer who is prescreened and will pay back with a loan. See, e.g., "Seattle homebuying startup Flyhomes draws $150 million in new funding" (*The Seattle Times*, Jun 10, 2021), "In a hot market, you can buy a home with cash – even if you don't have a lot of it" (*NPR*, December 18. 2021).

[3] According to the National Association of Realtors (2016), after a seller accepted an offer from a buyer, "issues related to obtaining financing" accounted for 37% of the delayed transactions and 14% of the terminated transactions.

the costly listing process and the illiquid nature of the real estate market in which properties not only take time to sell but also may not sell in the end. Almost one-third of residential listings exit the market without being sold, which imposes substantial expected loss for a seller that fails to close the current sales contract (Anenberg and Ringo 2022; Carillo and Williams 2019).

In a competitive market without financing risk, a seller should take the highest offer she receives, regardless of whether the offer is made with cash or mortgage. In reality, stringent policies for income and assets documentation for mortgage approvals translate into additional closing risk and subsequently expose the seller to re-listing risk, which delays or even entirely stalls the transaction. Hence, an offer could win over other offers not only through the best price but also through the fewest financial contingencies. A cash offer is the cleanest offer that a seller could obtain to minimize financing risk, thus the price discount associated with a cash offer relative to a mortgage offer should reflect the degree of financing risk in the housing market.

In this paper, using detailed housing transaction data for Los Angeles, we estimate financing risk by comparing the *sales price* of cash transactions with that of else equal mortgage transactions. We further estimate relative closing risk by comparing the *time-to-close* of these two types of transactions. We preface the empirical work with a theoretical analysis of the no-arbitrage condition for a seller facing two competing offers at time 0: a cash offer that closes with certainty at time 1 and a mortgage offer that closes with financing risk, where financing risk is characterized by both closing risk and re-listing risk. The stylized model is meant to capture the equilibrium mechanisms through which (1) the price discount for cash transactions is driven by financing risk; and (2) the closing delay for mortgage transactions is linked to the relative closing risk. By calibrating the model to the sample market, we solve for the equilibrium price discount, which guides the interpretation of the cash discount estimated later from the transaction level data and sheds light on the underlying mechanisms through which financing risk affects the housing market.

The model offers several useful insights. First, for financing risk to be quantitatively relevant for sellers, it requires both closing risk and re-listing risk to be present. Intuitively, if re-listing and re-selling is costless and riskless, closing risk only causes extra days in receiving the payment, with little impact for most sellers. On the other hand, in a world without closing risk, re-listing risk would not affect sellers and a mortgage transaction would be just as good as a cash transaction. Closing risk and re-listing risk do not work separately; rather they interact with each other, causing non-trivial expected loss for sellers who accept mortgage offers. Indeed, our calibration shows that even a small amount of closing risk turns on the possibility the property may never be sold again, generating a price premium as high as 3.78% for a mortgage offer in the sample market.

In addition, the model build a one-to-one relationship between relative closing risk and the difference in time-to-close for cash and mortgage transactions. The equilibrium cash discount depends on three factors: interest rate, closing risk and re-listing risk. While the interest rate and closing risk are determined by the financial markets, re-listing risk depends on search frictions in local housing markets. To the extent that matching in the housing market is characterized by increasing returns to scale (Ngai and Tenreyro 2014), larger markets are expected to have lower risk of selling a property back on the market and hence a smaller cash discount.

We then turn to the empirical analysis, which is our focus. A central challenge in estimating the cash effect on the sale price is possible clientele effects and unobserved heterogeneity. For instance, cash buyers may self-select into homes that have unobserved poor qualities, as mortgage applications on these houses are unlikely to be approved. This could lead to an overestimation of the cash discount. On the other hand, cash buyers may also represent wealthy people who have a higher willingness to pay for homes with unobserved luxury features, resulting in an underestimation of the cash discount.

To address these challenges, we estimate the effects of cash financing on both sales price and time-to-close by merging multiple CoreLogic datasets on housing transactions, deeds, tax assessment, and Multiple Listing Service (MLS) listing records for non-foreclosure residential house transactions in Los Angeles county. This unusually rich combined dataset offers several advantages. First, they contain detailed transaction-level information on home purchases, including house characteristics, home address, buyer and seller attributes, and financing. Importantly, our data include information on time-to-close (the number of days between the agreement date – when a buyer and a seller agree to transact – and the recording date –

when the transaction is legally recorded),[4] allowing us to estimate the delay associated with the mortgage approval process, which provides a necessary basis for understanding the cash discount. While the unobserved buyer, seller and house attributes may result in a spurious correlation between the sales price and the decision to purchase with cash, such correlation is unlikely a concern for time-to-close. Second, the data contain detailed listing information, including listing strategy (whether significantly underpriced), selling history (whether de-listed and re-listed), seller motivation, and time-on-the-market (the number of days between the listing date and the agreement date), which helps us explicitly control for the selection bias through listing channels. Third, focusing on a single county instead of a national market allows us to more accurately track buyers and sellers across multiple transactions by names, hence sharpening identification.[5] For example, the panel structure of the data allows us to control for time-invariant unobserved differences in houses, neighborhoods, buyers and sellers. In addition, exploiting the buyer's prior transaction experience also helps us understand what types of buyers are likely to be cash buyers, making it possible to find an exogenous source of variation that affects cash financing but is orthogonal to the sales prices.

Consistent with what one might expect, we find that cash purchases are not evenly distributed. Experienced buyers, flippers, and Chinese buyers are more likely to pay with cash. In addition, cash buyers are more likely to be attracted to properties that are significantly underpriced, listed for very long or very short time, or having unusual attributes. We take several steps to alleviate these selection concerns. First, we estimate the effects of cash financing by comparing time-to-close and sales price for the same house bought by buyers with different financing options, for properties bought with different financing options by the same buyer, and for properties sold by the same seller accepting different financing options. Second, we control for time-varying listing-, house- buyer-, seller-specific features, as well as time-varying property assessment values that capture house attributes observed by buyers and sellers but not by researchers. We also use census tract fixed effects interacted with year and month to control for time-varying market conditions at the neighborhood level.

[4]In the CoreLogic deed data, the agreement date is observed for California, but this is not the case for many other states. See footnote 20 for more details.

[5]The accuracy of using names to match buyers/sellers is reasonable when we restrict the sample to one metropolitan area as shown in Bayer, Mangumm and Roberts (2021), but it could be problematic when applied to a much larger geographic market such as the national market.

With this very rich set of controls and fixed effects at our disposal, we further strengthen our identification of the cash discount by using an instrument variable strategy. Following Bayer, Mangum, and Roberts (2021), we use buyer names to track the same individual who bought different homes at different times in Los Angeles. Controlling for buyers' demographics and previous home purchase experience including timing and locations, we show that a buyer's prior cash purchase indicator provides an exogenous source of variation that strongly affects her current financing method but is orthogonal to variations in the sales price due to buyer-house or buyer-seller sorting, making it a plausible instrument for the purpose of establishing the causal effect of the cash purchase on the sale price.

Several patterns emerge from our estimation. First, else equal, a cash purchase reduces time-to-close by about 29%, reflecting the relative degree of closing risk associated with accepting a mortgage offer. Moreover, the sales price associated with a cash offer receives about a 2-3.9% discount compared to an else-equal mortgage offer. This is equivalent to almost half of the total compensation real estate intermediaries earn on a transaction, suggesting that financing risk associated with mortgage offers in the housing market is substantial. For external validity, we find similar results when extending the sample to cover the national markets (top 100 U.S. cities) and institutional buyers. Note that while our unusually rich controls represent an improvement over the literature, admittedly, we would not be able to observe full information that buyers and sellers observe. In this sense, we cannot completely eliminate the possibility of sorting due to unobservables. Nevertheless, the estimated cash discount is robust to the inclusion of various fixed effects as well as an IV strategy, suggesting that we have done a reasonable job controlling for possible selection bias.

Moreover, the estimated cash discount matches most of the equilibrium cash discount calibrated to this market, suggesting that the model captures the underlying transaction process reasonably well and hence provides assuring support for the proposed mechanisms through which financing risk is capitalized into the cash discount. In particular, closing risk alone is not sufficient to explain the cash discount. Rather, closing risk turns on the possibility that a property back on the market may fail to sell, hence requiring a substantial premium to compensate the seller that accepts a mortgage offer.

Finally, we find that the cash discount is smaller in larger, hotter and more active markets.

Such markets are characterized with a lower re-listing or reselling risk if matching in the housing market exhibits increasing returns to scale. Hence, our results are consistent with the model's implication, highlighting the inseparability between financial frictions in the mortgage market and search frictions in the housing market as illustrated in the model.

Our findings give a new meaning to the conventional wisdom that cash is the king in the context of housing markets, in that cash buyers can pay a lower price by eliminating financing risk associated with a mortgage offer. Quantifying financing risk has profound implications for house price formation. In existing bargaining and auction models for housing (e.g. Albrecht et al. 2007; Carrillo 2012), potential buyers compete only on price. Such models do not incorporate the reality that buyers compete for a house along two dimensions – the price they are willing to offer and the contingencies attached to an offer. In the mortgage brokerage market, Woodward and Hall (2010) show that there is substantial room for cash payment (upfront fees) to affect the bargaining outcome on brokerage fees. In the housing market, understanding such cash discount is even more important, as the market is dominated by amateur buyers and sellers, and bargaining is important for price determination (Harding, Rosenthal, and Sirmans 2003).

Our newly gained evidence on cash purchases also contributes to a small but important literature on cash purchases by homebuyers (Asabere, Huffman, and Mehdian 1992; Hansz and Hayunga 2016) and by iBuyers (Buchak et al. 2021). In a recent complementary work, using the Zillow repeat sales data for the U.S. market, Reher and Valkanov (2023) document a puzzling 11% mortgage-cash premium that cannot be entirely explained by the risk that a mortgage transaction may fail to close. Our paper differs from theirs both empirically and conceptually. Empirically, leveraging unusually rich data that contain information not only on housing transactions but also on listings, tax assessments and deed records, we are able to flexibly control for sorting between cash buyers and houses/listings/sellers. We further estimate the delay associated with mortgage transactions, which provides a necessary basis for understanding the cash discount. Conceptually, Reher and Valkanov (2023) take a behavioral approach and link the mortgage premium to sellers' subjective beliefs and extreme uncertain aversion, while we instead rationalize the estimated cash discount as consistent with the underlying frictions in the transaction process, i.e., the interaction between mortgage closing risk and housing re-listing risk. While extreme uncertain aversion may pick up special situations for certain sellers, the interdependence of the mortgage market and the housing market is fundamental to most of real estate transactions. Our different approaches also yield different results. Like Reher and Valkanov (2023), we find that the closing risk alone is insufficient to justify the estimated cash discount. Going beyond that, we show that, for financing risk to be quantitatively relevant for sellers, it requires both closing risk and re-selling/re-listing risk to be present. Consistent with this, we show that our estimated 2-3.9% cash discount is in line with a calibrated equilibrium cash discount based on the closing risk and re-listing risk that we observed from LA.

More broadly, how collateralized borrowing affects asset price dynamics has been a repeated theme in asset pricing literature. It has been examined in the context of a variety of markets, including stocks (Garbade 1982), corporate asset sales (Shleifer and Vishny 1992), and land (Kiyotaki and Moore 1995). Compared to other financial markets, the housing market has been particularly relevant because many buyers have to finance their home purchase with mortgages and it takes time to close a transaction. In this paper, we show that there is substantial room for cash payment to affect the speed and price of housing transactions. With the increasing prevalence of all-cash purchases in recent years, a natural follow-up question is how cash payment affects local house price dynamics and monetary policies, which we explore in future research.

The structure of the paper is as follows. Section 2 presents a conceptual framework. Section 3 describes the data and presents stylized facts. Section 4 describes the estimation framework and presents the identification strategy. Section 5 presents and interprets the main results. Section 6 concludes.

2. Conceptual Framework

This section describes a conceptual framework that microfounds the cash discount our empirical analysis focuses on. We consider the situation of a home seller who faces two competing offers at time $t = 0$: an all-cash offer with price P_c and a mortgage offer with price P_m. For simplicity, we assume that when a cash offer is accepted, the transaction closes (i.e., legal transfer of the house) and the seller receives P_c at $t = 1$ with probability 1. When a mortgage offer is accepted, the mortgage transaction closes and the seller receives P_m at $t = 1$ with a hazard rate q. Should the mortgage transaction fail to close, with probability ω the seller successfully re-lists and re-sells the property to another mortgage offer P_m at the end of $t = 1$. Conditional on the property being re-listed and re-sold, the mortgage transactions closes at $t = 2$ with the constant hazard rate q.[6] This process repeats itself until infinity, as illustrated in Figure 3.

Financing risk refers to the risk to a seller that a mortgage transaction may not close on time and may fail to occur again because a mortgage contingency fails, captured by parameters $(1 - q,\ 1 - \omega)$. $1 - q$ reflects the closing risk that a sales agreement may fall through because a mortgage contingency fails. This closing risk is an unavoidable aspect of the pending process due to the time it takes to underwrite or process the loan qualification, financial documents, and all of the legal paperwork that needs to be ready for closing a mortgage offer. During this process, there is a possibility that the buyer may not qualify for financing for the sale to close.[7]

This closing risk is further amplified by the additional risk that the seller may not be able to successfully sell the house again, referred to as re-listing risk $1 - \omega$. The re-listing risk comprises of two possibilities: first, the seller chooses not to list the home after failing to close a mortgage transaction; second, conditional on a property being re-listed, it fails to be sold. Expired and withdrawn listings are a common feature of housing markets (Carrilo and Williams 2019; Anenberg and Ringo 2022).[8] In the sample market, conditional upon falling

[6] We make three simplifying assumptions. First, we assume a seller always takes a mortgage offer once she accepts a mortgage offer in the first time. Second, we assume that the mortgage offer (P_m) stays the same after a property is re-listed. This is because allowing for the changes in P_m alone cannot generate much cash discount. In addition, we do not find evidence from the sample market supporting price decline (or increase) associated with re-listing. Third, we assume away time-on-the-market for sellers when a property is re-listed. Allowing for it would generate additional delay associated with accepting a mortgage offer but have no quantitative impact on the cash discount derived below.

[7] According to the National Association of Realtors (NAR) Home Buyers and Sellers Generational Trends Report (2018), one of the most common reasons a pending sale falls through is that the buyer is not able to qualify for financing. Buyers sometimes submit a letter that they have been pre-approved or pre-qualified for a loan. But neither letter guarantees that a mortgage will be approved. There is always a possibility that a buyer has a change in their status, such as losing a job or acquiring additional debt. If there is a financing contingency in the agreement, the buyer could walk away without penalty.

[8] Exploiting individual residential listing records in 15 U.S. urban areas between 2004-2013, Carrilo and Williams (2019) find that expired and withdrawn listings are a common feature of real estate markets. For instance, in a suburb of Washington DC (Fairfax County, VA), over half of the properties listed in 2006 were withdrawn and as much as 60 percent of listings expired and/or were withdrawn during the peak of

out of contract, 75% of properties were re-listed but only 55% were re-sold successfully (i.e. with a sales agreement signed), indicating a substantial re-listing cost and re-selling risk for homesellers that fail to close a transaction. Together, the parameters cluster ($1-q$, $1-\omega$) captures the degree of financing risk associated with accepting a mortgage offer relative to a "safe" cash offer, reflecting the fact that stringent policies for income and assets documentation for a mortgage approval may significantly delay or entirely stall the transaction.

In equilibrium, a home seller is indifferent between accepting a cash offer P_c that closes with certainty at time 1 and a mortgage offer P_m that closes with financing risk ($1-q$, $1-\omega$).

$$P_c e^{-r} = P_m q\{e^{-r} + (1-q)\omega e^{-2r} + ((1-q)\omega)^2 e^{-3r} + \cdots\} \tag{1}$$

The left-hand-side of Equation (1) indicates the expected payoff for a seller if she accepts a cash offer at time 0; the right-hand-side indicates the expected payoff if she accepts a mortgage offer at time 0. Solving Equation (1) yields a closed-form solution for the cash discount, expressed as a percentage of the mortgage offer price:

$$\beta \equiv 1 - \frac{P_c}{P_m} = 1 - \frac{q}{1-(1-q)\omega e^{-r}} \tag{2}$$

Calibrating the cash discount must include estimates of the following parameters: (i) closing risk associated with a mortgage transaction relative to a cash transaction $1-q$; (ii) re-listing risk that the house fails to be re-listed or re-sold should the previous mortgage transaction not close, $1-\omega$, and (iii) discount rate, r.

We start from the parameter $q \equiv \frac{q_{\text{mortgage}}}{q_{\text{cash}}} = \frac{\Pr(close=1|cash=0)}{\Pr(close=1|cash=1)}$ where q_{cash} is normalized to be 1. Recall that $q_{\text{mortgage}} \equiv \Pr(close = 1|cash = 0)$ indicates the probability that a transaction closes conditional on accepting a mortgage offer, which cannot be imputed from

the financial crisis. More recently, using the Corelogic data that cover 263 counties in the U.S. between 2002-2021, Anenberg and Ringo (2022) impute the monthly sales hazard and withdrawal hazard over the entire sample. Their findings imply that about one-third of the listings exit the market without being sold.

the MLS data or the deeds data alone.[9] To impute q, we apply Bayes' rule:

$$\Pr(close = 1|cash = k) = \frac{\Pr(cash = k|close = 1)\Pr(close = 1)}{\Pr(cash = k)}, \quad k = 0, 1 \tag{3}$$

Given that $\Pr(close = 1|cash = 1) = 1$, we obtain

$$\Pr(cash = 1) = \Pr(cash = 1|close = 1)\Pr(close = 1).$$

Therefore, $\Pr(cash = 0) = 1 - \Pr(cash = 1|close = 1)\Pr(close = 1)$, and q_{mortgage} is given by

$$q_{\text{mortgage}} \equiv \Pr(close = 1|cash = 0) = \frac{\Pr(cash = 0|close = 1)\Pr(close = 1)}{1 - \Pr(cash = 1|close = 1)\Pr(close = 1)}. \tag{4}$$

From the MLS data, about 93% listings in Los Angeles county during the sample period closed after receiving offers, implying $\Pr(close = 1) = 0.93$. The deeds data for the same market show that about 10% transactions were financed by all cash, implying $\Pr(cash = 1|close = 1) = 0.1$. Plugging these numbers into (4) yields $q_{\text{mortgage}} = 0.92$ and $q \equiv \frac{q_{\text{mortgage}}}{q_{\text{cash}}} = 0.92$. In other words, else equal, compared to a normalized "safe" cash transaction, the risk that a mortgage transaction fails to close is about 8%.

Note that the model implies a one-to-one inverse relationship between the transaction risk and the time it takes to close a transaction.[10]

$$\frac{q_{\text{mortgage}}}{q_{\text{cash}}} = \frac{\tau_{\text{cash}}}{\tau_{\text{mortgage}}} \tag{5}$$

where τ_{cash} and τ_{mortgage} indicate the *unconditional* mean of time-to-close, for cash and mortgage transactions, respectively. Empirically time-to-close for failed transactions is not observed. However, given the 8% closing risk for mortgage relative to cash transactions, even

[9] As noted in Section 3, the Multiple Listing Service (MLS) data include both closed transactions and failed transactions, hence allowing us to impute the probability of closing among the listings that received offers, that is, $\Pr(close = 1)$. However, the MLS data do not contain information about whether the accepted offer was cash or mortgage. On other hand, the deeds data report information on whether the accepted offer was cash or mortgage for all closed transactions. This allows us to impute $\Pr(cash = 0|close = 1)$ and $\Pr(cash = 1|close = 1)$. However, the deeds data do not include failed transactions.

[10] In a stationary market, the time-to-close for a given transaction is distributed geometrically. One can derive the expected time-to-close as follows:

$$\tau \equiv E[t] = \sum_{t=1}^{\infty} tq(1-q)^{t-1} = \frac{1}{q}$$

The equation above indicates that the probability of closing a transaction is uniquely related to the expected time-to-close. Equation (5) follows from the equation above.

among closed transactions, time-to-close for mortgage transactions should be larger than that for cash transactions. Consistent with this, Table 1 reports that the conditional mean time-to-close is 23.4 days for *closed* cash transactions and 34.6 days for *closed* mortgage transactions. We will further test this implication in Section 5.1.

Turning to the re-listing risk parameter ω. Among the residential listings that were sold but not successfully closed in the sample market, roughly 55% were re-listed and re-sold successfully. Thus we take ω as 55%. Finally, we assume that a risk-free annual interest rate is 5%.[11] The implied cash discount for the LA market is thus given by

$$\beta = 1 - \frac{1 - 0.08}{1 - 0.08 \times 0.55 \times e^{-0.05/12}} = 3.78\% \tag{6}$$

Else equal, when competing with a mortgage offer, a cash offer should be compensated by a 3.78% price discount because it removes the risk to the seller that the sale of a property is delayed or even terminated.

The calibration analysis reveals two distinct but related sources of cash discount: financing risk captured by $(1 - q,\, 1 - \omega)$ and time preference captured by r. Among them, time preference plays a very small role. Indeed, the imputed price discount is quite robust to a wide range of r, making the cost of delay less relevant.[12] On the other hand, the delay itself matters as it is uniquely related to the relative closing risk associated with a mortgage offer. Longer expected delay $(\frac{\tau_m}{\tau_c})$ implies higher closing risk $(\frac{q_m}{q_c})$, and vice versa.

More importantly, Equation (6) demonstrates that the main driving force of the cash discount is financing risk, which contains both closing risk $(1-q)$ and re-listing risk $(1-\omega)$. If sellers can re-list and re-sell the property in a costless and riskless way, closing risk alone does not generate much risk premium. However, in the presence of re-listing risk, even a small amount of closing risk can generate a substantial risk premium. This is because in the event that a transaction fails to close, there is a substantial probability, measured by $1-\omega$, that the property exits the market and generates no revenue for the seller. This stems from

[11] We also experiment with a 3% annual interest rate as in Reher and Valkanov (2023), which does not change the equilibrium cash discount much.

[12] In the housing market, the true cost of delay may include an estimate of the inconvenience associated with the delay. For example, a seller may also look to buy in the same time. In this case, a delay in closing a home sale might create double mortgages, making the cost of delay much higher than the risk-free interest rate. While the level of inconvenience is difficult to measure, including it does not have a quantitative impact on the calibrated cash discount, given its insensitivity to r.

the illiquid nature of the real estate market in which properties not only take time to sell but also may not sell in the end. In the sample market, 33% of residential listings exit the market without being sold, resulting in substantial loss for sellers. By accepting a cash offer, a seller avoids not only the possibility that the current transaction does not close on time, but more importantly, the possibility that the property may not be sold again. Hence, a premium is required to compensate sellers who choose a mortgage offer over a cash offer.

While the calibration above is based on the Los Angeles market from 2005-2016 only, the mechanism revealed in equation (6) applies to other markets as well. The difference in cash discount between LA and an average U.S. market is mainly from re-listing risk, measured by $(1 - \omega)$. This is because r reflects the cost of capital and q depends on frictions in the lending process. Financial markets are national and hence one would expect r and q to be more or less similar across metropolitan areas. On the other hand, ω depends on the underlying search frictions in the housing market (Genesove and Han 2012). Unlike financial markets, housing markets are local. To the extent that matching in the housing market is characterized by increasing returns to scale (Ngai and Tenreyro 2014; Genesove and Han 2016), larger markets are expected to to have less dispersion in buyers idiosyncratic taste and a higher success rate of reselling (ω), thus implying a smaller cash discount.

Relationship to estimation: This paper is primarily empirical. The conceptual framework above guides our estimation and interpretation in several ways. First, we provide a theoretical rationale that links financing risk — including both closing risk $(1 - q)$ and re-listing risk $(1 - \omega)$ — to the cash discount. Based on the market statistics from the housing closing process in Los Angeles county between 2005-2016, the parameterized no-arbitrage condition yields an equilibrium cash discount of roughly 3.78%. This provides a useful benchmark for evaluating the cash discount we later estimate in Section 5.2. Second, the model establishes an inverse relationship between closing risk and the delay in time-to-close. This motivates us to estimate the effect of cash purchase on time-to-close, which provides a necessary basis for justifying the estimated cash discount as shown in Section 5.1. Third, holding the interest rate and closing risk equal, the equilibrium cash discount declines with the local housing re-listing risk. This generates testable implications that we take to the data in Section 5.4.

3. Data

This section describes our main datasets, examines the recent trends in cash transactions, and presents systematic differences between cash- and mortgage-transactions.

3.1 Data Description

The primary sources of our data come from two separate datasets from CoreLogic for Los Angeles county: (1) deeds (1990-2016) and tax assessment; and (2) Multiple Listing Service (MLS) data (2005-2016). The deeds and tax data are constructed from the county recorder's office as well as tax assessments in county assessors. The deed data include detailed information about all deed transfers, including the sale amount, mortgage amount, property type, address, the names of buyers and sellers, and the time it takes to close a transaction. The tax data include detailed information on house characteristics from the property assessment in 2016 as well as yearly historical assessment values between 2005-2016. The MLS data contain detailed information on each listing, including listing prices, time-on-market from the listing date until the agreement date (or the date when the listing was delisted from the MLS), and the listing history of the same property if it was listed multiple times before it was eventually sold or withdrawn. To maximize the information on each transaction, we merge the deeds data and the MLS data using parcel identification number, sale amount, and closing date to match two datasets. About 75% of transactions in the deed data can be matched with closed listings in the MLS data. The combined dataset thus covers the majority of housing transactions between 2005-2016.

For the main estimation, we choose Los Angeles county for a number of reasons. First, focusing on a single county versus a national market allows us to more accurately track buyers and sellers across multiple transactions by names, hence sharpening identification. For example, by observing the same buyer in different transactions, we can control for time-invariant unobserved heterogeneity specific to buyers. Exploiting the buyer's prior transaction experience also helps us understand what types of buyers are likely to be cash buyers, making it possible to find an exogenous source of variation that affects cash financing. Second, the Los Angeles county data contain information on time-to-close, which is novel to the literature and provides useful information on closing risk. This is not the case in many other counties. Third, Los Angeles county is the most populous county in the U.S. Fourth, the trend in the share of cash purchase from the Los Angeles sample is comparable to that from the National Association of Realtors (NAR) surveys that cover the entire U.S.[13]

Though our main sample covers transactions in Los Angeles, we also go beyond Los Angeles county and use the CoreLogic data to construct a national sample of the top 100 U.S. cities. This sample includes 100 counties, each of which is the largest county in its respective metropolitan statistical area. We exclude smaller counties because their deeds data quality is not as comparable to large counties. For example, many smaller counties have either a sizable proportion of observations with missing values on cash vs. mortgage purchases, or a small number of arm's length transactions.[14]

Note that our main estimation considers transactions by individual buyers. This is because, compared with individual buyers, institutional buyers have different pricing technology and liquidity needs (Buchak et al. 2021), financial portfolio and investment strategies, objectives and preferences (Han, Ngai, and Sheedy 2022), which affect not only their financing choice but also bargaining strategy, making the interpretation of the cash estimates confounded. Nevertheless, for external validity, Section 5.5 provides robustness checks using the national sample that includes transactions by both individual buyers and institutional buyers in the top 100 U.S. cities for a longer period 1998-2016.

Our analysis focuses on residential properties that consist of single family homes, duplexes, and residential condominiums. Though it would also be interesting to examine other types of properties such as commercial properties, housing markets for these other properties are not generally comparable to residential housing markets. We focus on only arm's length transactions. This means that we exclude non-arm's length transactions (e.g. between family members) or deed transfers involving non-transactions such as foreclosure transfers of prop-

[13]According to the NAR (2016, 2021) based on NAR realtor confidence index surveys, the fraction of all-cash home sales was 0.2 in 2009, 0.31 in 2013, 0.23 in 2016, and 0.24 in 2021. This trend is closely aligned with the trend in Los Angeles shown in Figure 1. Figure 2 further compares the cash purchase trend from the LA sample with that from other cities.

[14]Only about 500 counties in the CoreLogic data have the information on cash purchase available for over 80% of arm's length transactions, and among them, the median number of yearly arm's length transactions is less than 2,000 for about 400 counties.

erties between financial institutions. We further drop missing observations and outliers.[15]

Lastly, we exclude foreclosure sales for two reasons. First, properties at foreclosure auctions must be purchased with cash, and hence an all-cash purchase is by requirement rather than by choice. Second, foreclosure properties are sold by banks, and housing prices from foreclosure sales are not comparable to those from typical house transactions.

3.2 Trends in Cash Transactions

The CoreLogic data contain information on whether a transaction was carried out by all-cash or mortgage. Using this information, we define cash to be an indicator for cash purchase.[16] We further distinguish among four types of home buyers: institutional buyers and three types of individual buyers – "experienced" buyers, flipper buyers, and Chinese buyers. Institutional buyers are corporations or businesses that purchase residential properties. Experienced buyer indicates those who purchased any real estate property in the same county in the past. Following Bayer, Mangum, and Roberts (2021), we use owner names to match buyers across different transactions. This allows us to identify buyers with prior purchases.[17] Flipper buyer is the buyer who sold the house within two years of purchasing the house. Chinese buyer means that the home buyer's last name is in the list of Chinese last names that we have complied.[18] Note that the three types of individual buyers are not mutually exclusive. For experienced buyers, we further consider "downsized" buyers whose previous houses had more bedrooms, more bathrooms, and larger building square footage than their current houses.

Using the LA sample including both individual buyers and institutional buyers, we plot

[15] Some observations do not contain sales prices or house characteristics. For some observations, only the partial amount of the sales price is recorded, or the same property is sold on the same date using multiple transactions, likely involving multiple buyers, or multiple sellers, or multiple deeds for the same property. All these observations are not included in our sample.

[16] The data also include sale amounts and mortgage amounts, based on which we create a separate indicator variable for mortgage amounts equal to zero. We find that both variables for cash purchases are almost identical.

[17] The matching is not perfect because we cannot separate different individuals with the same name. However, this is less problematic when we restrict the sample to one metropolitan area as in Bayer, Mangum, and Roberts (2021), rather than much larger geographical areas. In addition, we exclude those who purchased different properties on the same date, since they are unlikely to be bought by the same individual.

[18] We are motivated to consider Chinese buyers because according to NAR (2015) the bulk of purchases by international clients were all-cash, and international clients from China purchased $28.6 billion worth of properties in 2015, exceeding all other international buyers. However, we do not interpret Chinese buyers as foreign buyers per se, because we cannot separate Chinese foreign buyers from Chinese local residents.

the fraction of cash purchases in Figure 1. A salient pattern from this figure is an unprecedented rapid growth in cash transactions during the last decade. The percentage of cash purchases in Los Angeles was negligible before 2000 and remained below 10% until 2007. After 2007, it started to increase and almost reached 30% in 2013. The rapid growth in all-cash transactions is not unique to Los Angeles. Figure 2 also shows a strikingly similar pattern for several other cities, where we plot the fraction of cash purchases for these cities from our national sample. The increase is clearly noticeable in foreclosure-concentrated areas such as Las Vegas and Miami. In Figure 4, we plot the fraction of all-cash purchase among different buyer groups in Los Angeles. While a similar pattern is observed for experienced buyers and downsized buyers, the growth in all-cash transactions is more pronounced among flippers and Chinese buyers. These patterns are also observed in other markets, as shown in Figure A1 in the Appendix which repeats the same plots using the national sample, excluding LA.

3.3 Cash vs. Mortgage Transactions

Table 1 reports the mean values of key variables for all transactions (column 1), cash transactions (column 2), and mortgage transactions (column 3).[19] Several interesting patterns emerge from the comparison of cash vs. mortgage transactions. First, on average, cash transactions are associated with a lower average sales price ($\$506,315$ for cash transactions and $\$528,050$ for mortgage transactions) and a shorter time-to-close, measured by the number of days between the agreement date and the recording date (23.4 days for cash transactions and 34.6 days for mortgage transactions).[20] The difference in time-to-close between cash transactions and mortgage transactions reflects the delay and uncertainties that a seller faces when accepting a conditional offer subject to financing, justifying higher average sales price for

[19]Table A1 in the Appendix presents similar summary statistics for the national sample without LA that does not include any information on the MLS.

[20]The CoreLogic deed data contain the sale date, which is the date when the transaction documents were signed, and the recording date, which is the date when the transaction was recorded at the county recorder's office. In many states, the difference between these two dates cannot provide the time-to-close, because the closing precedes the recording of the deed, so that the sale date in CoreLogic can be the closing date (when the property and the money are officially transferred), instead of the agreement date (when a buyer and a seller agree on their transaction and sign documents). In California, however, the closing of escrow occurs on the recording date, so that the recording date is also the closing date, and the sale date is the agreement date, which allows us to compute the time-to-close from the number of days between the agreement date (that is, the sale date in CoreLogic) and the recording date.

mortgage transactions.

On the other hand, average time-on-the-market, measured by the number of days between the listing date and the agreement date, does not vary much between cash and mortgage transactions (65.9 days for cash transactions and 64.3 days for mortgage transactions). To see how this difference is compared to the difference in time-to-close over time and across the distribution, we further examine their monthly variations and distributions in Section 5.3. The comparisons in Section 5.3 show that these variations and distributions for time-to-close are clearly different between cash transactions and mortgage transactions, whereas those for time-on-the-market are mostly similar, suggesting that the inverse relationship between cash discount and time-to-close is unlikely to be explained by time-on-the-market.

It is important to note that cash transactions are not independent of house, listing, buyer and seller characteristics: houses, listings and sellers with certain characteristics are likely to sort into certain buyers with cash offers, and these certain characteristics could potentially negatively impact house prices. This is shown in Table 1 that presents summary statistics for house-, listing-, buyer-, and seller-specific features. Compared to homes bought with mortgage, homes purchased with cash are slightly newer, smaller in square footage, parking space and the number of total rooms, having more atypical attributes relative to other homes sold in the neighborhood.[21] In addition, though single family homes account for the majority of residential properties in our sample, they account for relatively fewer cash transactions than mortgage transactions.

Turning to the listing features, on average, cash transactions involve listings that are slightly more likely to be delisted and re-listed, compared to mortgage transactions. Conditional on being listed, properties involved in cash transactions are more likely to be listed significantly lower than comparable listings' asking prices.[22]

[21] To measure how unique a home is, we follow Haurin (1988) and Glower, Haurin, and Hendershott (1998) to create Haurin's atypicality index in the following way. $\text{ATYP}_{ijt} = \sum_k |exp(a+b_k h_{ik}) - exp(a+b_k h^*_{jk})|/P_{ijt}$, where ATYP_{ijt} and P_{ijt} are the atypicality index and sales price for house i in district j at time t; h_{ik} is the kth physical attribute of house i; h^*_{jk} is the mean value of the kth physical attribute of houses in district j in the year of transaction; a and b_k are the intercept and slope estimates from a hedonic regression using the overall market sample in the year when the property is transacted. This index should be interpreted as the aggregate value of deviation of a property's characteristics from the sample mean of properties sold in the same district and in the same year-month, weighted by the hedonic price of that characteristics. Based on this, we construct a dummy that indicates whether the atypicality index is above the 75th percentile.

[22] We run hedonic regressions of the original listing price – the asking price of the very first listing among

Finally, cash transactions also involve different types of buyers and sellers relative to mortgage transactions. Compared to mortgage transactions, cash transactions are three times more likely to involve Chinese buyers, twice more likely to involve flipper buyers, 5% more likely to involve experienced buyers, while there is no significant difference in the experience of sellers or downsize purchases for two groups of transactions.

To examine how systematic these patterns are, we further regress a cash purchase indicator on house-, buyer-, seller-, listing-specific characteristics, controlling for house characteristics, market conditions and time-varying assessment value.[23] Compared with the raw comparison in Table 1, the conditional correlations in Table 2 are similar in direction, but smaller in magnitude and statistically insignificant in some cases. For example, cash transactions are 2.3% more likely to involve more atypical houses, 10% more likely for Chinese buyers and flipper buyers, 1.2% more likely for experienced buyers. On the the other hand, the correlation between cash transactions and seller attributes is weaker and less significant. Consistent with what one might expect, the source of variation in the financing method comes mostly from buyers and less from sellers.

Turning to the listing attributes, we find that a cash transaction is 1% more likely for listings whose asking pricing is 15% lower than the comparable property's, 1.3%-1.7% more likely if the property is listed for a very long or very short time (time-on-market above the 95th percentile or below the 5th percentile). On the other hand, there is no evidence that the properties that have been de-listed and re-listed is more (or less) likely to attract cash offer, perhaps because listing history is typically not available in the public listing websites.

Overall, Table 2 presents rich patterns about how house-, listing-, buyer-, and seller-specific features may differ across cash and mortgage transactions. To the extent that these differences might affect the sales price even in the absence of the payment method difference, this might introduce a selection bias into the estimated cash discount. In what follows, we will leverage rich data to control for these variations explicitly and exploit an IV strategy

multiple listings of the same property before it was sold – and then compute the predicted listing price. If the original listing price is 15% lower than the predicted listing price from hedonic regressions, we define a dummy for whether the listing's asking price is 15% lower than asking prices of comparable properties.

[23]Table A2 in the Appendix provides the results from similar regressions using the national sample including institutional buyers but not including additional control variables. Overall, the coefficients on buyer characteristics are similar.

based on buyers' purchase experience to address the selection issue.

4. Empirical Framework

Our empirical analysis focuses on estimating the effects of having a cash transaction on two housing transaction outcomes: time-to-close and sales price. As discussed in Section 2, the former is related to closing risk and the latter measures financing risk, both reflecting fundamental frictions a home seller faces when accepting a mortgage offer over a cash offer. For comparison, we also estimate the effect of having a cash transaction on time-on-the-market.

4.1 Econometric Model

We consider the following regression:

$$y_{jislt} = \beta \mathsf{cash}_{jislt} + X_j\alpha + \xi_j + \eta_{jt} + G_{it}\gamma + \lambda_i + S_{st}\varphi + \zeta_s + \mu_{jst} + \theta_{lt} + \epsilon_{jislt} \tag{7}$$

where y_{jislt} is the outcome variable — time-to-close, time-on-the-market, or real sales price (in 2010 dollar) — of a transaction in time t that involves house j in location l, buyer i, and seller s; cash_{jislt} is the dummy for a cash purchase in the transaction; X_j is a vector of observed house characteristics, ξ_j and η_{jt} are respectively time-invariant and time-varying unobserved house characteristics; G_{it} is a vector of observed buyer characteristics, λ_i is unobserved buyer characteristic; S_{st} is a vector of observed seller characteristics, ζ_s is unobserved seller characteristics; $(\beta, \alpha, \gamma, \varphi)$ is a vector of coefficients corresponding to cash_{jislt}, X_j, G_{it}, and S_{st}; μ_{jst} is unobserved listing characteristics specific to house j listed by seller s and sold in time t; θ_{lt} is time-varying location-specific unobservables, and ϵ_{jislt} is an idiosyncratic error term.

The key variable of interest is the dummy variable that indicates a cash purchase. The main challenge in estimating the cash effect is the possible correlation between a cash purchase decision and unobservables in (7) that affects the outcome variable. Let u_{jislt} denote the error term combining all unobservables, that is, $u_{jislt} = \xi_j + \eta_{jt} + \lambda_i + \zeta_s + \mu_{jst} + \theta_{lt} + \epsilon_{jislt}$. The error term u_{jislt} can be correlated with cash_{jislt} if poor quality houses are less likely to get financed by mortgage loans in which case the cash coefficient in the price regression would be biased downward. The error term u_{jislt} can be correlated with cash_{jislt} if wealthy people have a taste for better homes in which case the cash coefficient in the price regression would be upward biased. To address these concerns, we include a rich set of control variables on house, buyers, sellers, neighborhoods, as well as a flexible combination of fixed effects.

First, we include rich house characteristics in X_j, such as property size, the number of different kinds of rooms,[24] and information on assessed values collected in 2016.[25] These variables account for observed house characteristics. However, house characteristics may change over time in an unobserved way, as captured by η_{jt} in (7). To account for η_{jt}, we additionally use time-varying assessed values collected each year. The assessed home value contains information about a specific house that is observable to assessors but not to the econometrician and therefore provides a good control for unobserved house conditions. In particular, our yearly assessed values can account for unobserved house conditions unique to the year when the house was sold, thus controlling for the correlation between η_{jt} and cash_{jislt}.

Second, we include a rich set of time-varying buyer characteristics G_{it} and seller characteristics S_{st}. Experienced buyers, flippers, and Chinese buyers are more likely to experience cash transactions. We include dummy variables for these buyer groups as well as the number of prior transactions of the buyer in G_{it}. Similarly, we include the number of prior transactions of the seller, a dummy for whether the seller has sold houses in the same county in the past, as well as a dummy for Chinese sellers in S_{st}. By including these buyer and seller characteristics, we address the possible clientele bias associated with the cash effects.

Third, using listing information from the MLS data, we further construct a set of time-varying listing-specific variables. These include dummies that pick up very fast sales or very slow sales; a dummy that indicates whether a property's asking price is 15% lower than other properties recently sold in the neighborhood; a dummy that indicates whether a property is delisted and re-listed at least once; and a dummy that indicates whether a

[24]Specifically, we include the size of land; square footage information of the property; various building information, such as effective year built; #bedrooms; #rooms; #bathrooms; types of air conditioning; construction types of the property; types of the exterior walls; #fireplace; types of foundation; #parking spaces; parking types; heating types; pool; #stories; types of roof covering; roof types; kinds of view from building; location types of the parcel; types of building style.

[25]Assessed values include the logarithm of the property's land value as well as improvement.

property is delisted and re-listed more than four times. Cash buyers are more likely to look for listings that are either new or stale. Similarly, a house is more likely to accept a cash offer in the immediate window after having fallen out of contract, a time when sellers may have prepared to move or have higher discount rates than normal. A house listed 15% below its market value based on observed attributes may have a significant unobserved problem (e.g., foundation or earthquake risk), making it difficult to serve as the collateral and more likely to wait for and accept a cash offer. While it is not possible to observe a complete set of house characteristics that buyers and sellers would observe, listing information reveals a rich picture about house uniqueness, selling history, listing strategy and seller motivation, which in turn affects the arrival rate of and the acceptance tendency with cash offers. By including listing information, we explicitly control for such sorting between properties and cash offers.[26]

Despite our unusually rich data, there remain unobservables about properties, neighborhoods, buyers and sellers, which may introduce additional sorting that we cannot explicitly account for. To address this, we leverage the panel structure of our data and include a flexible set of fix effects as follows. First, we include house fixed effects for properties with multiple transactions during our sample period. Cash buyers may self-select into homes with unobserved poor qualities, as mortgage applications on these houses are unlikely to be approved. Such houses may also have a lower price or a longer time-to-close. Controlling for house fixed effects as well as time-varying assessed values alleviates this concern.

Second, we include buyer fixed effects for buyers who purchased houses in the same county multiple times during our sample period. Cash buyers may represent wealthy people who have a higher willingness to pay or choose to close faster. Controlling for buyer fixed effects helps address the selection bias associated with unobserved buyer types. Similarly, we include seller fixed effects for sellers who sold houses in the same county multiple times during our sample period, which controls for the correlation between ζ_s and cash_{jislt}.

Third, we include census tract×year×month fixed effects. A census tract typically contains 1,200 to 8,000 people, and tract×year×month fixed effects can thus control for time-varying market conditions at the neighborhood level. In most specifications, including those

[26] We thank the editor and an anonymous referee for this helpful suggestion.

with fixed effects described above, we include tract×year×month fixed effects. Though this reduces the sample size for the estimation, the resulting sample size is sufficiently large so that we can include tract×year×month fixed effects in most specifications to control for time-varying neighborhood unobservables. In addition, we use robust standard errors clustered at the census tract level in all estimations.

4.2 Instrumental Variable Strategy

Unlike time-to-close, sales price is a bargaining outcome between buyers and sellers. Thus, one may be particularly concerned about the endogeneity of the cash purchase dummy in a price regression. In particular, there may exist some time-varying unobservables that result in spurious correlation between the sales price and the decision to make or accept a cash offer. To address this concern, we additionally develop an instrumental variable strategy to further strengthen the identification of the cash discount from the price regression. The instrument we proposed for cash_{jislt} is $\mathsf{cash}_{j'is'l't'}$, that is, the previous cash purchase decision of buyer i of house j in location l when the same buyer i bought her previous house j' in location l' in time t' sold by seller s' ($l' \neq l$, $t' < t$, $s \neq s'$). Intuitively, home buyers who have previously bought with cash are more likely to buy their current home with cash, possibly due to habit persistence or consistent access to deep financial pockets, or limited tax benefits from using mortgage. For example, buyers who prefer to close quickly are likely to continue to make cash offers. In addition, buyers with multiple properties cannot claim mortgage interest deductions, in which case they may prefer all-cash to mortgage as long as they have sufficient cash holdings. While the buyer's preference for closing speed or benefit from a certain tax policy may not affect the sale price of the current property directly conditional on a rich set of controls, this could provide potentially exogenous variations that lead some buyers to continue to use cash financing in her current purchase. For this reason, we consider the buyer's prior cash purchase indicator as our potential instrument. In constructing the instrument, we exclude those who bought their previous house within a year, or those whose previous house was located within 10 miles from their current house for reasons illustrated later in this section.

In Table 3, we report the first-stage regression of the cash dummy on the instrument. In all columns, we include tract×year×month fixed effects and observed house characteristics. Column 2 adds buyer characteristics, while column 3 additionally includes seller characteristics. In column 4, we also include the quality-inflation-adjusted price for the buyer's previous transaction. In all columns, the coefficient on the instrument is large and statistically significant, indicating that the instrument is strongly correlated with the current cash purchase dummy even after controlling for various factors.

For the instrument to be valid in the sales price regression, one also needs to understand what drives variations in a buyer's prior cash purchase decision and whether these variations are exogenous to u_{jislt} in the sales price of her current purchase. To see this, we first regress the buyer's prior cash purchase dummy on a set of time-varying buyer and seller characteristics and market conditions. Several patterns emerge from Table A3. First, the coefficient on seller's experience is consistently small and statistically insignificant. Hence there is unlikely to be any sorting between those who previously bought a house with cash and the sellers in the current house transactions, suggesting that the instrument is exogenous to price variations due to seller characteristics. Second, those who previously bought a home with cash are more likely to be experienced buyers, flippers, and Chinese. Their current purchase is more likely to be in the neighborhood with a higher fraction of cash buyers and higher average house price, as expected. Moreover, we include the buyer's prior purchase price to control for the buyer's wealth which may be related to the sale price of the current purchase.[27] Overall, the significant correlation between the buyer's prior cash purchase and the buyer characteristics, local market conditions, and the buyer's previous house price suggests that these variables need to be included as control variables in the instrumental variable estimation. In other words, the identifying assumption is given by

$$E(u_{jislt} | \mathsf{cash}_{j'is'l't'}, X_j, G_{it}, S_{st}, W_{jslt}, p_{j'is'l't'}) = 0, \tag{8}$$

where W_{jslt} denotes a vector of control variables for listing characteristics, house uniqueness, and time-varying assessed values.

While u_{jislt} is unobservable, we provide an indirect test of (8) by examining the correlation between the instrument and the history of the property (i.e., the previous sales price of the

[27]Note that we use the CPI deflator and house characteristics to adjust the buyer's prior purchase price.

property) and of the seller (i.e., previous cash transaction indicator for the seller). The rationale is that if a house has unobserved attractive features that affect its current sales price, these features are likely reflected in its previous sales price. Similarly, if a seller is eager to close a transaction which affects price bargaining in an unobserved way, this may also be reflected by the seller's prior transaction experience.

Specifically, we regress the instrument on (i) the sale price of house j in its previous transaction before buyer i bought it at period t; and (ii) the cash transaction dummy for seller s's prior transaction before seller s sold house j at period t. The results are reported in Table 4 based on a sample of experienced buyers, since the instrument cannot be defined for the first-time home buyer. In Panel A, we restrict the sample to include only houses with previous sales, so that we can include the same house's previous sale price. Similarly, Panel B uses only sellers with prior transactions, so that we can include the same seller's prior cash transaction.

In column 1, we include tract×year×month fixed effects, observed house, buyer, and seller characteristics, as well as the buyer's previous house price. Column 1 in Panels A-B shows that the instrument is correlated with the same house's previous sale price and the same seller's prior cash dummy, suggesting that the assumption in (8) is not satisfied. However, note that column 1 uses the sample that includes buyers who purchased their previous house recently, or buyers whose previous house was near their current house. If a buyer's previous house was bought only several months ago, or was located only several miles away, unobserved buyer-time or buyer-location specific characteristics could determine both the buyer's cash decision in her previous transaction and the sale price in the current transaction. For example, an investor could identify multiple cheaper properties that became available around the same time or nearby, in which case this buyer-time or buyer-location specific information could be correlated with both the cash decisions and the sale prices in all these transactions.

In column 2, we thus exclude buyers who bought their previous house within a year, or buyers whose previous house was located within 10 miles of their current house. The coefficients on the same house's previous price and the same seller's prior cash dummy now become quantitatively irrelevant and statistically insignificant, which supports the exogeneity assumption in (8). Therefore, for the buyer's prior cash purchase to be a valid instrument in the price estimation, it is important to ensure that time t' and location l' are indeed very different from time t and location l.

The results from Tables 3-4 thus suggest that the validity of our instrument hinges on not only including various additional control variables, but also ensuring that the buyer's prior transaction is sufficiently separated from her current transaction in terms of time and location. Construction of the instrument requires detailed and accurate information on buyers' names and property locations, which is feasible with the LA county sample but difficult for the national sample. For this reason, we apply the instrumental variable approach to the estimation in the LA market.[28] Nevertheless, we still attempt to match buyers in the same county using their names, and construct the instrument for the national sample. Though our instrument for the national sample may contain measurement errors, we use it to examine external validity of our results in Section 5.5.

5. Empirical Results

In this section, we quantify the degree of frictions in lending and housing markets by estimating the closing speed and discount associated with cash offers versus mortgage offers for observationally identical homes.

5.1 Effects of Cash Purchase on Time-to-Close

We begin with time-to-close, which indicates the number of days it takes to record a transaction in the county recorder's office after a transaction agreement is made. In California, the closing of escrow occurs on the recording date, which allows us to compute the time-to-close in Los Angeles county.[29] A typical sales agreement contains numerous contingency clauses related to the acquisition of mortgage financing. The process of securing mortgage loans is complicated and lengthy, and mortgage applications can be disapproved even after buyers

[28]The quality of deed information from most other counties in the U.S. is much lower than that from LA county. Though we attempt to match different buyers using names in our national sample, typos and inconsistencies in names are nontrivial. Similar measurement errors occur in the property location information as well.

[29]This is not the case for most other markets. Hence we report the results on time-to-close for the LA market only.

and sellers agree on their sales. According to the National Association of Realtors (2016) confidence index surveys, after a transaction agreement was made, 33% were not settled on time and 6% were terminated. In particular, "issues related to obtain financing" accounted for 37% of the delayed transactions and 14% of the terminated transactions. In this section, we use the transaction-level data to estimate the effect of having a cash transaction versus a mortgage transaction on time-to-close for hedonically identical properties.

In Table 5, we regress the time-to-close (in the number of days) on an indicator for cash purchase. We use the sample including arm's length transactions of residential properties purchased by individual buyers. All columns include tract$\times$year$\times$month fixed effects to control for tract-level time-varying unobservables that might influence both cash purchase decisions and time-to-close. All columns also include observed buyer and seller characteristics. Columns 2, 3, 5, and 6 additionally control for observed house characteristics, while columns 3-6 add both time-varying assessed value and additional controls described in Table 2. Columns 4-6 include house fixed effects, buyer fixed effects, and seller fixed effects, respectively.

The estimated cash coefficient in columns 1-3 ranges from -10.08 to -10.15. Hence, the estimates are robust to different control variables. In column 4, we further include house fixed effects to control for unobserved house characteristics. The resulting estimate is very similar to those in columns 1-3. This is not surprising given our rich control of house characteristics and market conditions. In column 5, we instead include buyer fixed effects to control for unobserved buyer characteristics. This reduces the estimated cash coefficient to -7.8. In column 6, including seller fixed effects produces almost the same estimate as those in columns 1-3. Given that the average time-to-close for mortgage transactions is about 34.6 days (Table 1) and that a cash offer on average reduces time-to-close by about 10 days, this implies roughly 29% shorter time-to-close associated with cash transactions relative to mortgage transactions.

Note that the time-to-close estimate not only quantifies the closing delay associated with mortgage transactions but also qualitatively speaks about the risk that a transaction may be terminated if a mortgage offer is not approved by the bank. The model in Section 2 establishes a one-to-one link between closing risk and the expected time-to-close for all transactions, including those completed and failed transactions. The time-to-close we observed is only for completed transactions, not for all transactions. Nevertheless, using the MLS data for the sample market, we find that, conditional on having a purchase agreement signed, about 97% of the cash offers closed on time while only 82% of the mortgage offers closed on time. The ratio of the two is qualitatively consistent the closing risk implied from the estimated cash effect on the observed time-to-close for completed transactions.

5.2 Effects of Cash Purchase on Sales Price

We now turn to estimating price discount associated with accepting a cash offer. Table 6 presents the results from regressing the logarithm of real sales price (in 2010 dollar) on a dummy for a cash transaction. Column 1 includes tract$\times$year$\times$month fixed effects to control for time-varying local market unobservables that might influence both cash purchase decisions and sales prices. The cash coefficient estimate is -0.057, indicating that, all else equal, a cash buyer pays about 5.7% less for observationally identical homes purchased with mortgages. As noted earlier, a standard concern here is that the coefficient on the cash dummy can be biased in both directions. It can be biased downward if poor quality houses are less likely to get financed by mortgage loans, in which case the estimate of -0.057 could reflect a correlation between unobserved house characteristics and the likelihood of getting financed. It can also be biased upward if rich people have a taste for better homes, in which case the estimated cash coefficient would be contaminated by the positive correlation between unobserved buyers' wealth/taste and unobserved house conditions. Below we will explore various specifications with a rich set of controls on house, buyers, sellers and listing history/strategy, as well as various fixed effects and an instrumental variable strategy to establish the causal effect of having a cash purchase on the sales price.

In column 2, we add observed buyer and seller characteristics, which only slightly changes the estimate to -0.055. Column 3 adds observed house characteristics, which reduces the estimate by about one percentage point, so that the estimate becomes -0.046. Note that the estimation is based on the MLS-Deed merged sample (2005-2016). In a robustness check, we repeat the estimation in column 3 using the Deed sample data only. This allows for a longer time window 1998-2016 and yields an almost identical estimate (column 1 of Table A6).

Despite a longer time window, the Deed-only sample lacks listing and assessment information. In column 4 of Table 6, we return to the MLS-Deed sample and include additional control variables reflecting various confounding factors such as home uniqueness, listing strategy, and listing history. Listing price is often considered as a directing device in buyers' search (Han and Strange, 2015). In a similar way, buyers also use other information they read from the listing websites to decide whether to visit a property. For example, properties that have been significantly underpriced, or on the market for a very long time, or being de-listed and re-listed multiple times, or with highly unusual characteristics are often perceived as difficult-to-sell houses, which cash buyers may target. At the same time, these properties may have relatively unappealing house characteristics or motivated sellers, thus resulting in a lower price even in the absence of financing method difference. By including the listing strategy (whether significantly underpriced), selling history (whether de-listed and re-listed), time-on-the-market (whether in the top 5th or the bottom 5th percentile), and house uniqueness (atypicality index) in the price estimation, we explicitly control for the selection bias operated though listing information. Doing so reduces the estimated cash effect by 1.2 percentage points from -0.046 to -0.034.[30] In column 5, we add yearly assessed value to control for time-varying unobserved house characteristics, which reduces the estimated cash effect further down to -0.031. Failing to account for listing or assessment information yields an upward bias as large as 3.2 percentage points in the estimated cash discount when we use the Deed-only sample.[31] Together, this suggests that using the transaction data alone is not sufficient in estimating the cash effect on sales price and can yield a much larger discount than it actually is. By incorporating the listing and assessment information, our estimation represents an important improvement over the existing literature.

So far we have found that all else equal, a cash buyer pays about 3.1% less for an observationally identical house if it was purchased with a mortgage. To examine how sensitive this estimated cash discount is to the remaining unobservables, Table 7 presents a series of

[30]In alternative specifications, we include time-on-the-market directly in Table A4 and time-to-close in Table A5. The results remain almost the same.

[31]In column 1 of Table A6, we repeat the estimation in column 3 of Table 6 to the 1998-2016 LA Deed sample, which has a longer time window than the MLS-Deed sample (2005-2016) but lacks information on listing or assessment value. The remaining columns in Table A6 further enrich the baseline estimation with various fixed effects and the IV specification. Compared with Table 6, failing to account for listing and assessment information increases the estimated cash discount by up to 3.2 percentage points.

robustness checks where we include various fixed effects and implement an instrumental variable strategy. In all columns, we use tract×year×month fixed effects, observed buyer, seller, and house characteristics, as well as time-varying assessed value and additional controls listed in Table 6. Column 1 of Table 7 reports the same estimate from column 5 of Table 6 as a benchmark. Columns 2, 4, 6, and 8 use the same baseline specification as in column 1. While column 1 is based on the entire estimation sample, column 2 restricts the sample to properties with repeated transactions. Columns 4 and 6 respectively restrict the sample to transactions with repeated buyers and repeated sellers. Column 8 restricts the sample to buyers with prior transactions for whom we can construct our instrument. These columns are used as a basis to examine the change in the estimates due to additional fixed effects or an instrument, rather than the change in the estimation sample.

Controlling for house fixed effects changes the estimated cash discount from −3% moderately to −3.7% (columns 2 and 3), suggesting that the inclusion of rich house characteristics and time-varying assessed values might be sufficient to control for house characteristics. To maintain the sample size, all the remaining columns thus use observed house characteristics and yearly assessed values, as well as additional controls and tract-level monthly fixed effects to control for house-specific heterogeneity and time-varying local unobservables that might influence cash purchase decisions and house prices.

Controlling for buyer fixed effects changes the estimated cash discount from −3.6% slightly to −3.8% (columns 4 and 5). While it is impossible to observe every aspect of buyers, the robustness of the estimates here suggests that the rich set of time-varying buyers characteristics and listing attributes included in the main specification serve as a reasonable control for the clientele effects. Further, controlling for seller fixed effects reduces the estimated cash discount from −3.1% to −2% (columns 6 and 7). As noted in Section 3.3, the arrival rate of a cash offer is mostly determined by buyers, rather than sellers. This is even more so after we control for a rich set of house-, listing-, and buyer-characteristics. Thus we take the seller fixed effect estimate as a lower bound.

To further strengthen the identification, we instrument the cash dummy for the current transaction with buyer's prior cash, namely, a dummy variable that indicates whether the buyer used cash in the prior purchase. As discussed in Section 4.2, buyer's prior cash is a valid instrument, as long as we also include various additional control variables and ensure that the buyer's previous transaction is sufficiently separated from her current transaction in terms of time and location. In addition, this instrument needs to be constructed correctly by using accurate information on the buyer and the property's location. Column 9 of Table 7 reports that the estimate from the instrumental variable estimation is -3.9%.

To sum up, the estimated price discount associated with a cash purchase ranges between 2% and 3.9%, averaging around 3%. The estimates are statistically significant and consistent across specifications.[32] In other words, sellers are willing to accept about 3% lower prices to avoid financing risk associated with a mortgage offer. This amounts to a half of the total compensation that real estate intermediaries earn on a transaction, suggesting that financing risk associated with mortgage offers in the housing market is substantial.

Recall that in Section 2, the calibration of the sellers' no-arbitrage condition to the LA county between 2005-2016 yields an equilibrium cash discount of roughly 3.78%. Our estimated 2-3.9% cash discount matches most of the model-implied cash discount, suggesting that the canonical model constructed in Section 2 captures the home transaction process reasonably well. This provides assuring support for the proposed mechanisms through which financing risk is capitalized into the cash discount. In particular, the 2-3.9% estimated cash discount cannot be explained by the frictions in the lending process alone. Lending frictions make a mortgage offer 8% less likely to close than an else equal cash offer, as shown in Section 2. Importantly, this relative closing risk associated with mortgage offers makes sellers more likely to face re-listing risk and hence the possibility of a terminated transaction. The latter stems from search frictions in the housing market.

5.3 Time-on-the-Market

So far we have estimated the effects of cash payments on both sales price and time-to-close. Naturally one might be interested in another duration variable in the housing transaction process: time-on-the-market. In this section, we explicitly distinguish between these two

[32] An an additional robustness check, we perform a test for omitted variable bias. Specifically, we follow Mian and Sufi (2014) and compute the identified set $[\tilde{\beta}, \beta^*(R_{max}, \delta)]$ recommended by Oster (2019). Using information in columns 1 and 5 of Table 6, and assuming $R_{max} = 1$ and $\delta = 1$, we obtain $[-0.031, -0.015]$ for the recommended identified set, which is in line with the range of our estimates. We thank an anonymous referee for suggesting this test.

durations by measuring time-on-the-market (TOM) as the number of days between the listing date and the date when the transaction agreement is signed, and time-to-close (TTC) as the the number of days between the transaction agreement date and the legal date of transfer of ownership of the property.

Unlike TTC, TOM and sales prices are simultaneously determined housing market outcomes: sellers prefer a higher price and a shorter TOM. This gives a negative relationship between TOM and price, which is the opposite of the relationship between TTC and price. What we are interested, however, is not the tradeoff between TOM and price, but rather whether there is sorting between TOM and cash purchase that would affect the interpretation of the estimated cash discount. For example, a cash buyer may scoop houses that have unattractive features and hence remain listed and unmatched for a long time. Alternatively, a cash buyer may look for sellers that are desperate to sell and hence stay on the market only for a short time. In Section 5.2, we have addressed this type of sorting by controlling for various measures of TOM in the price estimation. To gain a better understanding about why TOM is unlikely to complicate the interpretation of the cash discount, we now present variations in TOM, in comparison with variations in TTC.

Figure 5 documents the changes in TOM (top figure) and TTC (bottom figure) over time. The average (median) TOM is 65 (41) days and the average (median) TTC is 34 (30) days, both with substantial seasonal and cyclical variations. Throughout the sample period, the average TTC for mortgage transactions remains about one-third above the average TTC for cash transactions. In contrast, the average TOM underlying cash and mortgage transactions are very close to each other. Figure 6 plots the entire distribution of TOM (top figure) and TTC (bottom figure) for cash and mortgage transactions. As expected, TTC for cash transactions tends to be shorter than that for mortgage transactions throughout the distribution. On the other hand, the distribution of TOM for the cash and mortgage transactions are nearly indistinguishable from each other. Figure 7 presents the bar chart of the mean fraction of cash transactions by TOM (top figure) and by TTC (bottom figure), respectively. The fraction of cash transactions hovers around 20% for TTC ranging from 0-20 days and reduces sharply to 5-9% once TCC exceeds 20 days. By contrast, the fraction of cash transactions remains around 10% for different ranges of TOM. While these figures are presented from different angles, they reveal a similar pattern: unlike TTC, the variation in TOM across mortgage and cash transactions is minimal.

To further examine the degree of sorting between cash transactions and TOM, we repeat the time-to-close estimation but replaces TTC with TOM as the dependent variable. As shown in Table 8, controlling for house characteristics and buyer/seller types, the coefficient on the cash purchase indicator is small and statistically insignificant. The result is robust to the inclusion of the house FE, buyer FE, seller FE, and the IV strategy. Finally, Table A4 shows that the estimated cash discount is quite robust to the inclusion of TOM.

To summarize, while TOM is interesting in its own right, is unlikely to complicate our interpretation of the relationship between the cash discount and time-to-close.

5.4 Exploring Heterogeneity in Cash Discount

The model in Section 2 shows that the equilibrium cash discount depends on three factors: interest rate, closing risk and resale risk. While the interest rate and closing risk are determined by financial markets, the re-listing risk is specific to local housing markets. To the extent that re-listing risk changes over time across markets, this generates testable implications that guide us to explore heterogeneities in the estimated cash discount.

In particular, re-listing risk reflects the underlying search frictions in the housing market. With increasing returns to scale to matching in the housing market (Ngai and Tenreyro 2014; Genesove and Han 2016), we expect that markets with fewer listings and fewer buyers have a lower transaction rate. Else equal, this implies a higher risk of failing to sell when a seller puts a house back on the market and hence a higher premium to compensate sellers who take mortgage offers. Taking this implication to the data, we devise three alternative measures of market size that drive variations in re-listing risk.

First, we compare cold market months with hot market months. Following Ngai and Tenreyro (2014) and Genesove and Han (2016), we denote the period between September and January as the cold market months, and the remaining period as the hot market months. In general, spring and summer times are thought to be more active as families with children want to to get settled down in their new residence before the start of the new school year. Similarly, we define the boom period in Los Angeles to be 2002-2006 and 2013-2016, so that we exclude the recession period around and after the Great Recession in the late 2000s. Naturally we expect the risk of failing to sell when a house is re-listed to be lower during the hot market months or the boom period. Finally, a more direct proxy for the size of the underlying market is the number of listings that are available for potential homebuyers to choose among. We therefore proceed by looking at the number of listings in the 5 digit ZIP code area where the buyer bought the house during the year or the month of transaction. A large number of listings are typically accompanied with a large number of buyers, which together imply a smaller risk of failing to sell.

In Table 9, we expand the sales price estimation by including an interaction between the cash purchase dummy and each of these measures. In all columns in Table 9, we use the baseline specification in column 1 of Table 7, because the cash estimates are robust to various specifications in Table 7, and the number of observations in column 1 of Table 7 is much larger than those in any other columns, thus allowing us to capture more variation across neighborhoods over time.

The results are consistent with the implication discussed above. The cash discount is larger during the cold market months (column 1), and smaller during the hot market months(column 2). The cash discount is also smaller during the boom period (column 3). Lastly, we use the MLS data to compute the yearly (or monthly) number of listings in each 5 digit ZIP code area, and create the dummy for a larger number of listings which is equal to 1 if the number of listing in a given ZIP code area for a given year (or month) is larger than the 90th percentile. Column 4-5 show that the cash discount is smaller when the market is more active as reflected by the number of listings.

To the extent that matching in the housing market is characterized by increasing returns to scale, larger and more active markets have a higher sale success rate for re-listed properties. Overall, the results above provide evidence that lower re-listing risk is associated with a smaller cash discount, lending support for the financing risk mechanism as conceptualized in Section 2.

5.5 External Validity

Our main estimation considers transactions by individual buyers in Los Angeles for the matched MLS-Deed data from 2005 to 2016. For external validity, we now generalize the estimation to include institutional buyers and housing markets outside of Los Angeles.

To this end, we construct a national sample of the top 100 U.S. cities (excluding LA) from 1998-2016 as discussed in Section 3.1. Table A1 provides summary statistics for the national sample. Table A7 in the Appendix presents the estimated cash discount based on the national sample. Given that our MLS data and time-varying assessed values are available for Los Angeles only, the results from our national sample cannot be directly compared with the results in Table 7 using our matched MLS-Deed data. Table A7 can be instead compared with Table A6 that uses the Los Angeles sample of individual buyers from the 1998-2016 Deed data.[33] To provide external validity, Table A7 includes both individual buyers and institutional buyers.

Comparing Table A6 and Table A7 shows that the estimated cash discount from an average U.S. city is about 1-2% larger than that from Los Angeles in most specifications, though the differences between two estimates can be even larger (column 5) or smaller (column 7).[34] A relatively larger cash discount in the national sample may not be surprising given that larger markets such as Los Angeles tend to have lower transaction risk and hence lower cash discount, as discussed in Section 5.4. Overall, the cash discount obtained from the Los Angeles sample of individual buyers are mostly comparable to those from the national sample of both individual buyers and institutional buyers. Given that our national sample does not have time-varying assessed values or MLS listing controls, however, the identification here is weaker compared with the main estimation based on the Los Angeles sample. Adding those additional controls is likely to produce a smaller and more reliable estimated cash discount for the national sample, as shown in Section 5.2.

[33] Both tables use the same specifications. Due to data limitation, we cannot include time-varying assessed values or MLS listing controls for the national sample. We identify the same buyers or sellers only within each county, using the same approach used for Los Angeles county. In other words, we do not attempt to identify the same buyers or sellers across different counties.

[34] An exception is column 9 with an instrumental variable, where the cash discount is smaller in Table A7. This might reflect the difficulty of identifying the same buyers and their previous transactions in many other counties, given that the quality of transaction data in many other counties is not as good as that in the Los Angeles data.

6. Conclusion

In Los Angeles, the fraction of all-cash home purchases quintupled during the last decade with the growth being more pronounced among experienced buyers. Else equal, we find that a cash purchase is associated with a roughly 29% shorter time-to-close and a 2-3.9% price discount. The former indicates higher risk of failing to close for a mortgage transaction than a cash transaction; the latter indicates a substantial amount of financing risk — the risk to a seller that a transaction may not close on time and may fail to occur again because a mortgage contingency fails. These estimates are robust to including a rich set of time-varying house-, listing-, buyer-, and seller-specific variables and a flexible combination of fixed effects as well as an instrumental variable strategy. The estimated cash discount matches most of the equilibrium cash discount from a canonical model calibrated to the sample market, providing support for the model's proposed mechanisms through which financing risk is capitalized into the cash discount. In particular, mortgage market frictions alone are not sufficient to explain the estimated cash discount. Rather, closing risk in the mortgage market turns on a possibility that a property back on the market may fail to sell, hence requiring substantial premium to compensate the seller that accepts a mortgage offer. The indispensable role of the the re-listing risk is further supported by the finding that cash discount is smaller in larger and more active markets.

References

Albrecht, James, Axel Anderson, Eric Smith, and Susan Vroman 2007. "Opportunistic matching in the housing market." *International Economic Review* 48(2), 641–664.

Amromin, Gene, Jennifer Huang, Clemens Sialm, and Edward Zhong 2018. "Complex mortgages." *Review of Finance* 22(6), 1975–2007.

Anenberg, Elliot and Daniel Ringo 2022. "Volatility in home sales and prices: supply or demand?" Working Paper.

Asabere, Paul, Forrest Huffman, and Seyed Mehdian 1992. "The price effects of cash versus mortgage transactions." *Journal of the American Real Estate and Urban Economics Association* 20(1), 141–150.

Bayer, Patrick, Kyle Mangum, and James Roberts 2021. "Speculative fever: investor contagion in the housing bubble." *American Economic Review* 111(2), 609–651.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru 2020. "Why is intermediating houses so difficult? evidence from iBuyers." Working Paper.

Carrillo, Paul 2012. "An empirical stationary equilibrium search model of the housing market." *International Economic Review* 53(1), 203–234.

Carrillo, Paul and Benjamin Williams 2019. "The repeat time-on-the-market Index." *Journal of Urban Economics* 112, 33–49.

Garbade, Kenneth 1982. "Federal reserve margin requirements: A regulatory initiative to inhibit speculative bubbles." in *Crises in Economic and Financial Structure* , Lexington, MA: Lexington Books.

Genesove, David and Lu Han 2012. "Search and matching in the housing market." *Journal of Urban Economics* 72(1), 31–45.

Genesove, David and Lu Han 2016. "Measuring the thinness of real estate markets." Working Paper.

Glower, Michel, Donald Haurin, and Patric Hendershott 1988. "Selling time and selling price: the influence of seller motivation." *Real Estate Economics* 26(4), 719–740.

Han, Lu, Rachel Ngai, and Kevin Sheedy 2022. "To own or to rent? the effects of transaction taxes on housing markets." Working Paper.

Hansz, Andrew and Darren Hayunga 2016. "Revisiting cash financing in residential transaction prices." *Real Estate Finance* 33(1), 33–46.

Harding, John, Stuart Rosenthal, and C. F. Sirmans (2003). "Estimating bargaining power in the market for existing homes." *The Review of Economics and Statistics* 85(1), 178–188.

Haurin, Donald 1988. "The duration of marketing time of residential housing." *Real Estate Economics* 16(4), 396–410.

Kiyotaki, Nobuhiro and John Moore 1997. "Credit cycles." *Journal of Political Economy* 105(2), 211–248.

Lambie-Hanson, Lauren, Wenli Li, and Michael Slonkosky 2022. "Real estate investors and the U.S. housing recovery." *Real Estate Economics* 50(6), 1425–1461.

Mian, Atif and Amir Sufi 2014. "What explains the 2007-2009 drop in employment." *Econometrica* 82(6), 2197–2223.

National Association of Realtors 2015. "2015 profile of home buying activity of international clients." NAR Research Division.

National Association of Realtors 2016. “Realtors confidence index survey: report on the December 2016 survey.” NAR Research Department.

National Association of Realtors 2021. “Realtors confidence index survey, December 2021.” NAR Research Group.

Ngai, Rachel and Silvana Tenreyro 2014. “Hot and cold seasons in the housing market.” *American Economic Review* 104(12), 3391–4026.

Oster, Emily 2019 “Unobservable selection and coefficient stability: theory and evidence.” *Journal of Business and Economic Statistics* 37(2), 187–204.

Reher, Michael and Rossen Valkanov 2023. “The mortgage-cash premium puzzle.” Working Paper.

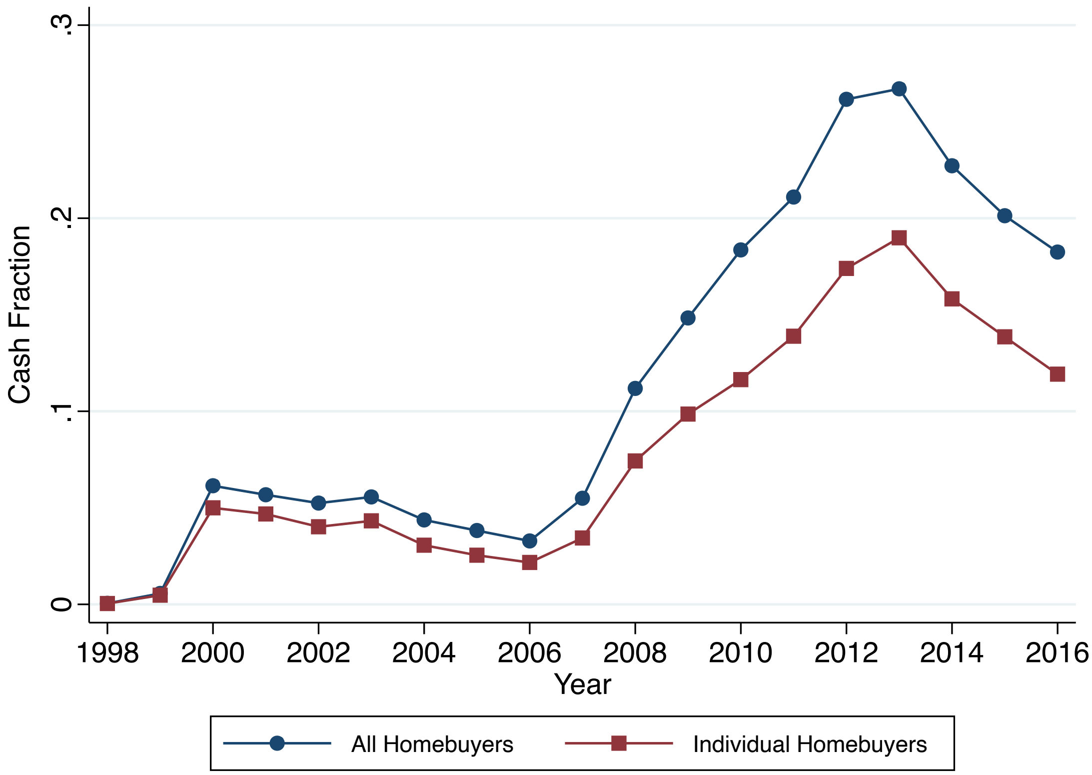
Shleifer, Andrei and Robert Vishny 1992. “Liquidation values and debt capacity: a market equilibrium approach.” *Journal of Finance* 47(4), 1343–1366.

Woodward, Susan and Robert Hall 2012. “Diagnosing consumer confusion and sub-optimal shopping effort: theory and mortgage-market evidence.” *American Economic Review* 102(7), 3249–3276.

Appendix: Definition of Variables

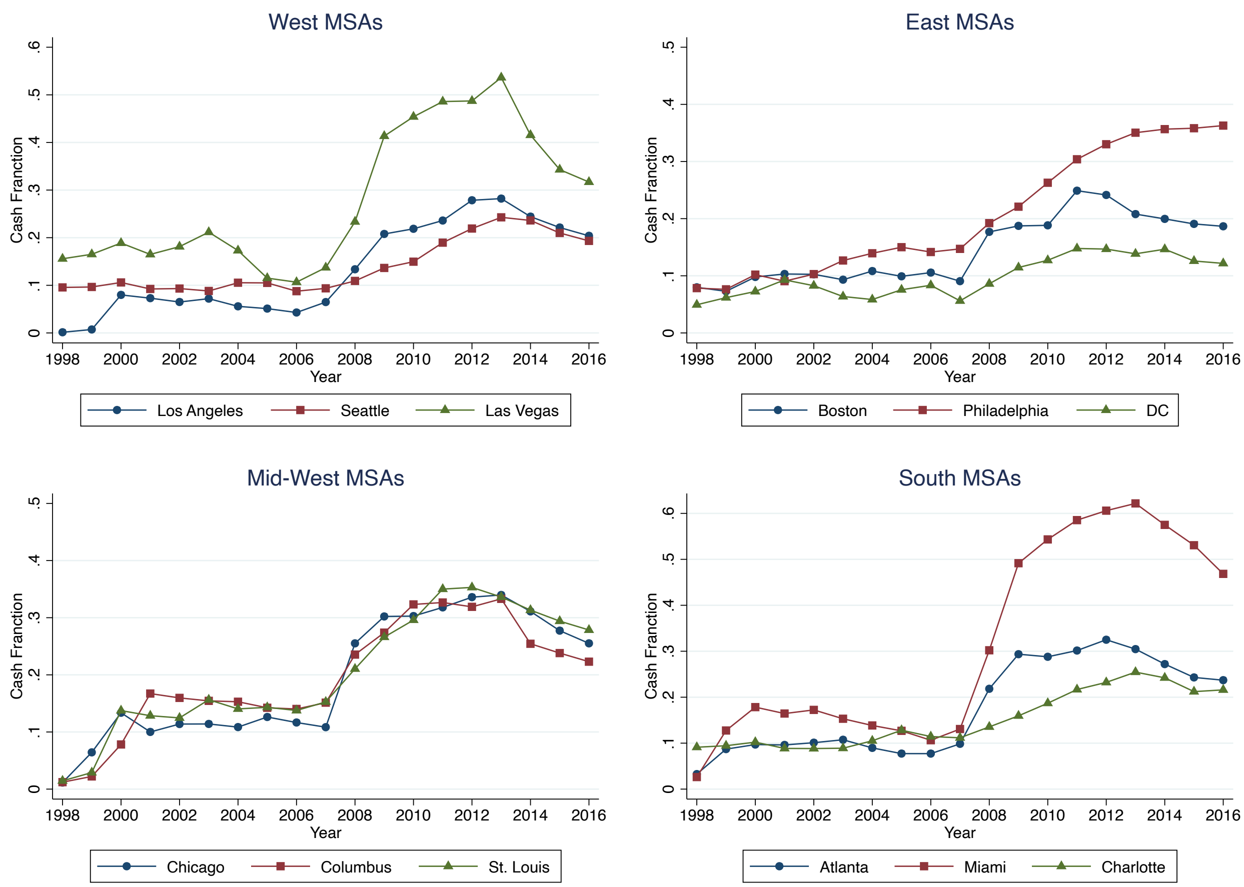
- sales price: real transaction price, deflated by Consumer Price Index.
- time-to-close: the number of days from the agreement date (when a buyer and a seller agree on their transaction and sign documents) to the recording date (when the deed and other recordable documents are recorded at the county recorder's office).
- time-on-the-market: the number of days from the original listing date to the agreement date.
- effective year built: the first year the building was assessed with its current components.
- log #prior transactions of the buyer: the log of 1 + the number of transactions in which the buyer purchased a house in LA county.
- experienced: buyer who purchased any house in LA county in the past.
- downsized: buyer whose previous house has more bedrooms, more bathrooms, and larger building square footage than the current house (since downsized buyers must have purchased houses before, they are also experienced buyers).
- flipper: buyer who sold the house within two years after purchasing the house;
- Chinese: buyer whose last name belongs to the list of Chinese last names.
- log #prior transactions of the seller: the log of 1 + the number of transactions in which the seller sold a house in LA county.
- experienced seller: seller who sold any house in LA county in the past.
- Chinese seller: seller whose last name belongs to the list of Chinese last names.
- asking price 15% lower than comparable asking prices: the dummy for whether the original listing price is 15% lower than the predicted listing prices from hedonic regressions
- atypicality index: constructed by following Haurin (1988), where the higher value means an atypical or distinct house
- delisting and relisting: refers to delisting a house from MLS and relisting it (often to reset the listing's days on market).
- house characteristics: include property type dummies; the size of land; building square footage; various building information, such as effective year built; #bedrooms; #bathrooms; types of air conditioning; construction types; types of exterior walls; #fireplace; types of foundation; #parking spaces; parking types; heating types; pool; #stories; types of roof covering; roof types; kinds of view from building; location types of the parcel; types of building style.
- time-varying assessed value: yearly total assessed value that includes both land value and improvement value available each year from 2005 to 2016.

Figure 1: Yearly Fraction of All-Cash Purchase in Los Angeles[a]



[a]Source: CoreLogic. The figure plots the fraction of all-cash purchases among arm's length transactions of buyers who purchased residential properties in Los Angeles county. All homebuyers include both individual buyers and institutional buyers. The sample excludes non-arm's length transactions as well as sales of foreclosed properties.

Figure 2: Cash Fraction in Some MSAs[a]



[a]Source: CoreLogic. Each figure plots MSA-level cash fractions in each region. Similar to Figure 1, the sample includes arm's length transactions of residential properties by both individual buyers and institutional buyers, but excludes non-arm's length transactions and sales of foreclosed properties.

Figure 3: Canonical Model of Housing Transactions

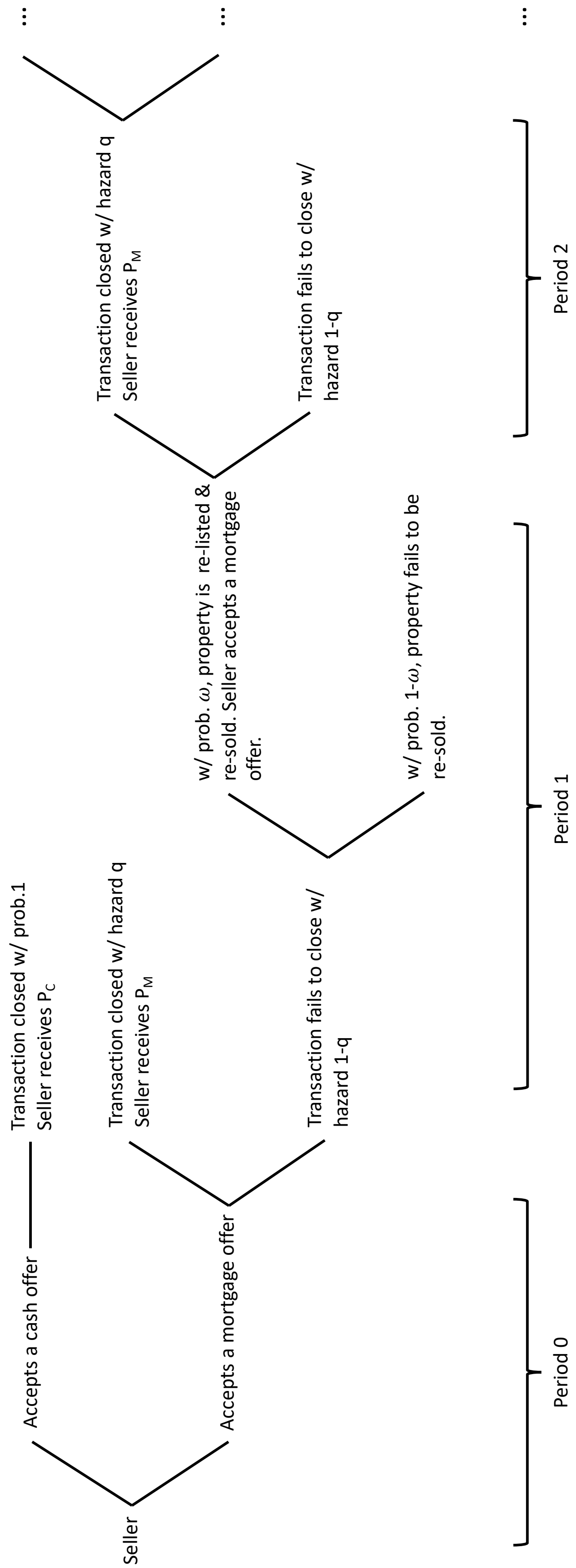
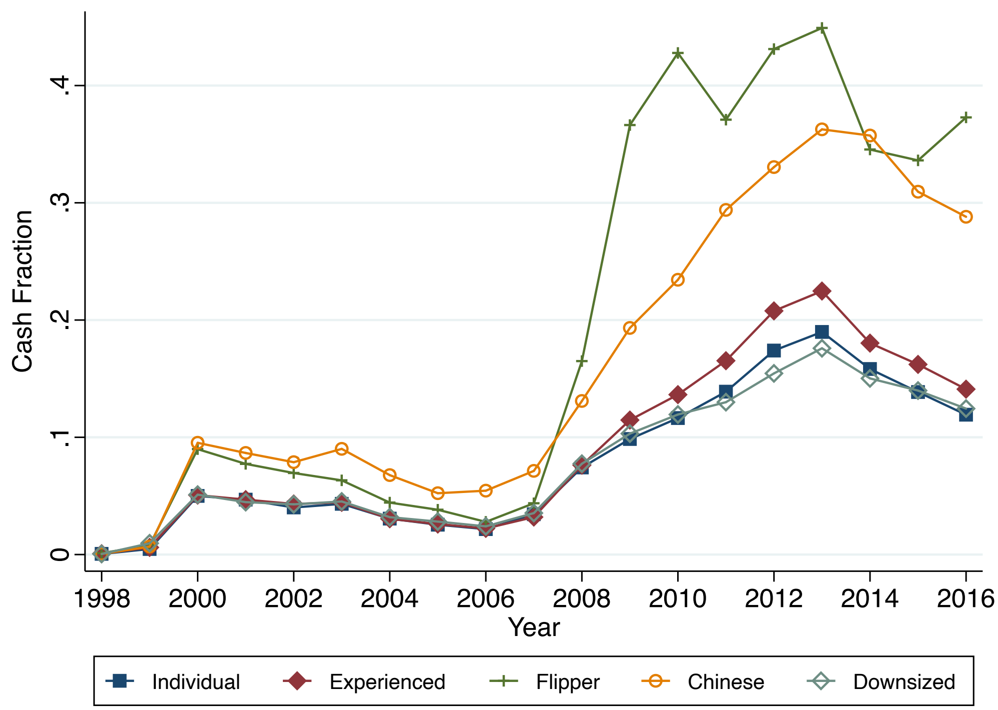
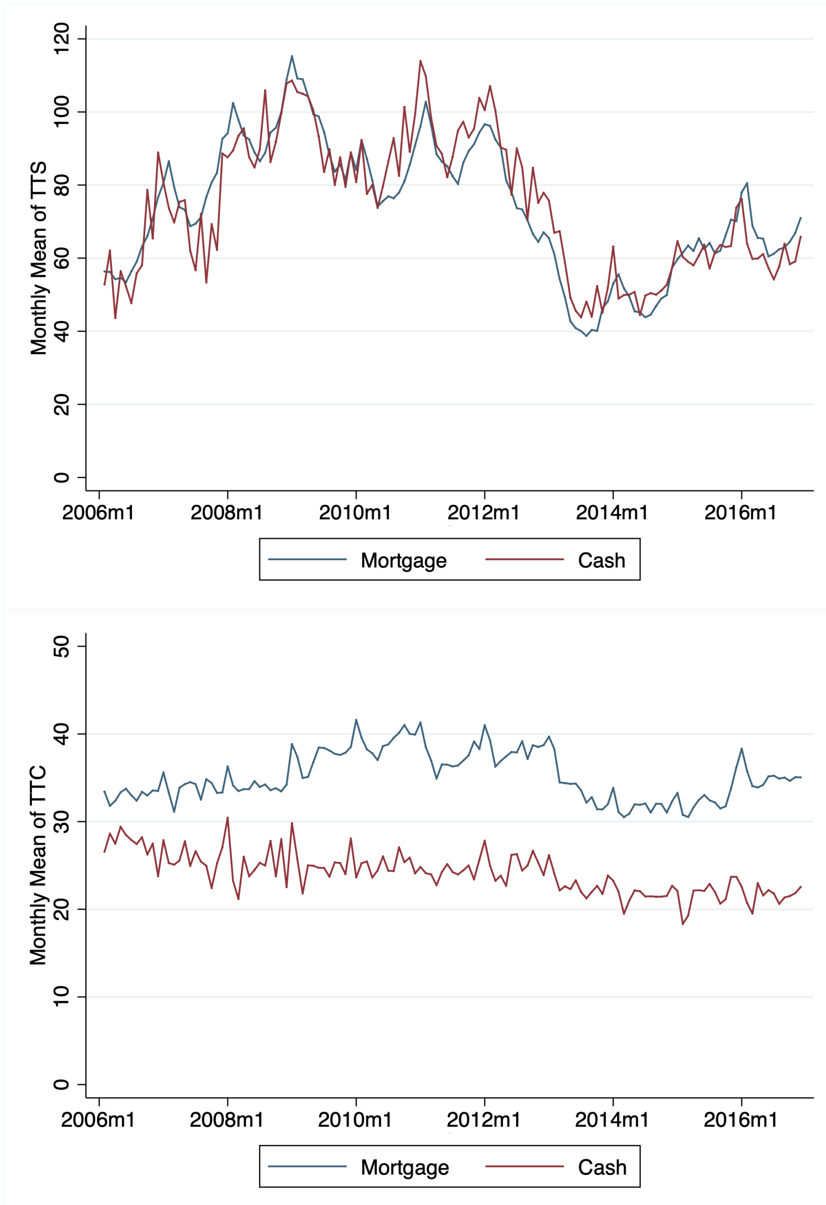


Figure 4: Cash Fraction Among Buyer Group in Los Angeles[a]



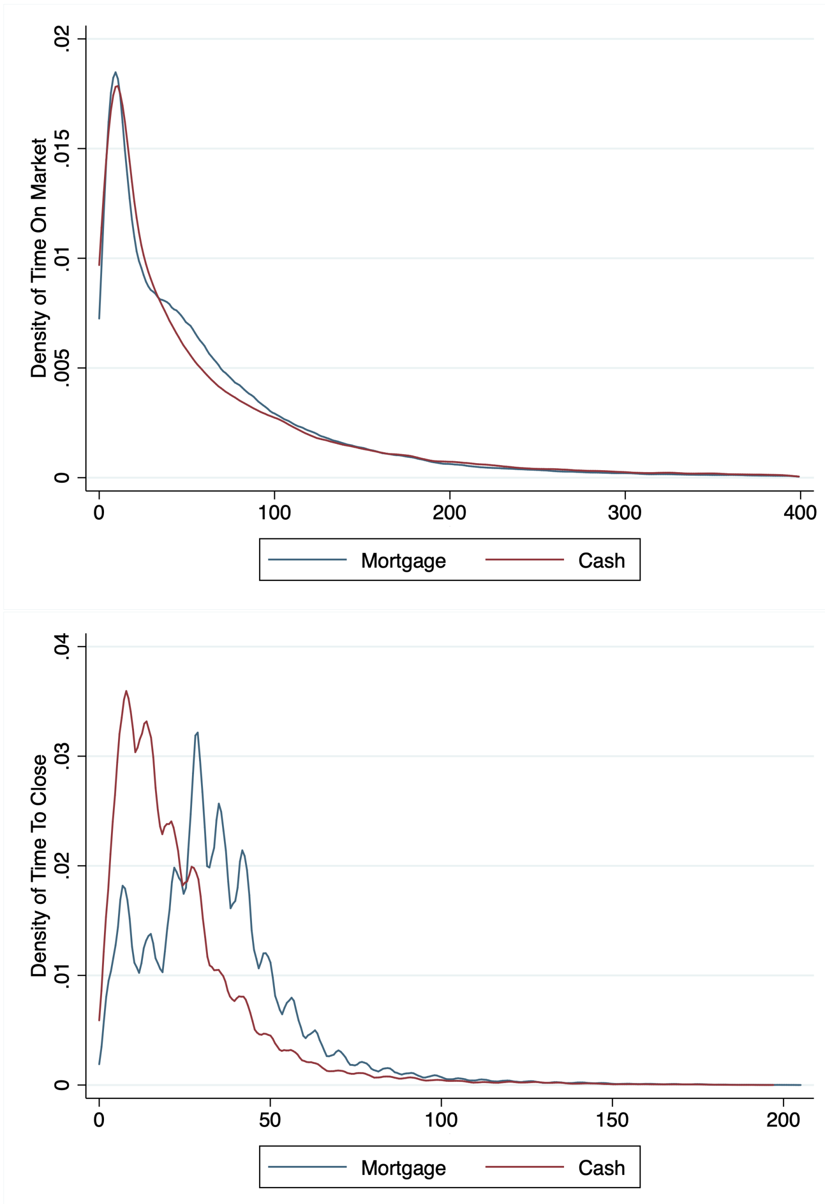
[a]Source: CoreLogic. The figure plots the fraction of all-cash purchases among each buyer group. See Appendix for the definition of each variable.

Figure 5: Monthly Average Time-on-the-Market and Time-to-Close[a]



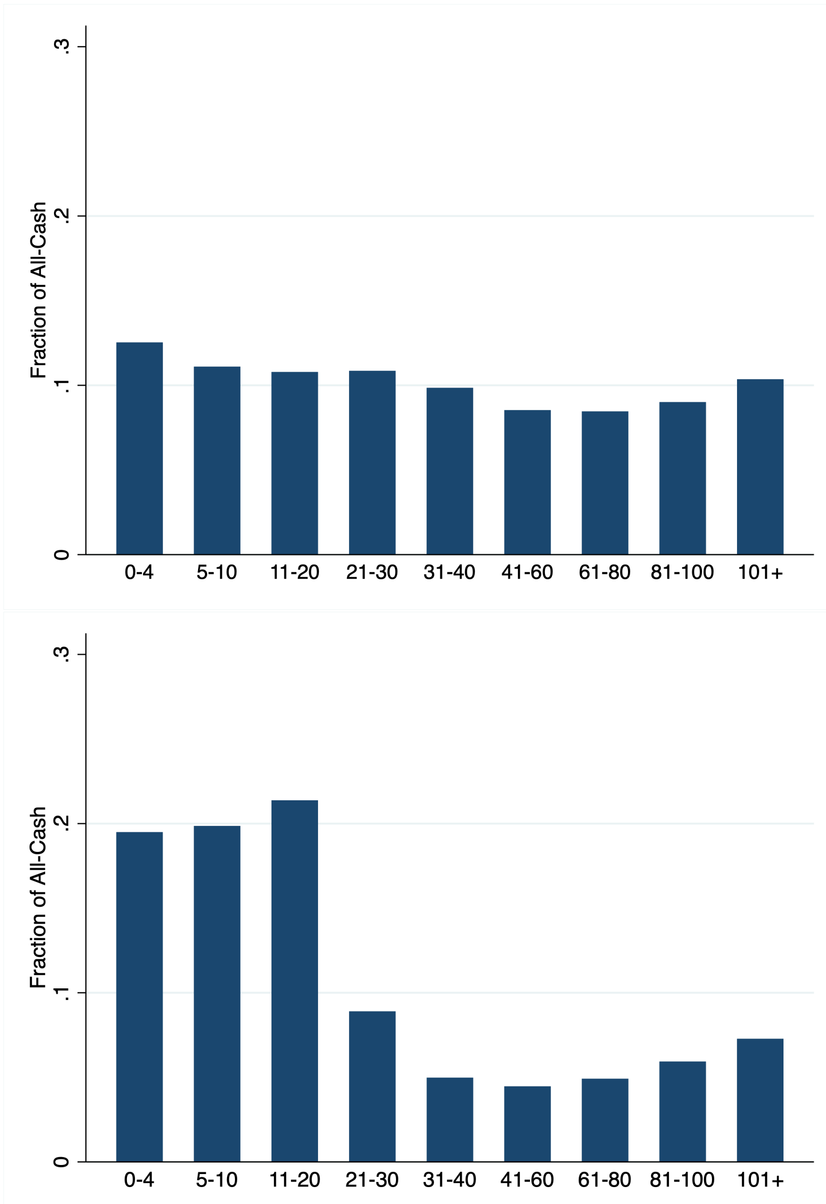
[a]Source: CoreLogic Deed and MLS data. The top plot shows the monthly average time-on-the-market for mortgage vs. cash transactions, while the bottom plot shows the monthly average time-to-close for mortgage vs. cash transaction.

Figure 6: Distribution of Time-on-the-Market and Time-to-Close[a]



[a]Source: CoreLogic Deed and MLS data. The top plot shows the probability density function of time-on-the-market for mortgage vs. cash transactions, while the bottom plot shows the probability density function of time-to-close for mortgage vs. cash transactions.

Figure 7: Cash Fraction Across Time-on-the-Market and Time-to-Close[a]



[a]Source: CoreLogic Deed and MLS data. The top plot shows the monthly average time-on-the-market for mortgage vs. cash transactions, while the bottom plot shows the monthly average time-to-close for mortgage vs. cash transaction.

Table 1: Summary Statistics for Los Angeles County MLS-Deed Data[a]

	All (1)	Cash (2)	Mortgage (3)
sales price (in 2010 dollar)	$525,835	$506,315	$528,050
time-to-close (#days)	33.5	23.4	34.6
time-on-the-market (#days)	64.5	65.9	64.3
#listings	1.44	1.51	1.43
asking price 15% lower than comparable asking prices	0.126	0.160	0.122
time-on-the-market below the 5th percentile	0.057	0.074	0.056
time-on-the-market above the 95th percentile	0.057	0.071	0.055
delisting and relisting at least once	0.269	0.298	0.266
delisting and relisting more than 4 times	0.016	0.022	0.016
atypicality index above the 75th percentile	0.254	0.346	0.244
building square footage	1596	1628	1592
effective year built	1967	1971	1967
#bedrooms	2.96	2.87	2.97
#total rooms	3.46	3.00	3.51
#bathrooms	2.18	2.27	2.17
#parking spaces	1.15	1.01	1.17
single family house	0.691	0.575	0.704
duplex	0.061	0.071	0.060
condo	0.248	0.354	0.236
log #prior transactions of the buyer	0.549	0.607	0.543
experienced buyer	0.395	0.441	0.390
downsized	0.108	0.110	0.108
flipper	0.067	0.136	0.059
Chinese buyer	0.113	0.277	0.094
log #prior transactions of the seller	0.446	0.443	0.447
experienced seller	0.268	0.270	0.268
Chinese seller	0.049	0.083	0.046
observations	536721	54690	482031

[a]The table reports the mean values of transaction-specific variables, house characteristics, dummies for buyer types, as well as variables available in Multiple Listing Service (MLS) data. The sample is constructed by merging transaction data and MLS data, and includes arms-length transactions of residential properties in Los Angels from 2005 to 2016, excluding sales of foreclosed properties and institutional buyers. Column 2 reports the mean of each variable among all-cash transactions, whereas column 3 reports the mean among mortgage transactions. See Appendix for the definition of each variable. Note that #listings is the number of listings of the same house before it was sold, so that it is equal to the number of delisting and relisting (before it was sold) plus 1.

Table 2: Cash Purchase Regressions for in 2005-2016 LA MLS-Deed Data[a]

	dependent variable: cash purchase dummy
	(1)
buyer characteristics	
log #prior transactions of the buyer	0.005**
	(0.001)
experienced	0.012**
	(0.002)
downsized	-0.003
	(0.002)
flipper	0.118**
	(0.003)
Chinese buyer	0.103**
	(0.003)
seller characteristics	
log #prior transactions of the seller	-0.003**
	(0.001)
experienced seller	0.006+
	(0.002)
Chinese seller	0.007*
	(0.003)
listing characteristics	
asking price 15% lower than comparable asking prices	0.010**
	(0.002)
time-on-the-market below the 5th percentile	0.017**
	(0.002)
time-on-the-market above the 95th percentile	0.013**
	(0.003)
delisting and relisting at least once	-0.002
	(0.001)
delisting and relisting more than 4 times	0.005
	(0.005)
house uniqueness	
atypicality index above the 75th percentile	0.023**
	(0.002)
tract×year×month	yes
house characteristics	yes
time-varying assessed value	yes
observations	451964
adjusted R^2	0.118

[a]The dependent variable is the dummy for whether properties are purchased by all-cash. The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. See Appendix for the definition of each variable and the list of variables included in house characteristics. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table 3: First Stage Regressions on Instrument[a]

	dependent variable: dummy for cash purchase			
	(1)	(2)	(3)	(4)
buyer's prior cash	0.170**	0.156**	0.156**	0.162**
	(0.012)	(0.012)	(0.012)	(0.012)
tract×year×month	yes	yes	yes	yes
house characteristics	yes	yes	yes	yes
time-varying assessed values	yes	yes	yes	yes
buyer characteristics	no	yes	yes	yes
seller characteristics	no	no	yes	yes
listing characteristics	no	no	yes	yes
house uniqueness	no	no	yes	yes
buyer's prior purchase price	no	no	no	yes
observations	31142	31142	31142	31142
adjusted R^2	0.094	0.110	0.111	0.115

[a]The dependent variable is the dummy for whether properties are purchased by all-cash. The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. All columns use only buyers with any previous transactions during 1990-2016, and also exclude buyers who purchased their previous house recently (within a year) or whose previous house was nearby (located within 10 miles from their current house). All columns include census tract×year×month fixed effects, and so if only one buyer with previous transactions is observed in a given census tract-year-month, that observation is dropped from the estimation. The buyer's prior cash is the dummy for all-cash purchase in the buyer's previous transaction. The buyer's prior purchase price is the quality adjusted price (in 2010 dollar) in the buyer's previous transaction. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table 4: Exogeneity of Instrument[a]

	dependent variable: buyer's prior cash purchase (1)	(2)
	A. only houses w/prior sale	
ln(house's previous sale price)	0.009*	0.002
	(0.004)	(0.007)
tract×year×month	yes	yes
house characteristics	yes	yes
time-varying assessed values	yes	yes
buyer characteristics	yes	yes
seller characteristics	yes	yes
listing characteristics	yes	yes
house uniqueness	yes	yes
buyer's prior purchase price	yes	yes
exclude nearby or recent prior purchase	no	yes
observations	71439	18260
adjusted R^2	0.092	0.064
	B. only sellers w/prior sale	
seller's prior cash transaction	0.045**	0.015
	(0.012)	(0.017)
tract×year×month	yes	yes
house characteristics	yes	yes
time-varying assessed values	yes	yes
buyer characteristics	yes	yes
seller characteristics	yes	yes
listing characteristics	yes	yes
house uniqueness	yes	yes
buyer's prior purchase price	yes	yes
exclude nearby or recent prior purchase	no	yes
observations	14610	3203
adjusted R^2	0.080	0.046

[a]The dependent variable is the buyer's prior cash purchase, which is the dummy for all-cash purchase in the buyer's previous transaction. The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. All columns include census tract×year×month fixed effects. All columns use only buyers with any previous transactions during 1990-2016. Column 2 additionally excludes buyers who purchased their previous house recently (within a year) or whose previous house was nearby (located within 10 miles from their current house). Panel A uses only houses with the previous sales price of the same house, while Panel B uses only sellers with any previous transactions. The buyer's prior cash is the dummy for all-cash purchase in the buyer's previous transaction. The buyer's prior purchase price is the quality adjusted price (in 2010 dollar) in the buyer's previous transaction. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table 5: Time-to-Close Regressions for 2005-2016 LA MLS-Deed Data[a]

	dependent variable: time-to-close (#days)					
	(1)	(2)	(3)	(4)	(5)	(6)
cash	-10.083**	-10.105**	-10.150**	-9.663**	-7.815**	-10.568**
	(0.149)	(0.150)	(0.150)	(0.920)	(1.260)	(1.131)
tract×year×month	yes	yes	yes	yes	yes	yes
buyer characteristics	yes	yes	yes	yes	yes	yes
seller characteristics	yes	yes	yes	yes	yes	yes
house characteristics	no	yes	yes	no	yes	yes
time-varying assessed value	no	no	yes	yes	yes	yes
listing characteristics	no	no	yes	yes	yes	yes
house uniqueness	no	no	yes	yes	yes	yes
house fixed effects	no	no	no	yes	no	no
buyer fixed effects	no	no	no	no	yes	no
seller fixed effects	no	no	no	no	no	yes
observations	454840	454840	451964	36650	14386	21072
adjusted R^2	0.096	0.097	0.096	0.046	0.106	0.305

[a]The dependent variable is the time-to-close which is the number of days from the agreement date to the recording date when the deed and other recordable documents are recorded at the county recorder's office – in California, the closing of escrow occurs on the recording date. The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. Columns 3-6 include yearly assessment values available for 2005-2016. Columns 4 includes only properties with repeated transactions during the sample period. Columns 5 includes only buyers with two or more transactions during the sample period. Columns 6 includes only sellers with two or more transactions during the sample period. Table 2 provides the list of variables included in seller characteristics, buyer characteristics, and listing characteristics. See Appendix for the definition of each variable and the list of variables included in house characteristics. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table 6: Price Regressions for 2005-2016 LA MLS-Deed Data with Controls[a]

	dependent variable: ln(sales price in 2010 dollars)				
	(1)	(2)	(3)	(4)	(5)
cash	-0.057**	-0.055**	-0.046**	-0.034**	-0.031**
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
buyer characteristics					
log #prior transactions of the buyer		-0.013**	-0.004**	-0.005**	-0.005**
		(0.001)	(0.001)	(0.001)	(0.001)
experienced		0.049**	0.015**	0.014**	0.014**
		(0.002)	(0.001)	(0.001)	(0.001)
downsized		-0.123**	0.015**	0.016**	0.018**
		(0.003)	(0.002)	(0.001)	(0.001)
flipper		-0.059**	-0.061**	-0.057**	-0.056**
		(0.002)	(0.002)	(0.001)	(0.001)
Chinese buyer		0.009**	-0.011**	-0.004**	-0.006**
		(0.003)	(0.002)	(0.001)	(0.001)
seller characteristics					
log #prior transactions of the seller		0.001	0.002*	0.004**	0.006**
		(0.001)	(0.001)	(0.001)	(0.001)
experienced seller		0.086**	0.015**	0.011**	0.003*
		(0.003)	(0.001)	(0.001)	(0.001)
Chinese seller		-0.004	-0.020**	-0.015**	-0.028**
		(0.004)	(0.002)	(0.002)	(0.002)
listing characteristics					
asking price 15% lower than comparable asking price				-0.157**	-0.150**
				(0.002)	(0.002)
time-on-the-market below the 5th percentile				-0.002*	-0.000
				(0.001)	(0.001)
time-on-the-market above the 95th percentile				-0.038**	-0.045**
				(0.002)	(0.002)
delisting and relisting at least once				0.004**	-0.007**
				(0.001)	(0.001)
delisting and relisting more than 4 times				-0.007*	-0.013**
				(0.003)	(0.003)
house uniqueness					
atypicality index above the 75th percentile				-0.092**	-0.086**
				(0.004)	(0.003)
tract×year×month	yes	yes	yes	yes	yes
house characteristics	no	no	yes	yes	yes
time-varying assessed value	no	no	no	no	yes
observations	454840	454840	454840	451964	451964
adjusted R^2	0.650	0.659	0.819	0.856	0.866

[a]The dependent variable is the logarithm of real sales price (in 2010 dollar). The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. See Appendix for the definition of each variable and the list of variables included in house characteristics. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table 7: Price Regressions for Individual Buyers in 2005-2016 LA MLS-Deed Data[a]

	dependent variable: ln(sales price in 2010 dollars)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
cash	-0.031**	-0.030**	-0.037**	-0.036**	-0.038**	-0.031**	-0.020**	-0.036**	-0.039+
	(0.002)	(0.004)	(0.004)	(0.008)	(0.009)	(0.006)	(0.006)	(0.005)	(0.023)
tract×year×month	yes	yes	yes	yes	yes	yes	yes	yes	yes
buyer characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
seller characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
house characteristics	yes	yes	no	yes	yes	yes	yes	yes	yes
time-varying assessed value	yes	yes	yes	yes	yes	yes	yes	yes	yes
listing characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
house uniqueness	yes	yes	yes	yes	yes	yes	yes	yes	yes
house fixed effects	no	no	yes	no	no	no	no	no	no
buyer fixed effects	no	no	no	no	yes	no	no	no	no
seller fixed effects	no	no	no	no	no	no	yes	no	no
buyer's prior purchase price	no	no	no	no	no	no	no	yes	yes
IV estimation									yes
1st stage F-stat									172.84
observations	451964	36650	36650	14386	14386	21072	21072	31142	31142
adjusted R^2	0.866	0.885	0.953	0.851	0.894	0.893	0.926	0.877	0.669

[a]The dependent variable is the logarithm of real sales price (in 2010 dollar). The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. Table 2 provides the list of variables included in seller characteristics, buyer characteristics, and listing characteristics. See Appendix for the definition of each variable and the list of variables included in house characteristics. All columns include census tract×year×month fixed effects, and so if only one house (or buyer or seller) with two or more transactions is observed in a given census tract-year-month, that observation will be dropped from the estimation, which explains why the number of observations is much smaller in columns 2-9. Column 1 of this table is the same as column 5 of Table 6. Columns 1-9 include yearly assessment values available for 2005-2016. Columns 2-3 include only properties with repeated transactions during the sample period. Columns 4-5 include only buyers with two or more transactions during the sample period. Columns 6-7 include only sellers with two or more transactions during the sample period. Column 9 reports the instrumental variable regression where the instrument is the dummy for all-cash purchase in the buyer's previous transaction. The buyer's prior purchase price is the quality adjusted price (in 2010 dollar) in the buyer's previous transaction. Column 8 uses the same sample as column 9. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table 8: Time-on-the-Market Regressions for 2005-2016 LA MLS-Deed Data[a]

	dependent variable: time on market (#days)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
cash	-0.597	-0.758	-0.668	-0.698	0.770	0.424	-0.546
	(0.473)	(0.472)	(0.465)	(0.649)	(1.247)	(0.806)	(3.242)
tract×year×month	yes	yes	yes	yes	yes	yes	yes
buyer characteristics	yes	yes	yes	yes	yes	yes	yes
seller characteristics	yes	yes	yes	yes	yes	yes	yes
house characteristics	no	yes	yes	no	yes	yes	yes
time-varying assessed value	no	no	yes	yes	yes	yes	yes
listing characteristics	no	no	yes	yes	yes	yes	yes
house uniqueness	no	no	yes	yes	yes	yes	yes
house fixed effects	no	no	no	yes	no	no	no
buyer fixed effects	no	no	no	no	yes	no	no
seller fixed effects	no	no	no	no	no	yes	no
buyer's prior purchase price	no	no	no	no	no	no	yes
IV estimation	no	no	no	no	no	no	yes
observations	454840	454840	451964	36650	14386	21072	31142
adjusted R^2	0.131	0.135	0.152	0.072	0.040	0.276	0.032

[a]The dependent variable is the time-on-the-market which is the number of days from the original listing date to the agreement date. The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. Table 2 provides the list of variables included in seller characteristics, buyer characteristics, and listing characteristics. Note that listing characteristics in this table exclude TOM below the 5th (or above the 95th) percentile, given that the dependent variable is TOM. See Appendix for the definition of each variable and the list of variables included in house characteristics. Columns 3-7 include yearly assessment values available for 2005-2016. Columns 4 includes only properties with repeated transactions during the sample period. Columns 5 includes only buyers with two or more transactions during the sample period. Columns 6 includes only sellers with two or more transactions during the sample period. Column 7 reports the instrumental variable regression where the instrument is the dummy for all-cash purchase in the buyer's previous transaction. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table 9: Heterogeneities in Cash Discounts[a]

	dependent variable: ln(sales price in 2010 dollars)				
	(1)	(2)	(3)	(4)	(5)
cash	-0.033**	-0.038**	-0.053**	-0.038**	-0.039**
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
cash × cold season	-0.009**				
	(0.003)				
cash × hot season		0.005+			
		(0.003)			
cash × boom			0.027**		
			(0.003)		
cash × yearly zip code listings above 90th				0.010*	
				(0.005)	
cash × monthly zip code listings above 90th					0.015**
					(0.004)
tract×year×month	yes	yes	yes	yes	yes
buyer characteristics	yes	yes	yes	yes	yes
seller characteristics	yes	yes	yes	yes	yes
house characteristics	yes	yes	yes	yes	yes
time-varying assessed value	yes	yes	yes	yes	yes
listing characteristics	yes	yes	yes	yes	yes
house uniqueness	yes	yes	yes	yes	yes
observations	451964	451964	451964	451964	451964
adjusted R^2	0.846	0.846	0.846	0.846	0.846

[a]The dependent variable is the logarithm of real sales price (in 2010 dollar). The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. Cold season (or hot season) is the indicator variable for the period from September until January (or from May until August). Boom is the dummy for housing boom periods in LA. Yearly zip code listings above 90th is the dummy for whether the number of listings (from the MLS) in the same ZIP code as a given property during the same year when it was sold is above the 90th percentile. Monthly zip code listings above 90th is the dummy for whether the number of listings (from the MLS) in the same ZIP code as a given property during the same month when it was sold is above the 90th percentile. Table 2 provides the list of variables included in seller characteristics, buyer characteristics, and additional controls. See Appendix for the definition of each variable and the list of variables included in house characteristics. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

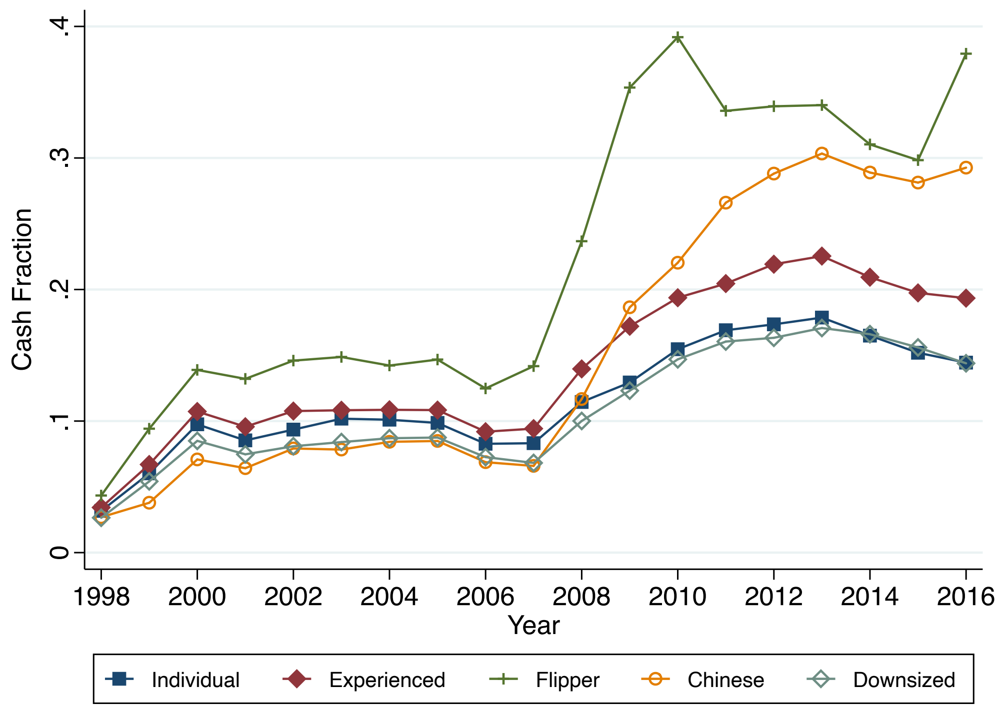
Online Appendix to

"Cash is King? Understanding Financing Risk in Housing Markets"

Lu Han
Wisconsin School of Business
University of Wisconsin – Madison
lu.han@wisc.edu

Seung-Hyun Hong
Department of Economics
University of Illinois
hyunhong@illinois.edu

Figure A1: Cash Fraction Among Buyer Group in the National Sample w/o LA[a]



[a]Source: CoreLogic. The figure plots the fraction of all-cash purchases among each buyer group. The national sample without LA includes transactions from the top 100 U.S. cities, excluding Los Angeles. Different buyer groups are defined as follows: individual refers to individual buyers; experienced buyer indicates buyers who purchased any house in their county in the past; downsized buyer is the buyer whose previous house has more bedrooms, more bathrooms, and larger building square footage than the current house (since downsized buyers must have purchased houses before, they are also experienced buyers); flipper buyer is the buyer who sold the house within two years after purchasing the house; Chinese buyer means that the last name of the house owner belongs to the list of Chinese last names.

Table A1: Summary Statistics for the National Sample w/o LA[a]

	All (1)	Cash (2)	Mortgage (3)
sales price (in 2010 dollar)	\$276,889	\$278,154	\$276,691
building square footage	1932	1948	1930
effective year built	1987	1989	1987
#bedrooms	2.64	2.29	2.70
#total rooms	3.29	2.55	3.41
#bathrooms	2.24	2.08	2.27
#parking spaces	0.34	0.29	0.35
single family house	0.83	0.68	0.86
duplex	0.01	0.02	0.01
condo	0.12	0.22	0.11
institutional buyer	0.046	0.214	0.020
individual buyer	0.954	0.786	0.980
experienced buyer	0.213	0.216	0.212
flipper	0.072	0.094	0.068
Chinese	0.023	0.028	0.022
observations	18701950	2539395	16162555

[a]The table reports the mean values of transaction-specific variables and house characteristics, as well as dummies for buyer types. The sample includes arms-length transactions of residential properties in the top U.S. cities without LA from 1998 to 2016, excluding sales of foreclosed properties. Column 2 reports the mean of each variable among all-cash transactions, whereas column 3 reports the mean among mortgage transactions. Sales price is real transaction price, deflated by Consumer Price Index. Effective year built is the first year when the building was assessed with its current components. Experienced, downsized, flipper, and Chinese buyers are defined in Figure 4's footnote.

Table A2: Cash Regressions for the National Sample w/o LA[a]

	dependent variable: dummy for cash purchase	
	(1)	(2)
institutional buyer	0.432**	0.429**
	(0.002)	(0.002)
experienced	0.024**	0.026**
	(0.000)	(0.000)
flipper	0.055**	0.055**
	(0.001)	(0.001)
Chinese	0.058**	0.057**
	(0.001)	(0.001)
tract×year×month fixed effects	yes	yes
house characteristics	no	yes
assessed value in 2016	no	yes
observations	18109436	18109436
adjusted R^2	0.232	0.232

[a]The dependent variable is the dummy for whether properties are purchased by all-cash. The sample includes arms-length transactions of residential properties in the top U.S. cities without LA from 1998 to 2016, excluding sales of foreclosed properties. See Appendix for the definition of each variable and the list of variables included in house characteristics. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table A3: Buyer's Prior Cash Purchase Regressions[a]

	dependent variable: buyer's prior cash purchase			
	(1)	(2)	(3)	(4)
ln(#prior transactions of buyer)	0.035**	0.035**	0.035**	0.036**
	(0.003)	(0.003)	(0.003)	(0.003)
downsized	0.008+	0.009	0.009	0.013+
	(0.005)	(0.006)	(0.006)	(0.007)
flipper	0.043**	0.043**	0.042**	0.037**
	(0.007)	(0.007)	(0.007)	(0.008)
Chinese	0.074**	0.074**	0.071**	0.072**
	(0.006)	(0.006)	(0.006)	(0.007)
ln(buyer's prior purchase price)	-0.031**	-0.031**	-0.031**	-0.030**
	(0.003)	(0.003)	(0.003)	(0.004)
experienced seller		0.004	0.003	0.003
		(0.005)	(0.005)	(0.006)
tract-year log price			0.032*	
			(0.015)	
tract-year cash fraction			0.157**	
			(0.026)	
tract×year×month	no	no	no	yes
house characteristics	yes	yes	yes	yes
time-varying assessed values	yes	yes	yes	yes
listing characteristics	yes	yes	yes	yes
house uniqueness	yes	yes	yes	yes
observations	31142	31142	31142	31142
adjusted R^2	0.047	0.047	0.048	0.054

[a]The dependent variable is the buyer's prior cash purchase, which is the dummy for all-cash purchase in the buyer's previous transaction. The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. See Appendix for the definition of each variable and the list of variables included in house characteristics. All columns use only buyers with any previous transactions during 1990-2016, and exclude buyers who purchased their previous house recently (within a year) or whose previous house was nearby (located within 10 miles from their current house). Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table A4: Price Regressions for Individual Buyers in 2005-2016 LA MLS-Deed Data with TOM as a Control[a]

	dependent variable: ln(sales price in 2010 dollars)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
cash	-0.031**	-0.030**	-0.038**	-0.037**	-0.037**	-0.031**	-0.020**	-0.036**	-0.038+
	(0.002)	(0.004)	(0.004)	(0.008)	(0.009)	(0.006)	(0.006)	(0.005)	(0.022)
TOM/100	-0.027**	-0.030**	-0.023**	-0.028**	-0.013*	-0.017**	-0.008*	-0.029**	-0.017**
	(0.001)	(0.002)	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)	(0.002)
tract×year×month	yes	yes	yes	yes	yes	yes	yes	yes	yes
buyer characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
seller characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
house characteristics	yes	yes	no	yes	yes	yes	yes	yes	yes
time-varying assessed value	yes	yes	yes	yes	yes	yes	yes	yes	yes
listing characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
house uniqueness	yes	yes	yes	yes	yes	yes	yes	yes	yes
house fixed effects	no	no	yes	no	no	no	no	no	no
buyer fixed effects	no	no	no	no	yes	no	no	no	no
seller fixed effects	no	no	no	no	no	no	yes	no	no
buyer's prior purchase price	no	no	no	no	no	no	no	yes	yes
IV estimation	no	no	no	no	no	no	no	no	yes
observations	451964	36650	36650	14386	14386	21072	21072	31142	31142
adjusted R^2	0.867	0.886	0.954	0.852	0.894	0.894	0.926	0.877	0.671

[a]The dependent variable is the logarithm of real sales price (in 2010 dollar). The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. We use TOM/100 which is the time on market divided by 100, so that the coefficient on TOM can be shown in three decimal places. Table 2 provides the list of variables included in seller characteristics, buyer characteristics, and listing characteristics. See Appendix for the definition of each variable and the list of variables included in house characteristics. Column 1 of this table is the same as column 5 of Table 6. Columns 1-9 include yearly assessment values available for 2005-2016. Columns 2-3 include only properties with repeated transactions during the sample period. Columns 4-5 include only buyers with two or more transactions during the sample period. Columns 6-7 include only sellers with two or more transactions during the sample period. Column 9 reports the instrumental variable regression where the instrument is the dummy for all-cash purchase in the buyer's previous transaction. Column 8 uses the same sample as column 9. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table A5: Price Regressions for Individual Buyers in 2005-2016 LA MLS-Deed Data with TTC as a Control[a]

	dependent variable: ln(sales price in 2010 dollars)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
cash	-0.031**	-0.031**	-0.038**	-0.036**	-0.040**	-0.031**	-0.020**	-0.036**	-0.041+
	(0.002)	(0.004)	(0.004)	(0.008)	(0.009)	(0.006)	(0.006)	(0.005)	(0.024)
TTC/100	-0.002	-0.009	-0.006	0.002	-0.027*	0.001	-0.006	-0.010	-0.011*
	(0.002)	(0.005)	(0.006)	(0.010)	(0.011)	(0.006)	(0.006)	(0.006)	(0.005)
tract×year×month	yes	yes	yes	yes	yes	yes	yes	yes	yes
buyer characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
seller characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
house characteristics	yes	yes	no	yes	yes	yes	yes	yes	yes
time-varying assessed value	yes	yes	yes	yes	yes	yes	yes	yes	yes
listing characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
house uniqueness	yes	yes	yes	yes	yes	yes	yes	yes	yes
house fixed effects	no	no	yes	no	no	no	no	no	no
buyer fixed effects	no	no	no	no	yes	no	no	no	no
seller fixed effects	no	no	no	no	no	no	yes	no	no
buyer's prior purchase price	no	no	no	no	no	no	no	yes	yes
IV estimation	no	no	no	no	no	no	no	no	yes
observations	451964	36650	36650	14386	14386	21072	21072	31142	31142
adjusted R^2	0.866	0.885	0.953	0.851	0.894	0.893	0.926	0.877	0.669

[a]The dependent variable is the logarithm of real sales price (in 2010 dollar). The sample includes arms-length transactions of residential properties from 2005 to 2016 in Los Angeles County transaction data merged with Multiple Listing Service data, excluding sales of foreclosed properties and institutional buyers. We use TTC/100 which is the time-to-close divided by 100, so that the coefficient on TTC can be shown in three decimal places. Table 2 provides the list of variables included in seller characteristics, buyer characteristics, and listing characteristics. See Appendix for the definition of each variable and the list of variables included in house characteristics. Column 1 of this table is the same as column 5 of Table 6. Columns 1-9 include yearly assessment values available for 2005-2016. Columns 2-3 include only properties with repeated transactions during the sample period. Columns 4-5 include only buyers with two or more transactions during the sample period. Columns 6-7 include only sellers with two or more transactions during the sample period. Column 9 reports the instrumental variable regression where the instrument is the dummy for all-cash purchase in the buyer's previous transaction. Column 8 uses the same sample as column 9. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table A6: Price Regressions for Individual Buyers in 1998-2016 LA Deed Data[a]

	dependent variable: ln(sales price in 2010 dollars)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
cash	-0.047**	-0.046**	-0.050**	-0.068**	-0.059**	-0.036**	-0.023**	-0.051**	-0.062*
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.028)
tract×year×month	yes	yes	yes	yes	yes	yes	yes	yes	yes
house characteristics	yes	yes	no	yes	yes	yes	yes	yes	yes
assessed value in 2016	yes	yes	no	yes	yes	yes	yes	yes	yes
buyer characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
seller characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
time-varying assessed value	no	no	no	no	no	no	no	no	no
listing characteristics	no	no	no	no	no	no	no	no	no
house uniqueness	no	no	no	no	no	no	no	no	no
house fixed effects	no	no	yes	no	no	no	no	no	no
buyer fixed effects	no	no	no	no	yes	no	no	no	no
seller fixed effects	no	no	no	no	no	no	yes	no	no
buyer's prior purchase price	no	no	no	no	no	no	no	yes	yes
IV estimation									yes
1st stage F-stat									234.23
observations	1260285	563577	563577	302478	302478	228841	228841	77466	77466
adjusted R^2	0.887	0.895	0.943	0.892	0.904	0.915	0.938	0.904	0.366

[a]The dependent variable is the logarithm of real sales price (in 2010 dollar). The sample includes arms-length transactions of residential properties from 1998 to 2016 in Los Angeles County, excluding sales of foreclosed properties and institutional buyers. Columns 2-3 include only properties with repeated transactions during the sample period. Columns 4-5 include only buyers with two or more transactions during the sample period. Columns 6-7 include only sellers with two or more transactions during the sample period. Columns 8-9 include only buyers with any previous transaction during 1990-2016, but exclude buyers who purchased their previous house within a year or whose previous house was located within 10 miles from their current house. All columns include census tract×year×month fixed effects. Column 9 reports the instrumental variable regression where the instrument is the dummy for all-cash purchase in the buyer's previous transaction. The buyer's prior purchase price is the quality adjusted price (in 2010 dollar) in the buyer's previous transaction. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.

Table A7: Price Regressions for the National Sample in 1998-2016 Deed Data, including Institutional Buyers[a]

	dependent variable: ln(sales price in 2010 dollars)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
cash	-0.059**	-0.057**	-0.066**	-0.088**	-0.118**	-0.060**	-0.029**	-0.058**	-0.046*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.005)	(0.021)
tract×year×month	yes	yes	yes	yes	yes	yes	yes	yes	yes
house characteristics	yes	yes	no	yes	yes	yes	yes	yes	yes
assessed value in 2016	yes	yes	no	yes	yes	yes	yes	yes	yes
buyer characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
seller characteristics	yes	yes	yes	yes	yes	yes	yes	yes	yes
time-varying assessed value	no	no	no	no	no	no	no	no	no
listing characteristics	no	no	no	no	no	no	no	no	no
house uniqueness	no	no	no	no	no	no	no	no	no
house fixed effects	no	no	yes	no	no	no	no	no	no
buyer fixed effects	no	no	no	no	yes	no	no	no	no
seller fixed effects	no	no	no	no	no	no	yes	no	no
buyer's prior purchase price	no	no	no	no	no	no	no	yes	yes
IV estimation									yes
observations	19851526	11458194	11458194	4015229	4015229	6391079	6391079	136893	136893
adjusted R^2	0.868	0.874	0.918	0.847	0.874	0.852	0.908	0.883	0.235

[a]The dependent variable is the logarithm of real sales price (in 2010 dollar). The sample includes arms-length transactions of residential properties from 1998 to 2016 in the top 100 U.S. cities, excluding Los Angeles. The sample also includes both individual buyers and institutional buyers. The sample excludes sales of foreclosed properties. Columns 2-3 include only properties with repeated transactions during the sample period. Columns 4-5 include only buyers with two or more transactions during the sample period. Columns 6-7 include only sellers with two or more transactions during the sample period. Columns 8-9 include only buyers with any previous transaction during 1990-2016, but exclude buyers who purchased their previous house within a year or whose previous house was located within 10 miles from their current house. All columns include census tract×year×month fixed effects. Column 9 reports the instrumental variable regression where the instrument is the dummy for all-cash purchase in the buyer's previous transaction. The buyer's prior purchase price is the quality adjusted price (in 2010 dollar) in the buyer's previous transaction. Robust standard errors clustered at the census tract level in parentheses. + denotes significance at a 10% level, * denotes significance at a 5% level, and ** denotes significance at 1% level.