iBuyers: Liquidity in Real Estate Markets?

Greg Buchak, Gregor Matvos, Tomasz Piskorski, and Amit Seru*

This Version: OCTOBER 2020

DO NOT CIRCULATE

We study the benefits and limits of dealer-intermediation in residential real estate through the window of "iBuyers," real estate companies with online acquisition platforms that buy and sell residential real estate. iBuyers provide liquidity by purchasing houses almost immediately. This allows households to bypass a lengthy listing process. In exchange, they buy at a 3.6% discount relative to comparable homes and earn a significant spread when selling. We show that iBuyers rely on a remote valuation technology and that they avoid transacting in houses where this valuation technology is less precise in expectation, suggesting that adverse selection is an important consideration for dealers in this market. Armed with these new facts, build a structural search model featuring adverse selection to quantify the economic forces underlying the tradeoffs of dealer intermediation in residential real estate. Our model suggests that dealers' primary value to homeowners is the enhanced speed of transaction. Additionally, adverse selection is a major impediment to intermediation, and counterfactually endowing intermediaries with a perfect valuation technology would quadruple the fraction of dealer intermediated transactions. The fact that intermediated houses are left unoccupied during the transaction is not quantitatively important. We next use our model to show that iBuyer technology, as opposed to merely having a balance sheet, is key in their growth. Finally, our model highlights an important limitation of iBuyers and real estate intermediation more generally: that while they provide liquidity, they mostly provide it in already liquid markets because their business model relies on reselling the houses quickly.

Keywords: Fintech, Proptech, iBuyers, Real Estate Market, Housing, Liquidity

---

* We thank Peter DeMarzo, Jeff Zwiebel and participants at Kellogg, Stanford, Columbia Workshop in New Empirical Finance, Duke Fuqua, NBER Summer Institute and SITE for useful comments. Piskorski and Seru thank the National Science Foundation Award (1628895) for financial support. Buchak is at Stanford Graduate School of Business (GSB), Matvos is at Northwestern University and the National Bureau of Economic Research (NBER), Piskorski is at Columbia University and NBER, and Seru is at Stanford GSB, the Hoover Institution, the Stanford Institute for Economic Policy Research (SIEPR), and NBER. *First Draft: November 2019.*

I. Introduction

Residential real estate, valued at more than $30 trillion, is the main asset of US households, accounting for about 70% of median household net worth.[1] Housing therefore combines a household's choice of habitation with a substantial investment of their wealth in a specific asset, two choices, which do not necessarily go hand in hand. Yet despite its central importance, the process of buying and selling homes is subject to significant frictions, making housing assets quite illiquid. For example, both buyers and sellers may have to incur expensive costs during the search process to buy or to sell a home. For sellers, it is costly to keep their homes on the market, while buyers may incur considerable pecuniary and non-pecuniary costs when visiting properties. Additionally, it is often difficult to align the timing of a home sale with the purchase of a new home or a move to a location. Consequently, the ability to trade homes and adjust housing consumption has first-order implications for overall economic activity, including household consumption and savings decisions, mobility, and employment.

These difficulties in homeowner-to-homeowner sales suggest a natural role for dealer intermediation: A selling homeowner could sell directly to an intermediary that carries the house on its balance sheet and resells once it finds a buyer. Yet, until recently, such transactions were rate. This has changed with the entry of iBuyers, a recent technological disruption in the real estate market, to provide novel evidence on the benefits and challenges of increasing liquidity in housing markets. iBuyers, such as Opendoor, RedfinNow, and Zillow Offers, are real estate companies that use an automated valuation model (known as an AVM) and other technology to make cash offers on homes quickly through their on-line acquisition platforms. Critically, they make an offer to purchase homes, often within hours of a seller asking for one. In other words, iBuyers promise to increase liquidity in housing markets.

In this paper, we use iBuyers as a window to study the tradeoffs in dealer intermediation in residential real estate, and the role that their technology has played in making dealer intermediation viable. We highlight three key economic forces surrounding selling homeowners and their interactions with dealers. The first force is speed: While a homeowner selling through traditional channels must go through an often slow listing process, the dealer intermediated transaction can happen much faster from the perspective of the seller. This speed is presumably valuable to at least a segment of the population. The second force is the utilization of the asset during the transaction. While a selling homeowner remains in possession of the house during the selling process, the dealer does not. Thus, potentially valuable service flows obtained from the house are lost in a dealer intermediated transaction, pushing against this form of transaction. The third force is adverse selection. There are many aspects of a house's quality that are observable to individual

[1] See US Census Bureau https://www.census.gov/data/tables/2016/demo/wealth/wealth-asset-ownership.html.

buyers and sellers but are potentially difficult for intermediaries to observe. For example, the noise of the neighborhood, the quality of the natural light, and so on, which can make small but important differences to the house's valuation. An intermediary in this market potentially has to deal with adverse selection, where only homeowners with lower-quality homes are willing to sell through the intermediary.

This paper proceeds in two main steps. We first use rich micro-level transaction and listing data to document new facts about iBuyers' business model. Second, armed with these new facts, we build and calibrate a realistic structural search model that explicitly incorporates these economic forces. We use the calibrated model to quantify their importance in residential real estate intermediation, which sheds light more broadly on why intermediation has been difficult in this market. We next use the model to disentangle the precise role of iBuyers' technology, and ask whether a "low-tech" dealer would be able to survive in this market. Finally, we use the model to study the limits to liquidity provision by intermediaries more broadly.

We start our analysis by documenting a significant growth in the iBuyer market share since their entry in 2014. For example, in Phoenix, Arizona, iBuyers market share had grown from less than 1% in 2015 to about 6% of all real estate transactions by 2018. We then turn our attention to the iBuyer business model.

In a standard housing transaction, the seller is a homeowner who currently occupies the house, and the buyer plans to occupy the house upon purchase. Traditional brokers help connect buyers and sellers. In consequence, this transaction requires matching a seller, who is ready to leave, with a buyer who is ready to move in at roughly the same time. The average time for a listed house to sell is roughly 90 days. In contrast, in a non-standard transition, a buyer (an intermediary) with a balance sheet can purchase the house from a seller as soon as the seller is ready to move and carry the house in inventory until the appropriate occupant-buyer appears. We demonstrate that iBuyers follow such business model and act as liquidity providers in the housing market: Sellers to iBuyers are significantly more likely to forgo the listing process and sell quickly and directly to iBuyers.

iBuyers purchase houses and hold them in inventory for only short-period of time: about a half of homes they buy are sold within three months and about three-quarter of homes within six months from acquiring. iBuyers acquire homes very quickly and majority of homeowners who sell to them do not list their houses. On the other hand, the iBuyers use traditional selling channels relying on multiple listing services to dispose their inventory.

iBuyers earn a positive spread (gross return) of about 5% on the houses they buy and sell. A large portion of this spread can be accounted for by a substantial discount at which iBuyers purchase homes from homeowners amounting on average to about 3.6% of home value. This estimate

compares houses purchased by iBuyers to those purchased by others in the same zip code, at the same point in time, and with the same characteristics. Overall, an average seller takes a discount of approximately $9,000 when selling to an iBuyer. Compared to other sellers with similar homes, iBuyers list the properties they acquire at higher prices, adjust listing prices more frequently, and are more patient. In doing so, on the sell-side, they achieve a selling premium averaging about 1.6% of home value. These facts suggest that one key aspect of iBuyers' business model is their ability to hold inventory on their balance sheet and trade more "strategically" than other sellers.

iBuyers do not make purchase offers for all homes and instead only focus on certain types of houses. We first show that iBuyers expansion was focused in homes that can be relatively accurately valued remotely (algorithmically). To demonstrate it we follow the strategy in Buchak et al. (2018) and ask how much variation in prices observable characteristics of the property and local market conditions can explain. Across specifications, characteristics explain a substantially higher share of iBuyer transaction prices, both purchases and sales, than those of other market participants. With zip times quarter fixed effects capturing local time-varying market trends, observable house characteristics explain over 80% of the variation in prices for iBuyer transactions, versus 68% of the variation in non-iBuyer transactions. These results suggests that iBuyers focus on homes that can be relatively accurately priced algorithmically using observable housing characteristics. This observation could help explain why iBuyers grew from businesses, which had previously been focused on collecting and analyzing house price data, such as Zillow.

We next show, using the pre-iBuyer entry empirical measures for housing liquidity, that iBuyers also predominately intermediate houses that are relatively more liquid. Houses with a "predicted probability" of selling within 3 months from a listing date – based on using measured housing characteristics -- of between 10-20% see close to zero iBuyer market share. In contrast those with a 70-80% predicted probability of selling within 3 months see an iBuyer market share of roughly 1.5% during our sample period. Overall, these results suggest that iBuyers predominantly buy homes that can be fairly accurately valued using (algorithmic) automatic valuation models and that can be quickly sold (i.e., are already fairly liquid).

If households are willing to pay a large premium for liquidity in the housing market, why is intermediation so limited? In other words, why have intermediaries like iBuyer not overtaken the market yet? Given the large premium households are willing to pay, illiquidity of real estate is a first order friction facing households, which intermediaries should help resolve. To investigate these questions, we develop a search-based equilibrium model of house trading with iBuyers. The model highlights when and how intermediaries such as iBuyers play a role in housing markets. We use this model, calibrated to data, to make sense of our empirical findings and conduct a set of

counterfactuals. These allow us to understand the impact of iBuyers on consumers and the housing market, the sources of their technological advantages, and limits to their liquidity provision.

We model the decision problem of iBuyers and individual homeowners as they buy, sell, and change homes. A homeowner is initially matched with a house from which she receives a flow utility benefit of housing services and pays a maintenance cost. With some probability, she becomes unmatched from her current house and begins the process of moving. We assume that the homeowner's balance sheet is constrained in the sense that she can only own one house at a time. In consequence, she must effect the move sequentially: she first attempts to sell her current house, and then looks to buy a new house. Once she has bought her new house, she once again becomes a matched homeowner.

We introduce a second type of agent: iBuyers. iBuyers can purchase homes from selling homeowners and list them again for resale. They differ from homeowners along three dimensions: First, they are not constrained to hold only one house at a time, and so their decision problem does not require that they sell a house they currently possess before trying to buy a new house. Second, iBuyers do not live in the house when trying to sell it. This means that they do not obtain utility flows from homeownership although they still must pay maintenance costs. Additionally, because nobody lives in the house, they potentially face maintenance costs that increase over time, representing, for example, the possibility of unobserved roof damage that leads to severe rain damage. Third, iBuyers possess a pricing technology that allows them to purchase almost *immediately* from selling homeowners, but it is imperfect in the sense that they may fail to identify costs associated with purchasing the home that homeowners themselves would observe.

To make quantitative statements with the model, we first calibrate several parameters externally in relation to existing literature, parameters that map directly to observable quantities, or through normalizations. We then use the method of moments to calibrate the remaining parameters by searching for parameters that map model-implied moments to moments we observe in the data. Our model matches the targeted moments based on 2018 data fairly well, with near exact matches for prices, iBuyer market shares, iBuyer purchase discount, time on market, relative time-on-market for iBuyers when selling their inventory, and the two sensitivities of market shares with respect to liquidity and pricing error. To provide external validation of the model, we bring the model's equilibrium predictions to untargeted reduced form moments, and show that the model predicts that iBuyer market share corresponds to decreases in equilibrium prices and increases in transaction speeds. Quantitatively, these elasticities match closely with reduced form evidence.

We first use the model to quantify the economic importance of the three key economic forces surrounding dealer intermediation in residential real estate. We should that speed is a critical factor, and that if iBuyers took as long as the typical listed sale, they would essentially vanish. The fact

that the house is unoccupied (and service flows unused) is quantitatively unimportant, driven primarily by the fact that while iBuyers do not use the house, the owners who are looking to move do not receive large utility flows from being in a house they want to move away from. We find that information asymmetries play a large role, and show that removing the information asymmetry between the seller and the iBuyer would increase iBuyer market share fourfold.

We next use the model to shed light on the source of iBuyers' technological advantage. We note that iBuyers differ from homeowners, because they derive no intrinsic value when holding on to the house. iBuyers therefore face a disadvantage relative to homeowners when it comes to holding real estate. They therefore need some other advantage to be willing to engage in intermediation. In particular, are iBuyers merely an intermediary with a balance sheet, or does their technology provide distinct benefits? We examine three aspects of iBuyer potential technological advantage over a potential intermediary with a balance sheet capacity: their valuation speed, their valuation accuracy, and their matching technology when selling.

Our analysis shows that transaction speed and valuation accuracy are the key elements of technological advantage of iBuyers over other intermediaries. The low-tech valuation intermediary with just balance-sheet capacity would counterfactually achieve 1-2% market share, in contrast to the iBuyer, which has five times that market share. This analysis also helps us understand why iBuyers are predominately active among homes that can be fairly accurately priced using (algorithmic) automatic valuation models. Among these homes their technological advantage is the highest, which is a key factor behind their significant market share. Among homes that are difficult to price, our model predicts a low market penetration, which is exactly what data reveals.

Finally, we examine the extent to which, and in what contexts, iBuyers are able to provide liquidity. A naïve prediction might be that iBuyers provide the greatest liquidity benefit in markets that are illiquid before iBuyers enter. However, the data show that iBuyers focus on the most liquid markets. The model provides several insights that highlights limits to the liquidity provision in the real estate market and the expansion of iBuyers. A key limitation in the intermediation of housing assets is that in contrast to a homeowner, the intermediary keeps the house vacant while listing and therefore does not derive the consumption benefit that the owner-occupant would. Additionally, empty houses face the risk of accelerated depreciation when they are not subject to constant upkeep. This acts as another disadvantage to iBuyers, particularly when the market is illiquid and the iBuyer expects the house to remain on market for some time. Finally, because iBuyers transact in houses quickly when approached by sellers, they need to be able to value the house quickly and not be subject to severe adverse selection. If houses are difficult to price, the scope for adverse selection is larger, suggesting iBuyers should focus on easy to price houses.

Overall, our model highlights why liquidity provision in real estate markets has been limited, despite its high potential benefits. Illiquid homes make it relatively more efficient for the seller to live in the house during the sale process rather than keep the house vacant exposing it to adverse maintenance shocks. Moreover since liquidity providers are subject to adverse selection, they need to focus on homes that they can quickly and fairly accurately value using their data-based statistical models. Liquidity provision is therefore efficient only when houses are already relatively easy to value and liquid. This ensures that such homes can be acquired and resold quickly while limiting the scope for adverse selection and adverse maintenance shocks. Our analysis suggests that iBuyers, with their current use of technology are not likely to impact a large part of the market – i.e., they will make already liquid and easy to value houses more liquid, rather than unlocking the sale of illiquid and harder to price homes. In other words, iBuyers comparative advantage allows them to add liquidity to the market, but in pockets where liquidity is the least valuable.

Our paper is related to the recent work focusing on the technological disruptions in the real estate marketplace. Most of this emerging literature has focused on the role of on-line fintech lenders (Buchak et al. 2018; Fuster et al. 2019) and the impact of sharing economy on the housing market (e.g., Calder-Wang 2019). We contribute to this literature by studding the emergence of iBuyers and their role in improving liquidity in the housing market. In doing so we also expand recent work analyzing the role of non-bank intermediaries in the housing market (e.g., Buchak et al. 2020; Jiang et al. 2020). Second, our work is related to the literature that studies how search-based frictions and other factors affect the housing trading market (e.g., Wheaton 1990; Genesove and Mayer 1997; Levitt and Syverson 2008; Piazessi and Schneider 2009, Anenberg and Bayer 2020; Guren 2018; Piazessi, Schneider, and Stroebel 2020). We contribute to this literature by studding how iBuyer technology may help alleviate some of these frictions by making housing markets more liquid. More broadly, our analysis of the intermediation in the housing market is also related to a large body of work focusing on intermediation and trading frictions in the decentralized asset markets (e.g., Grossman and Miller 1988; Duffie, Garleanu, and Pedersen 2005; Gavazza 2016).

II. Data and Institutional Background

II.A. Data Sources

Transactions Data: We use Corelogic deeds records data on housing transactions from five markets with large iBuyer presence as of 2018: Phoenix, Las Vegas, Dallas, Orlando, and Gwinnet County, a suburb of Atlanta. We use data between 2013 and 2018 and restrict the sample to arms-length, non-foreclosure transactions in single family homes or condominiums with transaction prices below $10 million and land footage below 50,000 sq. feet. The data reports each transaction tagged to a specific property, with seller name, owner name, transaction date, sale amount, and mortgage amount. Transactions without a recorded sale date are excluded. Merging these

transaction records with tax assessment files enables us to observe property-specific attributes, including the census tract, land square footage, building square footage, the number of stories, the year of construction, the type of air conditioning, garage, heating, sewer, water, and electricity. The assessment file also includes evaluations of the construction quality and location desirability. Table 1 Panel A provides summary statistics for this data.

Listings data: We use listings data from the Multiple Listings Service (MLS) provided by ATTOM Data. The data spans 2010 through 2018, and our main sample period is 2013 and 2018. Individuals, brokers, and companies selling their properties post listings on a set of common platforms, and we observe the combined data. The data is at the individual listing level. That is, for a given attempt to sell a house, the lister will make an initial listing with an asking price. As time passes, the seller may find an interested buyer at that price, or she may amend her listing with the different (typically lower) price, in order to attract other buyers. The listings data contain similar house-level information as well as the identifying information of the homeowner as the transaction-level data, listing agent, and buying agent. We aggregate the data to a "listing-spell," which captures a single period over which a homeowner, whether an individual or an iBuyer, attempts to sell her house. Each listing spell may contain multiple amendments and price changes, which we summarize for each transaction. Table 1 Panel B provides summary statistics for this data.

Redfin and Zillow: We use publicly available data from Redfin, which includes at the zip code level, the fraction of listings that sell within two weeks of listing, the average sale-price-to-list-price, and the average sale price. Additionally, we use house price indices from Zillow in robustness checks, which provide quality-adjusted transaction prices at a zip-quarter level.

Other data: We use the American Community Survey (ACS) data from the U.S. Census Bureau to measure several zip-level demographic characteristics including median income, median age, fraction of adults with a bachelor's degree or higher, population, fraction of the population that is white, and fraction of the population that pays over 50% of their disposable income on rent.

II.B. Classification of Buyers and Sellers

II.B.1 Identifying iBuyers

We classify parties to the transaction---buyers and sellers---as iBuyers or not. In our analysis, a party is an iBuyer if it is one of Opendoor, Offerpad, Knock, Zillow, and Redfin. These parties are not always recorded in the data as "Opendoor" or "Offerpad," but rather as a specific business entity such as "OD ARIZONA D LLC." "OD ARIZONA D LLC" is, in fact, a private company registered to the same address as Opendoor. To ensure our classification covers these separate

corporate entities, we use a set of regular expression search terms to make sure that the most common parties are included. Appendix A.2 details this classification procedure. The Corelogic transaction data captures more than 6,000 iBuyer purchases and sales.

II.B.2 Tracking Individuals over Time

From the Corelogic transaction records, we create a panel at the individual-year level, of household ownership and geographical location. From the baseline dataset of transactions between 2000 and 2018, we begin by extracting a list of unique names at the market level. To construct these names, we first remove all names with corporate or business markers[2] and absentee owners, and impose a series of cleaning and filtering steps.[3]

We next take the baseline transaction records and extract the last purchase transaction for each property for each year. We then define a person to be owning a house if at year *t*, he or she was the buyer in the most recent transaction in all years prior to *t* and he or she is not recorded as selling the house prior to year *t*. If by this methodology we identify a particular person as living in more than one house in a given year, we break ties in favor of the house with the latest sale date. Finally, if there are still duplicates, we choose the property that appears first in the dataset.[4] For each name in each market, this methodology yields at most a single owned home per year. However, there are years for which an individual will not own a house. We retain these years in order to keep the panel balanced; these missing years can be interpreted either as the individual living in a different market, or as the individual renting or living with another homeowner.

Putting this together, the panel provides key details of homeowner's behavior: We observe sales directly from deeds records. Then, conditional on selling her house, the homeowner either buys a new house in the same market or leaves the market entirely. Conditional on buying a new house, we measure changes in house size and leverage. Leaving the market means that the homeowner either owns a house in a new geographical area, or becomes a renter (in any geographical area). Thus, we observe not only sales, but also the homeowner's subsequent actions in terms of mobility and house consumption.

II.C. The Rise of iBuyers

iBuyers began significant growth in 2015 in Phoenix and between 2016 and 2018 in Las Vegas, Orlando, Gwinnet County, Georgia, and Dallas (Figure 1 Panel (a)). iBuyers had roughly 1%

[2] For example, "Corp," "INC," "LLC," and so on.
[3] We remove white space, names where the first or last name has fewer than one character, and names that appear on more than ten unique purchase transactions in the data set.
[4] Under this methodology, before this final step, 0.8% of homeowners appear to be living in more than one house in a given year.

market share in Phoenix in 2015; by 2018 this had grown to roughly 6%. Similar striking growth has occurred in Gwinnet (4%), Las Vegas (3%), Orlando (~2%), and Dallas (~2%).

iBuyers focus on a relatively narrow band of homes. As Figure 2 shows, they typically buy houses that are in the $100k - $250k price range, relatively new, of modest property (lot) size, and multistory. With respect to demographics, Figure 3 shows that they have the greatest market share in zip codes with younger, middle-class individuals: Those with median incomes between $70,000 and $90,000 (Panel (a)); those with average ages 30 and below (Panel (b)); those where residents possess bachelor's degrees (Panel (c)).

As their activity expanded, iBuyers inventory of houses expanded as well, both in terms of numbers of houses and dollar value (based on purchase price) of their houses (Figure 1, Panels (b) and (c)). By the end of 2018, iBuyers had roughly 1,500 houses in inventory, with a combined purchase price of roughly $350 million. iBuyers in 2018 purchased between 400 and 500 houses per quarter (Panel (d)). Their inventory turnover, defined as the number of sales in a quarter divided by their total inventory, is typically between 0.3 and 0.6 per quarter (Panel (e)). iBuyers typically hold inventory for a very short period of time, holding a house for a median period of roughly 100 days.

Based on completed transactions, iBuyers typically earn roughly a 5% spread between purchase and sale -- defined as the difference between the price at which they sell and price at which they buy, as a percentage of the acquisition price. The spread has been consistently positive over time, and the 25-75th percentile of realized spread on a per-house basis has been also positive for all but two quarters in 2015 (Figure 1 Panel (f)).

III. iBuyers' Business Model and Liquidity Provision

In a standard housing transaction, the seller is a homeowner who currently occupies the house, and the buyer plans to occupy the house upon purchase. Houses are advertised though listings and brokers connect buyers and sellers. This transaction requires matching a seller, who is ready to leave, with a buyer who is ready to move in at roughly the same time.[5] A natural alternative is a dealer-intermediary, which purchases the house immediately when the seller wants to move. The dealer holds the house in inventory and sells when the appropriate occupant-buyer appears. As we show above, iBuyers follow such a model: they purchase houses, hold them in inventory, and sell them, earning a spread.

[5] While an individual homeowner may be willing to temporarily own two houses to facilitate moving into a new home, such an activity requires a substantial amount of wealth on their part. This is not typical of most individual transactions.

III.A. iBuyers' Transaction Speed and Listing Dynamics

A homeowner who wants to sell her house traditionally works with a broker to list her house on a traditional listing platform. She then waits for an appropriate buyer and sells the house. This process can be slow: As Table 1 Panel (b) shows, the average time between a listing and a successful sale is 91 days. We begin by comparing this traditional process to a transaction intermediated through an iBuyer.

III.A.1. Selling to iBuyers avoids listing the house

We begin by examining the extensive margin of listing: To what extent do homeowners avoid the slow listing process by selling to an iBuyerWith the merged transaction-listing dataset, we estimate whether houses sold to or by iBuyers were more likely or less likely to be listed on MLS prior to purchase:

$$Listing_{izt} = \beta Buyer\_is\_iBuyer_{izt} + H_i'\mathbf{B} + \mu_{zt} + \epsilon_{izt} \quad (1)$$

An observation is at the deeds-records transaction level where *i* indexes a house in zipcode *z* at quarter *t*, and each deeds records transaction record may or may not have an associated listing in MLS. $Buyer\_is\_iBuyer_{izt}$ is an indictor for whether an iBuyer buys the house, and is a zero-one indicator for whether there is an MLS listing on the same property with a sale date within one week of the sale date in Corelogic. This indicator captures whether the sale is listed on MLS prior to the transaction. H_i is a vector of house controls, such as price, age, lot size, air conditioning type, garage type, heating type, location influence, and build quality. μ_{zt} is the zip times quarter (interacted) fixed effect. In other words, the variation we present is not simply variation across zip codes, but reflects house characteristics within a given zip code and a quarter.

The results in Table 2 Panel A show that selling to iBuyers provides substantially faster sale with a certain outcome. Sellers are roughly 27 percentage points (pp) less likely to list a property, if they sell it to iBuyer. If they list the house instead of selling directly to iBuyer, the time to sale can be substantial. When selling through a listing, slightly less than 50% of listings result in a sale within three months, and only 70% result in a sale within one year (Figure 4 Panel (b)). Even conditional on a sale, these findings imply that selling directly to an iBuyer may allow the seller to speed up the time of sale by on average 91 days.

III.A.2. iBuyers Sell Houses using Listings, set higher list prices which they adjust more often

iBuyers Sell Houses using Listings: We first show that iBuyers utilize the traditional listing process. We then show that they list houses at higher prices than homeowners, but lower prices

more aggressively if houses are not sold. We first estimate the probability that iBuyers sell houses using the listing process:

$$Listing_{izt} = \beta Seller\_is\_iBuyer_{izt} + H_i'\mathbf{B} + \mu_{zt} + \epsilon_{izt}$$

An observation is at the deeds-records transaction level where *i* indexes a house in zipcode *z* at quarter *t*, and each deeds records transaction record may or may not have an associated listing in MLS. $Seller\_is\_iBuyer_{izt}$ is an indicator that captures whether the sale is listed on MLS prior to the transaction. As before H_i is a vector of hedonics and. μ_{zt} is the zip times quarter fixed effect. iBuyers are roughly 12% *more* likely to go through the traditional listing process than other sellers (Table 2 Panel A). iBuyers do not appear to a differentiated technology for selling houses: they rely on the standard listing process instead of targeting buyers through a special website, for example.

iBuyers Set Higher List Prices and take longer to sell the house: We next examine differences in listing behavior between iBuyers and other sellers. Central in the listing decision is the tradeoff between the aggressiveness of the listing price and the speed of the sale. Levitt and Syverson (2008) shows that brokers list their own houses at higher prices than those of their clients.[6] We first examine whether iBuyers also follow the higher listing price strategy, and then examine whether they engage in any other strategies when selling houses. We compare iBuyers to two types of sellers: typical homeowners, who form the base category in the subsequent analysis, and *Flippers*, whom we define, as before, as absentee owners who re-list the house within one year of purchase. Flippers are a useful comparison group because they share some similarities with iBuyers: They are absentee owners, who likely purchase houses as an investment.[7] With this in mind, we estimate whether iBuyers listing prices differ from other listers using the following specification:

$$\log(List\ price_{izt}) = \beta Lister_{izt} + H_i'\mathbf{B} + \mu_{zt} + \epsilon_{izt} \quad (2)$$

Here *i* indexes a house in zipcode *z* at quarter *t*. The dependent variable is a listing price, not a sale price. $Lister_{izt}$ is an indicator for whether the lister is an ordinary homeowner, an iBuyer, or a Flipper. As before, we control for house characteristics in H_i and quarter x zip fixed effects μ_{zt}. We compare listing prices set by iBuyers to those of observably similar homes listed within a given zip code at the same point in time by individuals and home flippers. iBuyer listing prices are 2.1%

[6] Guren (2018) further shows that sellers do not set a unilaterally high or low list price because they face a concave demand curve: that increasing the price of an above-average-priced house rapidly reduces its sale probability, but cutting the price of a below-average-priced house only slightly improves its sale probability

[7] iBuyer's business model is not simply that of a large-scale house flipper. iBuyers are roughly 5% less likely than ordinary sellers to mention renovations, while Flippers are 15% more likely to do so. Thus, while Flippers appear to add value by renovating, iBuyers do not.

higher than ordinary sellers' listings on comparable properties (Table 2 Panel B). Flippers also list more aggressively than ordinary buyers, with a markup of 2.0%. In other words, both flippers and iBuyers appear to follow the strategy of listing at a high price.

The classic tradeoff of a higher listing price is that it results in a higher transaction price, but at a reduced probability of transaction and longer time on the market. On average, iBuyer houses spend approximately 6 days longer on the market (Table 1, Panel B). One issue that clouds simple comparison are "failed" sales (Figure 4, Panel b). Once iBuyers list the house, they are very likely to sell it. This is not surprising, since they are an intermediary who does not utilize the house. Homeowners, on the other hand, sometimes choose to pull the house from the market, and decide not to sell it at all. In fact, homeowners are 18pp more likely to have a "failed" sale once the house is listed (Table 2, Panel B). When examining *houses that eventually sold,* iBuyers' houses take longer to sell relative to other homeowners. Within that sample, iBuyers sell their homes more slowly especially in early months: they are 10% less likely to sell a home within 3 months than other sellers.

The second issue that arises is censoring. Since iBuyers are relatively new players in the market, perhaps their time on the market is a result of sales that did not yet have time to close. To properly account for "failed" sales and censoring we estimate a Cox proportional hazard model on sales propensity:

$$\lambda(t|X_i) = \lambda_0(t)\exp\left(\beta Buyer_i + H_i'\mathrm{B}\right) \tag{3}$$

Here, *i* refers to the individual listing spell, and *t* is the time between the initial listing and the sale. The dependent variable is the days to sale (which may be censored if a listing is withdrawn or does not lead to a sale). $Buyer_i$ is an indicator for whether the seller is an ordinary seller (the base category), a Flipper, or an iBuyer. H_i' is the vector of house hedonics (for example, square footage).

The hazard rate of an iBuyer sale from a listing is greater when we consider the fact that iBuyer sales are more likely to succeed. The magnitudes suggest that before conditioning on whether a sale occurs, iBuyer sales occur at roughly a 16% greater rate than non-iBuyer sales. In contrast, Column (9) shows that conditional on the sale occurring, the iBuyer hazard rate is significantly lower: conditional on a sale occurring, iBuyer sales occur at roughly a 14% lower rate. Columns (8) and (10) initial listing price to control for listing aggressiveness, and the addition of these controls do not meaningfully alter the results.

III.B. Returns for Liquidity Provision

As the previous section shows, iBuyers purchase houses quickly from homeowners and resell them. This section shows that in exchange for this service, iBuyers earn a positive gross return (spread) on average. This section decomposes the sources of return and shows that the return comes primarily through buying at a discount and selling at a premium, rather than through market appreciation.

III.B.1. iBuyer Earn a positive Bid / Ask Spread that is not Market Timing

In this section, we document that iBuyers earn a positive spread on their housing transactions even accounting for overall price changes in the market. The spread is one way to assess how much market participants seem to be willing to pay for the liquidity provision in the real estate market. Because of different holding periods of iBuyers and homeowners, we annualize the spreads, and define the annualized gross return (spread) on a given transaction,[8] as

$$Gross\ Return_{izt t'}^{Ann} = \left(\frac{Price_{izt'}}{Price_{izt}}\right)^{\left[\frac{1}{t'-t}\right]} - 1 \tag{4}$$

The subscript *i* denotes a house, *z* the zipcode of the house, and *t* the time of the purchase, and t' the time of the sale. iBuyers earn an annualized spread of 17.78% relative to homeowners' spread of 9.28% (Table 3). While iBuyer spreads are positive on average and exhibit significantly less volatility, they are also negative a significant fraction of the time, suggesting that iBuyers are sometimes willing to sell houses for a loss, even if they hold them for a short time (Figure 5).

To confirm that these differences are not driven by differences in market conditions or in the types of houses that iBuyers purchase, we regress annualized gross realized return on house hedonics and zip-quarter fixed effects at the transaction level:

$$Gross\ Return_{izt t'}^{Ann} = H_i'\mathrm{B} + \mu_{zt} + \epsilon_{izt t'} \tag{5}$$

$Gross\ Return_{izt t'}^{Ann}$ is the gross return of property *i* in zip *z* between its purchase time *t* and its sale time *t'*. All controls on the right-hand side are as of time *t*, the purchase date. H_i is a vector of house hedonics, and μ_{zt} is a vector of zip-quarter-of-purchase fixed effects. The regression therefore compares realized returns for purchases by iBuyers and non-iBuyers of similar houses as of the same date. Even controlling for differences in house types and local market conditions,

[8] The gross return does not capture other fees that iBuyers charge consumers as well as other operating costs including labor costs, financing costs, housing renovation costs, and property taxes.

iBuyers' annualized gross return is roughly 6.6% percentage points higher than those of typical individuals (Table 3). We separate the gross return into a component that is attributable purely to overall house price appreciation and the bid/ask spread. iBuyers earned roughly 1.5pp from overall market movements relative to the average household. The vast majority of iBuyers returns, 5pp, on the other hand comes from the bid / ask spread even accounting for their successful market timing (Table 4).[9]

III.B.2. Decomposition of Bid / Ask Spreads: iBuyer Purchase Discounts and Sale Premia

iBuyers can earn the bid ask spread by either purchasing houses cheaper, or selling them for more than an average household. We investigate the purchase discount and the selling premium using the following hedonic specification:

$$\begin{aligned} \log(Sale\ Price_{izt}) &= \beta(Buyer\ is\ iBuyer_{izt}) + H_i'\mathbf{B} + \mu_{zt} + \epsilon_{izt} \\ \log(Sale\ Price_{izt}) &= \beta(Seller\ is\ iBuyer_{izt}) + H_i'\mathbf{B} + \mu_{zt} + \epsilon_{izt} \end{aligned} \tag{10}$$

An observation is a house transaction, where i indexes a house in zipcode z at quarter t. $Sale\ Price_{itz}$ is the transaction price. $Buyer\ is\ iBuyer_{izt}$ and $Seller\ is\ iBuyer_{izt}$ are zero-one indicators for whether the buyer or seller is an iBuyer, respectively. H_i is a vector of house characteristics, and μ_{zt} is the zip code times quarter (interacted) fixed effect. Our specification therefore compares the price of properties acquired by iBuyers compared to observationally similar properties that transacted within a given zip code at the same point in time acquired by other market participants.

iBuyers earn both a purchase discount as well as a selling premium. The purchase discount represents most of the spread. Over the 2013-2018 sample, iBuyers' purchase prices were roughly 3.6% lower than other purchasers in the market (Table 4 Panel A). Hence, selling to iBuyers instead of another purchaser is costly: for the average house, the discount amounts to $9,000. Our results are consistent with the notion that liquidity that sellers receive from iBuyers is valuable since sellers are willing to pay for it. iBuyers' purchased homes at a discount but sold them at prices that were roughly 1.6% higher than other sellers. For the average iBuyer house with a price of approximately $250,000, this premium amounts to approximately $4,000. This result is consistent with the fact that iBuyers list houses at a higher price, resulting in higher transaction prices but a longer time to sale. While smaller than the purchase discount, the selling premium

[9] Because they hold multiple properties, iBuyers are also substantially more diversified than homeowners, and earn substantially higher risk adjusted returns. The mean annualized gross return on iBuyers portfolio is 24% with a standard deviation of 8.67% and is 11% with a standard deviation of 15.66% for homeowners

significantly contributes to the iBuyer average realized gross returns that we established earlier to be around 4.9% per transaction.

A natural alternative explanation for iBuyer's purchase discount is that they simply purchase houses that have worse characteristics on unobservable dimensions. For example, the house could be poorly kept up, have low curb appeal, or noisy neighbors. The positive sale premium suggests this is not the case. If a house has bad hedonics when purchased, then it would have similar hedonics when sold. This is especially true given the short time between iBuyer buy and sale, and the listing results in footnote that iBuyers are unlikely to renovate homes. iBuyer price discount do not seem to reflect different fees charged by iBuyers. For example, according to Opendoor, the largest iBuyer in our sample, the company charged the home sellers an average service fee of about 7.5% of home value per transaction, which is substantially larger than the typical real estate agent fees of 6% or less.[10]

These results taken together are consistent with the idea that iBuyers, in exchange for providing liquidity to homeowners looking for an instant sale, purchase the house at a significant discount. They go on to sell the house at a small premium by listing it at a higher price. The gross spread they earn is compensation for their liquidity provision.

III.C iBuyers Intermediate in Easy-to-price and Liquid Homes

We have documented that iBuyers act as liquidity providers, buying earning a bid ask spread and carrying properties on their inventories for a relatively short period of time. The large purchase discount indicates that sellers are willing to pay a significant amount for liquidity provision in this market. A high liquidity demand suggests a natural role for dealer intermediation. Yet, until the entry of iBuyers, such transactions were rare. To better understand why liquidity provision in real estate markets is difficult, we examine which market segments iBuyers chose to enter. We use the characteristics of these segments as an indicator that in these markets, intermediation is easiest. We focus on the role of information and underlying market liquidity. These two forces play a central role in intermediation. Moreover, iBuyers tout their algorithmic pricing as an advantage. This Several iBuyers are offshoots of firms, which specialize in collecting house price data, as well as pricing houses. Appendix A.1 shows screenshots from Opendoor's website."

III.C.1. iBuyers' use algorithmic pricing

Intermediaries who purchase and sell assets must be concerned about adverse selection. This is especially the case in the housing market, where houses have diverse characteristics, some of

[10] See https://www.opendoor.com/w/pricing (accessed on January, 2020).

which are difficult to measure, and homeowners have an advantage in the knowledge of these characteristics. We examine how iBuyers price houses that they purchase, and find that a simple algorithm of hedonics, which account for local market conditions can explain a large part of their pricing strategy. We estimate the following hedonic regression:

$$\log(Sale\ Price_{izt}) = H_i'\mathbf{B} + \mu_k + \epsilon_{izt} \tag{11}$$

An observation is a house transaction, where i indexes a house in zipcode z at quarter t. $Sale\ Price_{izt}$ is the transaction price. H_i is the vector of house characteristics. We include a sequence of fixed effects, μ_k with k indexing the level of saturation. We include zip, quarter, zip and quarter, and zip times quarter fixed effects.

Across specifications, observable characteristics explain a substantially higher share of variation in iBuyer transaction prices – both for purchases and sales – relative to transaction prices of other market participants (Table 4). House characteristics explain roughly 50% of the variation in price for transactions involving iBuyers relative to 40% for transactions by other participants. With zip times quarter fixed effects capturing local time-varying market trends, observable characteristics explain over 80% of the variation in prices for iBuyer transactions, versus 68% of the variation in transactions by other market participants. These results suggest that a simple algorithm of hedonics, which account for local market conditions can explain a large part of iBuyers' pricing strategy. Moreover, non-iBuyer real estate buyers use other inputs to determine prices that does not seem to be captured in the iBuyer algorithm. Such information can either arise from other participants using difficult to encode information that is available—soft information—or information acquired through a thorough and lengthy inspection, which iBuyers do not conduct because they offer a speedy closure. If iBuyers' pricing is not contingent on this additional information, which is likely also known to property sellers they may be vulnerable to adverse selection.

The reliance on easy to assemble hard information also differentiates iBuyers from several other fintech participants. For example, fintech lenders in the mortgage market use standard hard information than other market participants, when pricing mortgages (Buchak et al, 2018).

III.C.2. iBuyers' intermediate in easy-to-value and liquid segments of the market

As we discuss in Section II.C, iBuyers entry is very selective, both geographically, and, conditional on geography, which types of homes they purchase. In this section, we use iBuyers' revealed preferences to study why intermediation in real estate markets is so limited. As we suggest above, because iBuyers rely on algorithmic pricing, they are potentially vulnerable to adverse selection. If that is the case, they should intermediate in segments, in which their informational disadvantage

is smallest: those in which their algorithmic pricing of hard hedonic information works well. The underlying liquidity of the house should also play a first order role in both the demand and supply of liquidity. More illiquid houses have higher demand for liquidity provision. On the other hand, it may be also more difficult to the supply liquidity for these houses. We look at iBuyer entry to better understand which force dominates.

We first estimate which houses can be easily priced with an algorithm using of hard information, i.e. which types of houses have a small pricing error when priced with hedolics. To ensure that iBuyers' pricing decisions do not mechanically affect our classification, we estimate a hedonic pricing model using 2008-2012 data, which precedes iBuyers. We then study which markets iBuyer chose to enter from 2013 through 2018. To estimate which houses are easy to price, we estimate a model of the form of Equation (11), which regresses log sale prices on house hedonics at the level of house *i*, in zip code *z*, at quarter *t*, on the training sample defined above. Houses with higher standardized (squared) residuals $e_{izt}^2 = \frac{\hat{\epsilon}_{it}^2}{\sigma_{\epsilon^2}}$ from this specification are those that are not priced well by hedonics.[11] We then predict which house characteristics make them difficult to price:

$$e_{izt}^2 = H_i'\mathbf{\Delta} + \zeta_{\mathrm{izt}} \tag{13}$$

As before, H_i' is a vector of house hedonics. A positive coefficient on a particular house characteristic, for example, means that on average, houses with that characteristic will have greater pricing errors when priced with a hedonic model. The results are intuitive: newer houses, larger houses, and multistory houses command higher prices (Table 5). Houses with large property sizes, on the other hand, are more difficult to price with a simple hedonic model.[12] Intuitively differences in land characteristics can results in different opportunities to develop the land. The larger the land, the larger the valuation differences.

We follow a similar process when trying to estimate how the ex-ante liquidity of a market relates to iBuyer entry. We use the listing data to estimate a hedonic model to predict whether a given listing sells within 90 days form the listing date, our measure of a liquid market from 2008-2012:

$$Sells\ Within\ 90\ Days_{izt} = H_i'\mathbf{B} + \epsilon_{izt} \tag{14}$$

[11] The normalization allows us to assess changes residuals in units of standard deviations of prediction pricing errors.
[12] As a robustness check, we estimate the same analysis for a larger training sample spanning 2006-2012. The remarkable stability in the coefficient estimates provide confidence that iBuyers using different sets of data would come to largely similar conclusions.

$Sells\ Within\ 90\ Days_{izt}$ is a zero-one indicator for whether house *i* in zip *z* listed at time *t* sells within ninety days of its listing. Using the estimates from these two equations, we construct for every house its predicted standardized pricing error $\hat{e}_{izt}^2$ and underlying liquidity $Sells\ Within\ 90\ Days_{izt}$.

Figure 6 graphically shows that iBuyer market shares are highest in segments of easy to price and liquid houses. We formally test whether iBuyer choice of which houses to purchase from 2013-2018 is related to the measures of house valuation accuracy and liquidity using the following specification:

$$iBuyer_{izt} = \beta \hat{e}_{izt}^2 + \gamma \widehat{Sells\ Within\ 90\ Days}_{izt} + \mu_{zt} + \epsilon_{izt} \tag{15}$$

As with the earlier specification, an observation is a house transaction, where *i* indexes a house in zipcode *z* at quarter *t*. $iBuyer_{itz}$ is a zero-one indicator for whether the buyer is an iBuyer. $\hat{e}_{izt}^2$ is the predicted pricing error normalized to the standard deviation of the pricing errors, and $\widehat{Sells\ Within\ 90\ Days}_{izt}$ is the predicted probability of a listing selling within 90 days. As before, we control for quarter x zip fixed effects μ_{zt}. This means we study which houses iBuyers choose to intermediate, conditional on having entered a geographic market.

A house with a one standard deviation greater predicted pricing error is 3.7% less likely to be purchased by an iBuyer (Table 6). This is a large effect relative to the base rate of iBuyer purchases over this sample period of roughly 0.60% and demonstrates the importance of price predictability in iBuyer participation. The predicted liquidity of a house is also strongly associated with iBuyer intermediation. An increase in the probability of selling within 90 days of 10pp corresponds with an increase in iBuyer market share of 0.2%. These results suggest that iBuyers indeed intermediate in more liquid houses and those which are easiest to price. We argue that iBuyers are reluctant to transact in houses with a high pricing error, because these types of houses expose them to adverse selection. Then, if iBuyers *do* buy such houses, they earn smaller profits.

These results suggest that two forces limit the provision of dealer intermediation in the real estate market despite its high potential benefits. To provide liquidity, intermediaries need to transact in homes quickly and are therefore subject to adverse selection. Even firms, which specialize in algorithmic pricing such as iBuyers are at an information disadvantage relative to other buyers, who can take the time to conduct a thorough investigation. Second, while low liquidity of a house increases the demand of homeowners for liquidity, it also decreases the intermediaries' ability to supply liquidity. On net, our results suggest that the second effect dominates. As we illustrate in the next section, the results is natural when applied to real estate intermediation. Illiquid homes make it relatively more efficient for the seller to live in the house during the sale process rather

than keep the house vacant exposing it to adverse maintenance shocks. Critically, liquidity provision is therefore efficient only when houses are already relatively easy to value and liquid—when additional liquidity is least valuable.

IV. Equilibrium Housing Trading Framework with iBuyers

In this section we develop an equilibrium model of house search and matching, in which we introduce an intermediary, which purchases houses from households, holds them, and resells them to other households—iBuyers. We study the equilibrium effect of the pricing technology available to the intermediary, and the associated adverse selection problem, as well as the speed at which it can close transactions, and thus provide liquidity to sellers. We calibrate the model to the data, to explore the qualitative and quantitative forces, which constrain the provision of liquidity in the market, even when demand for liquidity is high.

IV.A Model Setting

The model is in continuous time in which all agents discount the future at rate ρ. Figure 7 shows the timeline of the model within a period, tracing out the role of homeowners and iBuyers through the transaction. A homeowner is initially matched with a house from which she receives a flow benefit (consumption value less costs). With some probability, she becomes unmatched from her current house and begins the process of moving. We assume that the homeowner's balance sheet is constrained and can only own one house at a time. Therefore, to buy a new house she must sell her old house. Once she finds a new house she likes, and purchases it, has bought her new house, she again becomes a matched homeowner. The transactions among homeowners occur in standard search market, in which sellers list houses and are randomly matched with buyers.

We depart from the standard setting by introducing an intermediary which can provide liquidity in the market: iBuyers. Instead of listing houses and waiting for a buyer, sellers can sell them directly to iBuyers. iBuyers then list houses for resale, using the standard listing process. iBuyers are a balance sheet intermediary: they are not constrained to hold only one house at a time. On the other hand, iBuyers do not live in the house when trying to sell it—the house remains vacant. This means that they do not obtain utility flows from homeownership although they still must pay maintenance costs.

We endow iBuyers with three different characteristics, which can affect their ability to intermediate in this market. The first is speed; in other words, the reason why intermediaries are valuable to sellers in the first place is because they can execute the transaction without waiting for a buyer. The second is information. As we document above, even if iBuyers have great algorithmic pricing, they are still at an information disadvantage relative to homeowners, and other potential

sellers, who can take time to thoroughly screen a purchase. The third is matching technology. While we see little evidence of differences in matching homeowners with houses study the effect of these characteristics on the effectiveness of intermediation in this market, this is a popular explanation of iBuyers advantage, and we explore its consequences in the model. We study the model by varying how changes in these characteristics affect the equilibrium outcomes in the market, and the profits from intermediation, to see if intermediation is viable.

IV.A.1. Market and information structure

Homeowners: At any instant, a homeowner is one of three states, between which she transitions over time. These states are denoted $\{h, s, b\}$. h denotes a matched homeowner, who is happy with the house in which she currently resides. s denotes a selling homeowner: a homeowner who is unhappy with her current house and is in the process of selling it. b denotes a buying homeowner: a homeowner who currently does to own a house and is looking for one. The total homeowner population has an exogenous mass $M = 1$, with $\{m_h, m_s, m_b\}$ denoting respectively the endogenous mass of h-types, s-types, and b-types. Homeowners become unmatched at rate μ.

Matched h-type homeowners own a house producing flow utility $\bar{u}_i = \bar{u} + \tilde{\epsilon}_i$. $\bar{u}_i$ captures the benefits of living in the house such as housing services, proximity to work, and so on, net of holding costs. $\bar{u}$ capturing the flow component common across homeowners; $\tilde{\epsilon}_i$ allows for idiosyncratic differences in homeowners utility flows from their current property. When homeowners become unmatched, they receive utility flow $\underline{u} < \bar{u}$. This represents the idea that while they still obtain some utility benefits from occupying the house, the house is no longer ideal. For example, their place of work may have changed, increasing their current commute, or they have had children and desire to switch to a different school district. Or, if they sold their current house, they occupy a non-ideal rental residence.

Listings: Selling households can list their house and wait to be randomly matched with potential buyers. Buyers and s-type households meet at an aggregate rate $F^{hh}(m_s, m_b) = \lambda m_s m_b$. λ measures the underlying liquidity of the market. A high λ implies a market in which matching is faster, all else equal, because buyers and sellers can, for example, meet on the internet. Of course, the final matching rate $F^{hh}(m_s, m_b)$ is endogenous. The probability of a given seller matching with a buyer declines with the overall number of sellers. Let subscripts s and b, denote the rate for an individual buyer or seller to match; then $F_s^{hh}(m_s, m_b) \equiv F^{hh}(m_s, m_b)/m_s$. Given a listing price p, a matched buyer accepts the offer with endogenous probability $\pi(p)$.

iBuyers: iBuyers are intermediaries which purchase houses directly from households, hold them until sale, and sell the houses using a listing. Instead of listing their house, households can sell it to iBuyers. The iBuyer transaction closes in τ days, for example. Closing delays arise because of documentation, but also because of inspections of the property. This parameter allows us to study the importance of immediacy in the business model of an intermediary in this market and highlights the tradeoff between speed and precision of information, which we introduce below. We also allow households to differ in their preferences over transacting with iBuyers. For example, some households have to move urgently, or they are technologically savvier. We capture these preferences with an idiosyncratic utility shock, ϵ_{ib}^{i}. ϵ_{ib} is distributed type-1 extreme value distribution with scale parameter σ_{ib}.

Upon purchasing the house, iBuyers sell the house using listings. iBuyers can be more or less effective at finding buyers for their properties than households, $F^{ib}(m_s, m_b) = \lambda_{ib} F^{hh}(m_s, m_b)$. $\lambda_{ib} > 1$ would imply that iBuyers are better able to find buyers than other sellers in the market, potentially by using their own websites for listings. While the house is on the market, iBuyers pay maintenance costs m^L, which we normalize to 0. Because the house is unoccupied, iBuyers cost may increase over time, for example, if the roof leaks, and no one notices because the house is vacant. With probability η, the costs become high, m^H. The increase in costs is important to rationalize two facts that we document in the data.

House quality and information: iBuyers close housing transactions without a lengthy inspection, and instead use algorithms to set prices. We model the potential information disadvantage that Buyers face as a "repair cost" r. If the house can be sold as is, $r = 0$ with probability $1 - \phi_R$. Alternatively, with probability ϕ_R the house may need some repairs, that are not visible at first blush, then $r = R$. These repair costs are known to the seller, who has lived in the house. Moreover, these costs can be uncovered by a thorough inspection by a potential buyer. For tractability, we assume that inspections by b-types uncover this information completely, and that the repair cost is paid by the seller.[13] Because they perform more cursory inspections, iBuyers receive a noisy signal v of repair costs. They observe whether a house is "Good" and "Bad," where

$$\phi(G|r = R) \equiv \phi_{G|R} = \xi$$
$$\phi(G|r = 0) \equiv \phi_{G|\sim R} = 1 - \xi$$

A low ξ implies a better technology, which has a low probability of classifying a high repair cost house as good. In essence, one can think of iBuyer technological problem as trading off speed of closing τ with accuracy ξ. iBuyers then condition houses prices on the signal, $p_b^{ib}(v)$. Upon

[13] This assumption is without loss of generality, but allows for listed houses to be homogenous in quality, increasing model tractability

purchasing the house, iBuyers find out whether the house will actually require repairs before listing, and pay the required cost.

IV.A.2. Homeowners' problem

Homeowners choose their actions to maximize their expected utility. Let $\{v_h, v_s, v_b\}$ denote the value functions of *h*-, *s*-, and *b*- type homeowners, respectively.

Matched homeowners do not need to take any actions. At any point, their consumption flow is that of their current house $\bar{u} + \tilde{\epsilon}_i$, and the continuation value of living in the current house. The latter depends both on how likely they are to become unmatched, and, conditional on being unmatched, how likely their house is to require repairs, and their utility of selling to an iBuyer. Formally, a matched homeowner *i* has the following value function:

$$(\rho + \mu)v_h^i = \bar{u} + \tilde{\epsilon}_i + \mu \int_{r,v,\epsilon_{ib}^i} \max\{v_s - r, \delta(\tau)\left(p_{ib}^b(v) + v_b\right) + \epsilon_{ib}^i\} \, dG\left(\epsilon_{ib}^i, v, r\right)$$

The integration is over the repair cost, the probability of being unmatched (a seller), and idiosyncratic value of selling to an iBuyer. These random variables are jointly distributed as $dG\left(\epsilon_{ib}^i, v, r\right) = d\mathrm{E}^{\mathrm{ib}}\left(\epsilon_{ib}^i\right) dG(v, r)$. Given the above decision problem, $\pi_{ib}(p, r) = 1 - \mathrm{E}^{\mathrm{ib}}[v_s - r - \delta(\tau)(p + v_b)]$.

v_h^i can be expressed as a sum of a common component v_h, which is how the average homeowner values their house, and the idiosyncratic home valuation $\tilde{\epsilon}_i$, how homeowner *i* values her house relative to the average homeowner . Hence, for the remainder of the paper we focus on v_h and $\epsilon_i \sim \mathrm{E}(\epsilon_i)$, with $v_h^i = v_h + \frac{\tilde{\epsilon}_i}{\rho+\mu} \equiv v_h + \epsilon_i$. We interpret ϵ_i as the capitalized idiosyncratic flow utility from the house. ϵ_i is distributed type-1 extreme value distribution with scale parameter σ_m.

Selling homeowners have two choices. They either sell the house to the iBuyer or choose to list it. If the homeowner chooses to list the house, she pays the repair cost, if any, and becomes a seller. She sets the listing price to maximize her expected utility and faces a standard tradeoff: the higher the house price p, the higher the profit once the house is sold. On the other hand, the endogenous probability that a buyer accepting the offer $\pi(p)$ declines, and she cannot start buying a new house until she sells the old one. Formally, her value function is:

$$\rho v_s = \underline{u} + F_s^{hh}(m_s, m_b) \max_p \pi(p)(p + v_b - v_s)$$

If she accepts the iBuyer's offer, she gets the price and becomes a buyer after the closing period, τ the value of which she discounts by $\delta(\tau)$.

Buyers: b-type homeowners have sold their houses and are trying to purchase a new house. They decide whether to purchase a house for the lister price. Upon encountering the seller and seeing the house, the buyer's idiosyncratic valuation, ϵ_i, realizes, she pays a viewing cost κ, and she chooses whether to accept or to continue looking. Her value function is given by:

$$\rho v_b = \underline{u} + \sum_j F_b^j(m_j, m_b) E\left[\max\{v_h + \epsilon_i - p_j, v_b\} - v_b - \kappa\right]$$

Where j indexes the seller type. The buyer accepts the offer if $v_h + \epsilon_i - p_j > v_b$. Thus, $\pi(p) = 1 - E[v_b + p - v_h]$.

IV.A.3. iBuyers' problem

When a homeowner becomes unmatched, iBuyers inspect the house for repair costs. iBuyers then offer a price $p_{ib}^b(v)$ that depends on their quality signal. A homeowner, who knows the acutal repair cost r accepts the price with probability $\pi_{ib}(p, r)$. Because repair costs are asymmetric information, this is the source of adverse selection in the model. Upon acceptance, the iBuyer pays the repair cost, if any, and takes possession of the home. Let v_{ib}^L and v_{ib}^H denote the value functions for a low- and high-maintenance iBuyer, respectively. The iBuyer's expected profit when making an offer to an unmatched homeowner given signal v is as follows:

$$v_{ib}^{offer}(v) = \max_p \int_r \pi_{ib}(p, r)(v_{ib}^L - p - r) dF(r|v)$$

After buying the house, the iBuyer's decision is like a listing seller. In the beginning, her maintenance costs are low, m^L. When setting the price, she trades off the probability of a sale, and the profits she realizes conditional on a sale. Moreover, she sets the price accounting for the fact that her costs might increase in the future:

$$\rho v_{ib}^L = m^L + \eta(v_{ib}^H - v_{ib}^L) + F_s^{ib}(m_s, m_b) \max_p \pi(p)(p - v_{ib}^L)$$

If iBuyer's maintenance costs increase to m^H, she sets the price to maximize expected profits given her match rate:

$$\rho v_{ib}^H = m^H + F_s^{ib}(m_s, m_b) \max_p \pi(p)(p - v_{ib}^H)$$

IV.A.4. Population dynamics

We now present the population dynamics. There are five prices in the market at any point in time: the price at which households list their houses, p_{hh}, the prices at which iBuyers buy high and low quality houses, $p_{ib}^{b}(v)$, and the listing price that iBuyers start with when their maintenance costs are low, p_{ib}^{L}, and the listing price of iBuyers once the maintenance costs increase p_{ib}^{H}. For a vector of prices $\boldsymbol{P} \equiv \{p_{ib}^{b}(v), p_{hh}, p_{ib}^{L}, p_{ib}^{H}\}$, we define $\pi_{ib}(\boldsymbol{P})$ as the unconditional probability that a seller sells to an iBuyer, given by integrating jointly over repair costs and signals:

$$\pi_{ib}(\boldsymbol{P}) \equiv \int_{v,r} \pi_{ib}\left(p_{ib}^{b}(v), r\right) dG(v,r)$$

Then, first three equations describe how the population of different types of households change over time

$$\frac{dm_h}{dt} = -\mu m_h + \sum_{j \in \{hh, ib^L, ib^H\}} F^j\left(m_j, m_b\right)\pi\left(p_j\right)$$
$$\frac{dm_s}{dt} = \mu m_h\left(1 - \pi_{ib}(\boldsymbol{P})\right) - F^{hh}(m_s, m_b)\pi(p_{hh})$$
$$\frac{dm_b}{dt} = \mu m_h \pi_{ib}(\boldsymbol{P}) + F^{hh}(m_s, m_b)\pi(p_{hh}) - \sum_{j \in \{hh, ib^L, ib^H\}} F^j\left(m_j, m_b\right)\pi\left(p_j\right)$$

The next two equations describe the changes in the stock of iBuyer houses depending on their maintenance costs:

$$\frac{dm_{ib}^{L}}{dt} = \mu m_h \pi_{ib}(\boldsymbol{P}) - \eta m_{ib}^{L} - F^{ib}(m_{ib}^{L}, m_b)\pi(p_{ib}^{L})$$
$$\frac{dm_{ib}^{H}}{dt} = \eta m_{ib}^{L} - F^{ib}(m_{ib}^{H}, m_b)\pi(p_{ib}^{H})$$

Additionally, the total population is fixed, and that the total number of houses is fixed which implies:

$$m_h + m_s + m_b = M$$
$$m_s + m_{ib}^{L} + m_{ib}^{H} = m_b$$

The first equation captures the fact that the total number of matched, buying, and selling houses is equal to the total population, M. The second equation captures the fact that the housing stock is fixed, which means that for every house on the market, there is exactly one potential buyer.

IV.A.5. Equilibrium

We look for a stationary equilibrium. The equilibrium is a set of prices such that the first-order conditions determining prices are satisfied, the value function equations are satisfied and iBuyers earn positive profits in expectation, and the state variables $\{m_h, m_s, m_b, m_{ib}^L, m_{ib}^H\}$ are constant as determined by their laws of motions and iBuyers beliefs about housing repairs are consistent as described above. There is a first-order condition for each of the three listing sellers---households, low-cost iBuyers, and high-cost iBuyers, as well as a first-order condition for each value of the iBuyer's signal when buying.

IV.B Model Calibration

We first calibrate the model. We calibrate several parameters externally in relation to existing literature, parameters that map directly to observable quantities. We then use the method of moments to calibrate the remaining parameters by searching for parameters that map model-implied moments to moments we observe in the data. Table 7 Panel A describes these moments and our model's fit. Panel B details the externally calibrated parameters, and Panel C details the parameters calibrated through the method of moments. We go into further detail in subsequent sections.

V.B.2 Externally calibrated parameters

We follow Anenberg and Bayer (2020) and set the discount rate ρ to 0.05. The census[14] estimates that individuals move roughly 9.1 times after they turn 18, or at a rate of roughly 0.152 (9.1 / 60 years) per year. Agents in our model move after becoming unmatched, and consequently this number corresponds to the unmatching rate μ in our model. Additionally, we assume that houses needing serious repairs mention renovations in their listings. Thus we set the probability that a house needs serious repairs, p_f, to the fraction of listings mentioning renovation, which is 0.109 in the MLS data. Finally, we assume for the baseline analysis that iBuyers' time to close, τ, is one day—we explore how changing this parameter affects iBuyers' ability to intermediate in Section VI. Finally, we set the meeting cost κ to 0.577. For reasons discussed below, this is the expected value a potential owner receives when viewing a house with zero net utility plus a Type-1 EV utility shock, accounting for the potential buyer's option to refuse the offer. Table 7 Panel B summarizes these choices.

[14] https://www.census.gov/topics/population/migration/guidance/calculating-migration-expectancy.html

V.B.3 Parameters calibrated to the data: Identification

We calibrate the remaining 10 parameters by matching moments in the equilibrium model with the empirical target in 2018, the most recent year in our data. We summarize the parameters and the moments in Table 7. As the estimates highlight, the model can match the data quite well even quantitatively. Discussing the identification of the remaining parameters also presents an opportunity to provide exposition on the economics underlying the model.

Match utilities: The gap between mean utilities of housing between matched and unmatched households, $\bar{u}$ and $\underline{u}$, are reflected in listing prices and iBuyer discounts. The main benefit of buying a house for an unmatched buyer is to receive exactly this increase in utility. A buyer is willing to pay more for a house when this gap is larger, and if buyers are willing to pay more, then the seller will list at a higher price. By selling directly to an iBuyer, can become rematched sooner and obtain this utility increase faster. Thus, when the gap is larger, the iBuyer purchase discount grows.

We can use iBuyers selling behavior to compare their cost of holding the house to the cost of an unmatched household. If living in an unmatched house has lower utility for a homeowner than the maintenance costs faced by iBuyers, then they would face more urgency in selling the house. They would list it at a lower price than iBuyer. This is consistent the iBuyers selling at a premium, which we document in the data.

Match rates: The match intensity, λ, is related to how long a house stays on the market. When λ is high, matches occur more frequently. Likewise, iBuyer's relative match advantage / disadvantage, λ_{ib}, maps to time on market for iBuyer listings. A higher λ does not mechanically map into shorter duration on the market. With a higher λ buyers can reject sellers and see more offers per unit time, which reduces seller markups. This force tends to reduce time on market. On the other hand, sellers see more buyers per unit time, and therefore may want to raise prices, which, other things equal, tends to increase time on market. The model can replicate the average 90-day time on the market. iBuyer homes spend approximately 6 days longer on the market but do so at higher list prices. Our calibration suggests that iBuyers do not possess a better matching technology than other sellers, with a $\lambda_{ib} = 0.97$.

iBuyer cost increases: Because iBuyers houses are vacant, they face the possibility that maintenance costs increase if issues are undetected for a while. The cost increase is given by the parameter c, and the Poisson arrival rate of the cost increase is given by the parameter η. η most directly impacts the probability that iBuyers decrease prices before selling. We observe this moment in the data, and from the model we can calculate the probability that an iBuyer reduces prices before successfully finding a seller. The cost increase c to a first order governs the size of

the price reduction: if possessing the house suddenly becomes more expensive, iBuyers will want to reduce prices to sell the house more quickly. Additionally, these parameters broadly impact iBuyer market share, through the fact that a higher c or η increase expected iBuyer costs and therefore market share.

Most interestingly, these parameters influence the relationship between iBuyer market share and market liquidity. In particular, for a fixed c and η, the increased costs are less likely to realize in more liquid markets because the probability that iBuyers will sell the house before the cost shock arrives is greater. Thus, the model generates a negative relationship between iBuyer market share and market liquidity, as measured in the data by the probability that homes sell within 90 days. Recall that our reduced form results also found a negative relationship. Our model generates a predicted derivative, d(iBuyer share)/d(P(sells in 90 days), which is another moment we match in the data to discipline the calibration.

Repair cost and iBuyer assessment noise: Houses may need repairs costing R, and iBuyers receive signals regarding whether they need repairs with false positive and true negative probabilities ξ. In our model, repair costs R are the source of unexplained house price variation. A larger R relative to the price generates greater unexplained variation. We take this to the data by mapping it to the residual mean square error of the regression of house prices on house hedonics, with the interpretation that unexplained variation captures unobserved repair costs (or the costs of improving the house quality to what is required to sell it).

Next, because in our model, homeowners can observe whether the house needs repairs but iBuyers cannot, there is adverse selection, with homeowners being adversely selected against on average. This occurs because iBuyers offer a single pooling price conditional on a noisy signal. Homeowners with houses needing repairs attempt to pool with homeowners with houses not needing repairs. Thus, homeowners not needing repairs will therefore list themselves, while homeowners needing repairs will attempt to sell through iBuyers, in a classic lemons problem.

The noisier the iBuyer's signal is, the more severe the adverse selection becomes, and fewer and fewer homeowners will benefit from selling through iBuyers. Thus, more noise, ξ, leads to a lower iBuyer market share in the model. We thus calculate the market share derivative with respect to pricing error, d(iBuyer share)/d(Mean pricing error). We map this moment to its empirical analog, which we obtain from regressing whether an iBuyer is involved in a transaction with the hedonic pricing error on the house, which in the data, provides a negative coefficient.

Shock variances: Finally, we need to identify the variances on the taste shocks for buyers upon seeing a listing, σ_m, and the taste shocks for sellers when deciding whether to sell to iBuyers, σ_i. A low variance means that an agent's action is more likely to align with what the non-random

components in the model would predict. For instance, if a buyer's deterministic utility from purchasing a house is -1, and her deterministic utility from the outside option is 0, a low variance implies that she will choose the deterministically optimal action---in this case, the outside option---more frequently. In economic terms, high variances correspond to more market power for the agent not receiving the shock, e.g., the lister.

In the context of our model, because iBuyers offer to purchase houses at discounts, a higher σ_i makes potential sellers more likely to sell to iBuyers, and conversely, a lower σ_i makes potential sellers less likely to sell to iBuyers. Thus, iBuyer market share, conditional on the offer price, is informative about σ_i.

σ_m impacts sellers' market power. Other things equal, greater variance in σ_m will lead sellers to increase markups. Importantly, however, this logic applies asymmetrically between ordinary sellers and iBuyers. Ordinary sellers are impatient to sell in part because they want to move from a bad match to a good match. iBuyers, on the other hand, are less impatient, and thus more able to take advantage of their market power to wait for a high quality match. Thus, σ_m impacts both house prices overall, but in particular iBuyer markups.

V.C Model Fit and Validation

We now discuss how the model matches the targeted moments and present evidence of the model's fit of non-targeted moments.

V.C.2. External validation

We perform two external validation exercises with the model. The first examines the equilibrium impact of iBuyer entry, which we tie to reduced form evidence on the impact of iBuyer entry on house prices and liquidity, which we detail here. The second examines variation in which consumers in the model utilize iBuyer and tie these predictions to reduced form evidence on post-iBuyer household mobility.

To test our model's predictions on prices and liquidity, we simulate the calibrated model with and without iBuyers. In each case, the model produces equilibrium house prices and times on market. As shown in Figure 8 Panels (a), our model predicts that without iBuyers, median house prices (in the calibrated market) is roughly $275,000, and roughly $262,000 thousand with iBuyers. This represents a reduction of roughly 4.7%. Recall that in our calibration, iBuyer market share is 4.9%, giving an elasticity between prices and iBuyer market share of close to -1. Panel (b) shows that the average time on market decreases from 115 days to 90 days once iBuyers enter. Assuming an exponential distribution, this corresponds to an increase of roughly 11.5% of listings selling within

two weeks to roughly 14.5% of listings selling within two weeks,[15] or an elasticity between share of houses sold within two weeks and iBuyer market share of roughly 0.65.

We next look for reduced form validation of these quantities. To do so we use regional variation and exploit Redfin data at the zip code level. The data reports listing prices, and the fraction of listed homes in a zip code that sell within two weeks of listing. We investigate how these variables change with iBuyer market hare using the following specification:

$$\Delta Liquidity_z = \beta iBuyerShare_z + X_z'\mathbf{\Gamma} + \epsilon_z \tag{34}$$

$$\Delta Price_z = \beta iBuyerShare_z + X_z'\mathbf{\Gamma} + \epsilon_z \tag{35}$$

$\Delta Liqidity_z$ and $\Delta Price_z$ are the 2013-2018 changes in the percentage of houses selling within two weeks of listing and percent change in house prices. X_z' is the vector of zip-level controls, and $iBuyerShare_z$ is the growth in iBuyer Market share from entry in 2015 to 2018.[16] The coefficients on iBuyer market share map directly to the elasticities reported above. Because iBuyer entry may be endogenously correlated with liquidity (indeed, our model suggests that it is), we instrument for iBuyer market share using the physical characteristics of the housing stock transacting before iBuyer entry. Specifically, we use the following to predict which homes iBuyers purchase in Phoenix in 2018:

$$iBuyer_{izt} = H_i'\mathbf{B} + \epsilon_{izt} \tag{3}$$

Then, at the zip code level, we calculate the predicted iBuyer market share among houses that transacted between 2011 and 2014 in the zip code, defining:

$$\%\,\widehat{iBuyer}_z = \frac{1}{N_z} \sum_{i \in z; t \in (2011-2014)} \widehat{\imath Buyer}_{\imath zt} \tag{36}$$

As before, $\widehat{\imath Buyer}_{\imath zt}$ is the house-level prediction for whether iBuyer would buy the house, *i* indexes over all houses in zip code *z*, and time *t* spans 2011 to 2014. We use this measure to instrument for $iBuyer\ Share_z$. Thus, our instrument for iBuyer market share in 2018 (equivalently, iBuyer market share growth from 2015 to 2018), is the predicted iBuyer share based on the physical homes transacted between 2011 and 2014.

[15] This is calculated as $1 - \exp\left(-\frac{1}{115} \times 14\right)$ and $1 - \exp\left(-\frac{1}{90} \times 14\right)$.

[16] Because iBuyer market share began at 0%, this measures both changes and levels.

Table 8 Column (1) shows that the first stage effect is very strong: A 1% increase in predicted share based on the physical characteristics of houses transacting in 2011-2014 is associated with a 0.65% increase in actual iBuyer market share in 2018. Columns (2) through (7) show the OLS, reduced form, and IV results for the liquidity and price outcomes. Consistent with the model's predictions, we find elasticity of price to iBuyer market share roughly around -1, depending on the specification, and between 0.15 and 2 for fraction of houses selling within two weeks of listing, again, depending on the specification. Broadly, our model makes predictions that are qualitatively and quantitatively consistent with this reduced form exercise.

To summarize, our model predicts that iBuyers specifically target those individuals with an idiosyncratically high valuation of transacting early. To take this prediction to the data, we show that individuals living in iBuyer-type houses are significantly more likely to leave the market (e.g., move to another city) once iBuyers enter. We interpret these households as those placing the greatest value on moving early, consistent with our model's predictions.

VI. The Economics, Technology, and Limits of iBuyer Liquidity Provision

In this section we use the calibrated model to probe the underlying economics of residential real estate intermediation. We first discuss and quantify the key economic forces that make intermediation here valuable but difficult. We next use the model to examine the role of iBuyer technology and ask specifically whether a balance sheet alone would be sufficient for intermediation in this market. Finally, we examine the limits of iBuyer liquidity provision, particularly as it relates to the ex-ante liquidity of a market.

VI.A. The Economics of Intermediation

As discussed earlier, dealer-intermediated residential real estate transactions alter three key economic forces as compared to the traditional broker-intermediated transaction. The first force is speed: A listing homeowner has to go through the lengthy listing process, whereas selling to the dealer-iBuyer takes place almost instantaneously. Our model reflects this through the fact that iBuyers buy directly, but that closing potentially takes some time, τ. The second force is occupancy: A listing homeowner (typically) remains in her house during the listing process, allowing her to derive utility flows and perform routine maintenance. In contrast, a dealer-iBuyer leaves the house unoccupied, foregoing the utility flows and potentially failing to perform routine maintenance. Our model reflects this through differences in flow utilities, $\underline{u}$, and the chance that an iBuyer's house might become more expensive to maintain, η. The third force is information: Homeowners are better equipped to observe hard-to-measure differences in value than dealer-

iBuyers, who value the home remotely. Our model reflects this through the precision of the signal, specialized in our calibration as ξ. In this section, we quantify the importance of each of these channels by altering these parameters away from their calibrated values and assessing the impacts on predicted iBuyer market share.

VI.A.1. Speed

We investigate the quantitative importance of iBuyers' transaction speeds by varying the parameter governing time to close, τ. The baseline value is $\tau = 1$ days, and we calculate counterfactual iBuyer market shares by setting $\tau = 90$ days, roughly the time an average homeowner takes to sell her house through a listing. We leave all other parameters equal to their calibrated values, and compare the baseline iBuyer market share to this counterfactual "slow" iBuyer's market share.

Figure 9 Panel (a) shows the results. The first column shows the market share for the baseline case, roughly 5%. The second column shows the market share for the slow iBuyer case, roughly 0%. This result shows the perhaps unsurprising importance of speed in the viability of an iBuyer transaction. Without their speed advantage, iBuyers offer essentially nothing of benefit to households that they could not effect more better themselves.

VI.A.2. Occupancy

We next investigate the quantitative importance of the fact that selling homeowners occupy their houses, while iBuyers do not. This fact brings two potential disadvantages to iBuyers. First, while homeowners are not situated in their ideal home, they still live in a house and derive *some* consumption benefit from it. Second, the fact that they are in the home means that they engage in routine maintenance, which protects from potential adverse events like a roof caving in. iBuyers potentially face both of these drawbacks when intermediating. To examine these impacts, we consider a counterfactual where iBuyers rent the house to a homeowner (e.g., the original selling homeowner) in order to produce the utility flow of an unmatched homeowner, $\underline{u}$, and set the rate of maintenance cost increases η to zero.

Figure 9 Panel (a) shows the results. The third column shows the market share for the iBuyer offering second-best occupancy. We find essentially no difference in iBuyer market share, suggesting that this margin is not quantitatively important. The intuition can be seen from our calibrated parameters: The unmatched flow utility is actually *negative*, suggesting that it is better

to leave the house unoccupied than to have the unmatched homeowner living in it. The interpretation is that the unmatched homeowners utilizing iBuyer services are better served by living somewhere else than remaining in their old house. This is broadly consistent with our earlier reduced from evidence showing that owners of iBuyer houses tend to leave the market entirely when selling.

VI.A.3. Information

Finally, we investigate the role of information asymmetry in limiting residential real estate intermediation. The intuition is that ordinary homeowners, both buyers and sellers, are better able to detect minor issues with the house than iBuyers running a remote valuation. We call these issues "repair costs" but the interpretation can be broadened to include hard-to-detect items like loud neighbors, good midday light, and pleasant ambiance. In our model, the parameter ξ controls the noise in the iBuyer's signal. To investigate the importance of adverse selection, we set $\xi = 0$ and recompute the equilibrium iBuyer market share.

Figure 9 Panel (a) again shows the results. The fourth column shows that with a perfect valuation technology, iBuyer market share jumps from 5% (Column (1)) to 20% (Column (4)). Intuitively, given repair costs, iBuyers are forced to offer lower prices. These lower prices are unappealing to homeowners whose houses are in good condition and do not require repairs. As homeowners with high quality houses leave the market, the discount that iBuyers have to offer increases further, driving both high- and low-quality home homeowners out of the market. Giving iBuyers a perfect valuation technology eliminates this death spiral. The large quantitative impact here suggests that this is an important economic force in residential real estate intermediation.

VI.B. iBuyer Technology and Balance Sheets

We now use the model to attempt to understand the source of iBuyers' advantage. We note that iBuyers differ from homeowners, because they derive no intrinsic value when holding on to the house. iBuyers therefore face a disadvantage relative to homeowners when it comes to holding real estate. They therefore need some other advantage to be willing to engage in intermediation. In particular, are iBuyers merely an intermediary with a balance sheet, or does their technology provide an important distinct benefits? We examine three aspects of iBuyer potential technological advantage over a potential intermediary with a balance sheet capacity: their valuation speed, their valuation accuracy, and their matching technology when selling.

To shed light on it, Panel (b) of Figure 9 shows the comparison of market shares of intermediaries having the baseline iBuyer technology, a fast and inaccurate technology, and a slow and accurate

technology. In particular, for speed, we compare the iBuyer baseline closing speed of 1 day to a more typical 25 day closing time. For accuracy of valuation technology, we compare a valuation technology with iBuyers' precision to an uninformative valuation technology (uninformative signal regarding the true value of the house).

Intuitively, our counterfactual considers the case of an intermediary with a balance sheet but unsophisticated valuation technology. This intermediary can either undertake the accurate valuation at the expense of slow purchase speed, or they can purchase quickly at the cost of performing the valuation inaccurately. iBuyers, in contrast, can do these tasks both quickly and relatively accurately. We examine the counterfactual market shares of these intermediaries. The results highlight the fact that merely having balance sheet capacity with no technological advantage is not sufficient. The low-tech valuation intermediary would counterfactually achieve between 1 and 2% market share, in contrast to the iBuyer, which has two to five times that market share. These findings indicate that transaction speed and valuation accuracy are the key elements of technological advantage of iBuyers over other intermediaries. They also help explain why intermediation in real estate market has been fairly rare prior to the iBuyer entry.

This analysis helps us understand why iBuyers are predominately active among homes that can be fairly accurately priced using (algorithmic) automatic valuation models as we established in Section III.F. Among these homes their technological advantage is the highest, which is a key factor behind their significant market share. Among homes that are difficult to price our model predicts a low market penetration, which is exactly what we find in the data.

VI.C. The Limits of iBuyer Liquidity Provision

Finally, we examine the extent to which, and in what contexts, iBuyers are able to provide liquidity. Intuitively, a naïve prediction might be that iBuyers provide the greatest liquidity benefit in markets that are illiquid before iBuyers enter. However, the data show that iBuyers focus on the most liquid markets. To understand the economic mechanism behind this finding we consider iBuyer market shares and percent changes in transaction speeds as a function of pre-iBuyer transactions speeds as captured by the parameter governing the matching rate, λ. Pre-iBuyer transaction speeds serve as a measure of ex-ante liquidity.

Figure 10, Panel (a) show that the association between pre-iBuyer entry market liquidity and the iBuyer market share implied by our model. The results show that greater ex-ante liquidity is associated with greater iBuyer market share. The model provides two mechanisms leading to this result. First, iBuyers maintain empty houses, and when maintaining an empty house, there is the possibility that an unobserved negative shock (e.g., a collapsed or leaking roof) occurs which significantly increases maintenance costs. Thus, the longer that an iBuyer has to hold the house, the more likely it is that this shock occurs. In a thinner markets where sales are slow, iBuyers must

either face this risk, or alternatively reduce their sale prices to sell the house faster. In both cases, low liquidity markets put iBuyers at a disadvantage.

As iBuyers face more adverse holding costs and lower profits, they eventually have to compensate by offering lower prices to potential sellers. This, due to the adverse selection that iBuyers face, has the effect of collapsing the market for good houses that do not need repairs: Given these low prices, good types would rather list houses themselves, and thus only bad types sell their houses to iBuyers. As this part of the market disappears, iBuyers' market share collapses dramatically. Similarly, for liquidity provision, while iBuyer entry reduces the time it takes to become rematched regardless of ex-ante liquidity, the effect is stronger when markets are more liquid. This is because in the more liquid markets, more iBuyers enter, and thus have both a greater direct and indirect effect, though competitive pressures, on time to rematch.

To summarize, iBuyer benefits are largest in the markets that you might think need them the least. This is illustrated in Panel (b) of Figure 10 showing that iBuyers have much more pronounced relative effect on the housing turnover (time-to-rematch) of already liquid homes.

Taken together our findings highlight why liquidity provision in real estate markets has been limited, despite the high potential benefits to market participants. Difficult to price and illiquid homes make it relatively more efficient for the seller to live in the house during the sale process rather than keep the house vacant. Liquidity providers, such as iBuyers, are subject to adverse selection, and have to keep houses vacant while holding them for resale exposing themselves to potential maintenance shocks. Liquidity provision is therefore efficient only when houses are already easy to price and relatively liquid. This ensures that such homes can be acquired and resold quickly while limiting the scope for adverse selection and adverse maintenance shocks. This argument suggests that iBuyers, with their current use of technology are not likely to impact a large part of the market – i.e., they will make already liquid and easy to price houses more liquid, rather than unlocking the sale of illiquid and harder to price homes.

VI. Discussion and Conclusion

In this paper, we studied the growth of "iBuyers," online real estate companies that buy and sell residential real estate, which have gained significant market share since 2015, to provide novel evidence on the effects and challenges of making housing markets more liquid. We show that these firms act as liquidity providers, buying low and selling high, and carrying properties on their inventories for only a short period of time. Like in the case of online fintech lenders in mortgage origination, consumers appear to greatly value the convenience that iBuyers offer, and sell their properties to them at a considerable discount.

We also document considerable limitations to the liquidity provision by iBuyers. Their pricing relies more on hard information and hence they do not enter market segments with difficult-to-

value homes to also limit the scope of adverse selection. Moreover, since homes are empty during the intermediation phase and subject to adverse upkeep shocks, iBuyers tend to focus on properties that can be resold relatively quickly.

We rationalize these findings within a search-based housing trading model with iBuyers. Our model highlights why liquidity provision in real estate markets has been limited, despite its high potential benefits. While the intermediary keeps the house vacant while listing, forgoing these consumption benefits are not quantitatively important. Rather, intermediaries are subject to significant adverse selection which significantly limits their expansion in harder-to-value homes. However, despite the difficulties in fast, remote valuation, we show that the transaction speed and valuation accuracy that iBuyers possess are nevertheless an important innovation over other potential dealer intermediaries. In contrast to iBuyers, the low-tech valuation intermediary with just balance-sheet capacity would counterfactually achieve a negligible market share, which explains why prior to the iBuyer entry intermediation in the housing market has been rare.

We conclude by making a few observations regarding our findings. First, it is possible that development of better pricing algorithms and collection of new data could considerably expand the range of properties that iBuyer could accurately price in the future. Our analysis suggests, however, that for this to meaningfully affect their market penetration and hence the liquidity provision, the iBuyers would need to be able to resell such properties relatively quickly. In other words, a substantial increase in the market penetration of iBuyers -- and hence in the liquidity provision in the market -- will require not only technological improvements in algorithmic pricing but also in the ability to match homes quickly with subsequent buyers. Our analysis suggests that iBuyers at present do not possess such technological advantages when selling their inventory.

Second, the growth of iBuyer market share we focus on occurred during relatively good times in the housing market (2013-2018), when on average most of the properties were holding or appreciating in value. It is unclear how viable the iBuyer business model would be during an economic downturn accompanied by a decline in house prices. On the one hand, an increase in expected time to resell the property and challenges in accurately pricing homes during rapidly changing economic environment may considerably limit if not shut-down altogether the liquidity provision by iBuyers.[17] This indeed occurred during the early stages of COVID-19 pandemic when most of iBuyers temporarily suspended their operations.[18] On the other hand, the economic downturn could increase the share of homeowners that value convenience of a quick sale, making liquidity provision by iBuyers more valuable. We leave the analysis of viability of iBuyer business model across various economic environments for future research.

[17] In addition, iBuyers could face considerable financial stress due to their need to finance a large inventory of homes that may be declining in value.

[18] https://finance.yahoo.com/news/zillow-redfin-realogy-opendoor-suspend-045712124.html

Finally, our findings have broader implications for balance sheet intermediation of consumption goods. Assets such as homes that are relatively illiquid, harder to price, and have high utilization value have seen little intermediation in the past. Only recent technological advances in valuation accuracy and speed of transacting facilitated by on-line acquisition platforms have allowed some inroads into provision of intermediation services in such markets. On the other hand consumption goods such as cars are relatively more liquid, easier to price, and have relatively lower carry cost, which explains why intermediation in such markets have been historically at much higher levels.

References

Anenberg, E., and P. Bayer, 2020, Endogenous Sources of Volatility in Housing Markets: The Joint Buyer-Seller Problem, forthcoming in the *International Economic Review*.

Buchak, G., G. Matvos. T. Piskorski, and A. Seru, 2018, Fintech, Regulatory Arbitrage and the Rise of Shadow Banks, *Journal of Financial Economics* 130, 453-483.

Buchak, G., G. Matvos. T. Piskorski, and A. Seru, 2018, Beyond the Balance Sheet Model of Banking: Implications for Bank Regulation and Monetary Policy, Working Paper.

Calder-Wang, S., 2019, The Distributional Impact of the Sharing Economy on the Housing Market, Working Paper.

Duffie, D., N. Garleanu, and L. H. Pedersen, 2005, Over-the-Counter Markets, *Econometrica* 73, 1815–1847.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery, 2019, The Role of Technology in Mortgage Lending, *Review of Financial Studies* 32, 1854–1899.

Gavazza, 2016, An Empirical Equilibrium Model of a Decentralized Asset Market, *Econometrica* 85, 1755-1798.

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

Grossman, S. J., and M. H. Miller, 1988, Liquidity and Market Structure, *Journal of Finance* 43, 617–633.

Guren, A., 2018, House Price Momentum and Strategic Complementarity, *Journal of Political Economy* 126, 1172-1218.

Jiang, E., G. Matvos, T. Piskorski, A. Seru, 2020, Banking without Deposits: Evidence from Shadow Bank Call Reports, Working Paper.

Levitt, S.D., and C. Syverson, 2008, Market Distortions when Agents are Better Informed: The Value of Information in Real Estate Transactions, *Review of Economics and Statistics* 80, 599-611.

Ngai, L. R., and S. Tenreyro, 2014, Hot and Cold Seasons in the Housing Market, *American Economic Review* 104, 3991-4026.

Piazzesi, M., and M. Schneider 2009, Momentum Traders in the Housing Market: Survey Evidence and a Search Model, *American Economic Review* 99, 406–411.

Piazzesi, M, M. Schneider, and J. Stroebel, 2020, Segmented Housing Search, *American Economic Review* 110*,* 720-59.

Wheaton, W. C., 1990, Vacancy, Search, and Prices in a Housing Market Matching Model, *Journal of Political Economy* 98, 1270–1292.

Table 1: Summary Statistics

This table shows summary statistics for the main datasets used in the paper: the Corelogic transaction deeds records between 2013 and 2018 (Panel A), and the MLS listings data between 2013 and 2018 from ATTOM Data (Panel B). Data are from Phoenix, Orlando, Dallas, Gwinnet County, and Las Vegas. For a given variable we show number of observations, mean, standard deviation (S.D.), and 5 through 95 percentiles of the data. For the Corelogic data (Panel A), *Sale price* is the sale price, *Land sq ft* is the assessed land square footage, *House age* is the age of the house at the time of transaction. *iBuyer is buyer* indicates when an iBuyer is buying the property. *iBuyer is seller* indicates when an iBuyer is selling the property. *All other* is all transactions not involving an iBuyer as either buyer or seller. For the MLS data (Panel B), *first list price* is the first listed price of the property on MLS. *Mentions renovation* is an indicator for whether the listing mentions "renovation," "refurbish," or "remodel." *Total listings* is the number of listings in the listing spell. *Has sold* is an indicator for whether the property ultimately sells. *Days on market* is the number of days between initial listing and sale (only among sold listings), and *Discount to list* is the ratio of the sale price to the initial listing price minus one (only among sold listings). *iBuyer* is when an iBuyer is listing the property. *Flipper* is when an absentee owner who has owned the house for less than one year before listing is listing the property. *Other* is when a non-iBuyer, non-lister seller is listing the property.

Panel A: Transaction Data (Corelogic)

Variable	N	Mean	S.D.	5%	25%	50%	75%	95%
Sale price								
iBuyer is buyer	5,887	251,982	194,389	146,653	191,000	230,400	281,350	390,000
iBuyer is seller	7,384	269,795	206,276	164,664	208,000	245,000	295,000	398,000
All others	885,451	280,251	372,086	82,500	156,000	218,175	305,000	582,500
Land sq ft								
iBuyer is buyer	6,003	7,094	3,880	2,800	5,227	6,580	8,073	12,324
iBuyer is seller	7,460	7,208	3,900	2,614	5,227	6,664	8,273	12,946
All others	966,261	9,074	6,948	2,614	5,720	7,405	9,798	21,622
House age								
iBuyer is buyer	5,978	20	12	4	12	17	28	45
iBuyer is seller	7,431	21	12	5	12	17	29	46
All others	954,313	27	19	6	13	22	40	63

Table 1: Summary Statistics, Continued

Panel B: Listing Data (ATTOM Data)

Variable	N	Mean	S.D.	5%	25%	50%	75%	95%
First list price								
Other	924,490	325,627	213,028	124,500	193,000	265,000	380,000	750,000
iBuyer	1,353	241,471	75,131	160,000	193,000	226,000	269,900	383,000
Flipper	63,328	319,038	209,176	119,900	185,900	260,000	379,000	740,000
Mentions renovations								
Other	924,490	0.097	0.296	0	0	0	0	1
iBuyer	1,353	0.033	0.179	0	0	0	0	0
Flipper	63,328	0.282	0.450	0	0	0	1	1
Total listings								
Other	924,490	3.686	2.335	1	2	3	5	8
iBuyer	1,353	6.997	4.073	3	4	6	9	15
Flipper	63,328	3.837	2.835	1	2	3	5	9
Has sale								
Other	924,490	0.705	0.456	0	0	1	1	1
iBuyer	1,353	0.960	0.196	1	1	1	1	1
Flipper	63,328	0.668	0.471	0	0	1	1	1
Days on market								
Other	652,039	91.124	85.670	26	42	63	106	249
iBuyer	1,299	96.810	70.013	29	45	74	127	242
Flipper	42,332	89.835	79.029	26	44	66	106	229
Discount to list								
Other	651,153	-0.033	0.061	-0.133	-0.057	-0.024	0	0.032
iBuyer	1,299	-0.028	0.043	-0.101	-0.052	-0.021	0	0.026
Flipper	42,270	-0.036	0.059	-0.132	-0.060	-0.027	0	0.023

Table 2: iBuyer Transaction Behavior

This table examines iBuyer transaction behavior on the intensive and extensive margin of listing. Panel A shows the propensity to list a property on multiple listing service (MLS) using merged transaction-listing data with sale dates between 2013 and 2017. The table presents the OLS estimates from a regression of the dummy variable that takes value of one if the seller lists the property on MLS and is zero otherwise on the dummy variable taking the value of one if the property buyer was an iBuyer and is zero otherwise (Columns 1-3), and the dummy variable taking the value of one if the property seller was an iBuyer and zero otherwise (Columns 4-6). Columns (1) and (4) have no controls or fixed effects. Columns (2) and (5) have zip-quarter fixed effects. Columns (3) and (6) additionally include square footage, house age, and whether the house is multistory. Panel B compares the pricing dynamics of listings where an iBuyer is the seller to listings of home flippers and regular homeowners (excluded category). *Log first price* is the log of the first listing price. *Mentions renovations* is an indicator for whether the listing description describes the house as being renovated, i.e., includes "renovation," "refurbish," or "remodel." *Total listings* is the number of price adjustments in a given listing spell. *Leads to sale* is an indicator for whether the listing leads to a sale. *Days on market* is the number of days between the first listing and the sale. *Sale-to-list* is the sale price divided by the initial listing price. Columns (1)-(6) and (11) are linear specifications; (7—10) are Cox Proportional Hazard Rate models. Columns (1)-(4) and (7—8) consider all listings. Columns (5), (6), (10), and (11) consider only listings leading to sales. Columns (6), (8), and (10) include the log of the initial listing price. A *Flipper* is an absentee owner who lists the house within one year of purchasing it. The *iBuyer* and *Flipper* categories are measured relative to the base category of other lister. All columns include house controls including square footage, whether the house is multistory, and house age. The linear models include zip times quarter fixed effects. Data are from Corelogic and MLS provided by ATTOM Data between 2013 and 2018 at the combined listing level. Standard errors are in parentheses.

Panel A: Propensity to List on MLS

	Dependent variable: Lists on MLS					
	(1)	(2)	(3)	(4)	(5)	(6)
Buyer is iBuyer	-0.314	-0.272	-0.275	-	-	-
	(0.006)	(0.006)	(0.006)	-	-	-
Seller is iBuyer	-	-	-	0.121	0.135	0.128
	-	-	-	(0.006)	(0.005)	(0.005)
Hedonics	N	N	Y	N	N	Y
Zip-Quarter FE	N	Y	Y	N	Y	Y
Observations	822,081	822,081	807,102	822,081	822,081	807,102
R^2	0.003	0.134	0.171	0.001	0.133	0.169

Panel B: iBuyer Listing Behavior

Specification	Linear	Linear	Linear	Linear	Linear	Linear	Hazard	Hazard	Hazard	Hazard	Linear
Dependent variable	Log list price	Mentions renovations	Total listings	Leads to sale	Days on market	Days on market	Days on market	Days on market	Days on market	Days on market	Sale-to-list price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
iBuyer	0.023	-0.057	1.338	0.136	27.115	26.191	0.309	0.309	-0.019	-0.020	-0.005
	(0.007)	(0.007)	(0.037)	(0.009)	(1.700)	(1.639)	(0.024)	(0.024)	(0.024)	(0.024)	(0.001)
Flipper	0.008	0.142	0.172	-0.012	2.180	1.394	-0.080	-0.079	-0.069	-0.070	-0.002
	(0.001)	(0.001)	(0.006)	(0.001)	(0.317)	(0.307)	(0.005)	(0.005)	(0.005)	(0.005)	(0.0002)
Hedonics	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Initial list price	N	N	N	N	N	Y	N	Y	N	Y	Y
Zip-Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sample	All				Sales only		All		Sales only		
Observations	1,348,518	1,348,518	1,348,518	1,348,518	800,182	789,168	1,348,518	1,348,518	800,182	800,182	789,168
R^2	0.748	0.176	0.489	0.357	0.392	0.413	0.042	0.043	0.049	0.050	0.176

Table 3: Gross Returns: iBuyers versus Individuals

Panel A shows holding periods and the realized gross housing investment return for iBuyers and non-corporate individuals. Observations are all purchases in Corelogic data between 2013 and 2018 where the buyer sells the house during this sample period. Column (1) shows the average holding period in years, defined as the number of years between purchase and sale. Column (2) shows *Gross Return*, calculated as the percentage change in house price from purchase to sale. Column (3) shows annualized gross returns calculated by annualizing gross returns by the holding period. Column (4) shows the number of houses a purchaser purchases in a given quarter conditional on purchase. Column (5) shows the annualized portfolio returns, calculated by averaging the annualized returns of all houses purchased by a single buyer in a single quarter. The top number in each row is the mean; the bottom number in parentheses is the standard deviation. Panel B shows the regression of holding period returns on house controls and zip-quarter fixed effects and the iBuyer dummy taking the value of one if the property is purchased by an iBuyer and zero otherwise. Columns (1) and (2) show annualized gross return in decimals. Columns (3) and (4) show the *Index Return*, defined as the percentage change in median house prices in the three-digit zip code from the quarter of purchase to the quarter of sale, in decimals. *Non Index Return* is the residual: *Gross Return* minus *Index Return*. Columns (1), (3), and (5) include no controls or fixed effects. Columns (2), (4), and (6) include house hedonic controls including square footage, house age, and whether the house is multistory (and excluding price). Standard errors are in parentheses. The table excludes extreme observations where the total or annualized return is greater than 50% in absolute value.

Panel A: Raw returns

	Holding period given sale (years)	Gross Return (Raw %)	Gross Return (Ann %)	Quarterly portfolio size	Portfolio return (Ann %)
	(1)	(2)	(3)	(4)	(5)
Individuals	2.65	24.66	9.28	1.02	11.21
	(1.28)	(19.83)	(9.09)	(0.69)	(15.66)
iBuyer	0.39	4.91	17.78	113	24.15
	(0.38)	(6.39)	(19.72)	(117)	(8.67)

Panel B: Spread regressions

	Dependent variable:					
	Gross Return (ann)		IndexSpread (ann)		NonIndexSpread (ann)	
	(1)	(2)	(3)	(4)	(5)	(6)
iBuyer	0.074	0.066	0.004	0.015	0.068	0.050
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
House Controls	N	Y	N	Y	N	Y
Zip x Quarter FE	N	Y	N	Y	N	Y
Observations	102,140	94,499	102,140	94,499	102,140	94,499
R^2	0.014	0.225	0.0004	0.581	0.011	0.231

Table 4: iBuyer Discounts, Premia, and Pricing Strategies

Panel A examines the extent to which iBuyers buy and sell properties at a premium or discount relative to similar properties that did not involve iBuyers, i.e., individual or corporate owners. The table presents the OLS estimates of the logarithm of the acquisition price on the dummy variable taking the value of one if the property buyer was an iBuyer and is zero otherwise (Column 1 and 2), and the dummy variable taking the value of one if the property seller was an iBuyer and is zero otherwise (Column 3 and 4). The house controls are those as in previous tables, except for price: *House age* is the difference between the transaction date and the year of construction. *Land square footage* is the tax-assessed property square footage. *Multistory* is an indicator for whether the house has greater than 1 stories (including partly-multilevel houses that have "1.5" stories.) Other house characteristics are air conditioning type, garage type, heating type, location influence, and build quality. Columns (1) and (3) user zip times quarter fixed effects. Columns (2) and (4) include zip-year fixed effects, to explore the effect that seasonality has on iBuyer pricing. Data are from Corelogic between 2013 and 2018. Standard errors are shown in parentheses. Panel B examines the extent to which the physical house characteristics and local economic conditions can explain the variation in pricing of properties that iBuyers intermediate in. The table shows the R^2 from regressions of log house price on house characteristics and fixed effects for transactions where iBuyers are buyers, where iBuyers are sellers, and other transactions that do not involve iBuyer, using Corelogic transaction data between 2013 and 2018. Each row differs in the fixed effects it includes. In the first, there are no fixed effects. In the second row, there are zip fixed effects. In the third row, there are quarter fixed effects. In the fourth row, there are zip and quarter fixed effects. In the fifth row, there are zip times quarter fixed effects. A *High iBuyer market* is a zip code in above the 75th percentile for iBuyer market share over the sample period, 2013-2018. (1) represents how much hedonics (and fixed effects) explain price variation when iBuyers purchase. (2) measures this for when iBuyers sell. (3) measures this for transactions in which no iBuyer is involved. (4) and (5) split the no-iBuyer transactions into those taking place in markets where iBuyers are common (4), and markets where iBuyers are uncommon (5).

Panel A: iBuyer Purchase Discount and Sale Premium

	Dependent variable: Log(sale amount)			
	(1)	(2)	(3)	(4)
Buyer_is_iBuyer	-0.036	-0.031	-	-
	(0.005)	(0.005)	-	-
Seller_is_iBuyer	-	-	0.016	0.019
	-	-	(0.004)	(0.004)
House Controls	Y	Y	Y	Y
Zip-Quarter FE	Y	N	Y	N
Zip-year FE	N	Y	N	Y
Observations	822,166	822,166	822,166	822,166
R^2	0.705	0.693	0.705	0.693

Panel B: Determinants of iBuyer Transaction Prices

			All markets		*High iBuyer Market*	*Other Markets*
Hedonic controls	Fixed effects	iBuyer buyer	iBuyer seller	No iBuyer involved	No iBuyer involved	No iBuyer involved
		(1)	(2)	(3)	(4)	(5)
Y	None	0.483	0.471	0.401	0.520	0.417
Y	Zip	0.675	0.671	0.625	0.593	0.637
Y	Qtr	0.552	0.508	0.443	0.592	0.449
Y	Zip + Qtr	0.740	0.712	0.676	0.669	0.673
Y	Zip x Qtr	0.833	0.803	0.684	0.674	0.683

Table 5: House Characteristics and Algorithmic Pricing Errors

This table shows the estimated relationship between property characteristics, house prices, and pricing errors. We use Corelogic transaction data between 2008 and ends at the end of 2012., which corresponds to the data that an entering iBuyer would have available when making a pricing decision once they enter the market. Column (1) shows the regression of log of house price on house characteristics and Column (2) shows the regression of squared pricing errors (normalized by mean price) on house characteristics. This residual is obtained directly from Column (1), squared, and divided by the standard deviation of the residuals. Omitted house characteristics include garage type, heating type, air conditioning type, and house quality. Columns (3) and (4) show a robustness checking using data between 2006 and 2012, with Column (3) corresponding to the pricing model and (4) corresponding to the errors model. Standard errors are shown in parentheses.

	2008-2012 (Main specification)		2006-2012 (Robustness)	
	Dependent variable:			
	Log(house price)	Squared deviation from predicted price	Log(house price)	Squared deviation from predicted price
	(1)	(2)	(3)	(4)
House age (omitted category: > 50 years)				
Age < 5 years	0.364***	-0.060***	0.363***	-0.065***
	(0.003)	(0.002)	(0.003)	(0.002)
Age 5-15 years	0.320***	-0.060***	0.313***	-0.062***
	(0.003)	(0.002)	(0.002)	(0.001)
Age 15-50 years	0.111***	-0.048***	0.111***	-0.050***
	(0.003)	(0.001)	(0.002)	(0.001)
Land square footage (omitted category: > 25k)				
Square footage < 5k	-0.808***	-0.016***	-0.796***	-0.019***
	(0.004)	(0.002)	(0.003)	(0.002)
Square footage 5-10k	-0.532***	-0.023***	-0.531***	-0.027***
	(0.003)	(0.002)	(0.003)	(0.002)
Square footage 10-25k	-0.271***	-0.014***	-0.279***	-0.016***
	(0.003)	(0.002)	-0.796***	-0.019***
Multistory	0.172***	-0.002**	0.176	0.0002
	(0.001)	(0.001)	(0.001)	(0.001)
Other house characteristics	Y	Y	Y	Y
Observations	680,640	680,640	889,661	889,661
R^2	0.665	0.029	0.696	0.024

Table 6: Limits to iBuyer Technology: Easy-to-Price and Liquid Homes

This table shows the regression of whether an iBuyer purchases the house on predicted pricing errors, $\hat{e}^2_{izt}$ based on house hedonics and the predicted probability that a house sells within 90 days of listing. $\widehat{SellsWithin90Days}_{izt}$ based on house hedonics and MLS data. All columns include zip times quarter fixed effects. Standard errors, clustered at the property level, are shown in parentheses. $\widehat{SellsWithin90Days}_{izt}$ relies on knowing the last sale price, which is not available for all properties, which explains the drop in observations from Column (1) to (2) and (3). Standard errors, clustered at the property level, are shown in parentheses.

	Dependent variable:		
	iBuyer buyer (%)		
	(1)	(2)	(3)
$\hat{e}^2$	-3.686	-	-3.796
	(0.272)	-	(0.469)
$\widehat{SellsWithin90Days}$	-	1.919	1.716
	-	(0.192)	(0.197)
Zip x Quarter FE	Y	Y	Y
Observations	557,164	259,771	259,771
R^2	0.026	0.030	0.031

Table 7: Equilibrium Housing Trading Model with iBuyers: Calibration and Fit

This table provides details of the model calibration. Panel A shows targeted moments in the data and calibrated model. Panel B shows parameters calibrated externally or as normalizations, together with their values and sources. Panel C shows parameters calibrated through the method of moments, where parameters are chosen to match the model-predicted moments to the empirical moments in the data as shown in Panel A.

Panel A: Moments Targeted in Calibration and Fit

Moment	Data (2018)	Model
Median house price ($k)	262.000	262.424
iBuyer market share (%)	4.881	4.918
iBuyer purchase discount (%)	2.673	2.862
iBuyer selling premium (%)	0.050	0.009
Time on market (days)	91.000	90.895
ΔiBuyer Time on market (days)	6.000	6.205
P(iBuyer reduces listing price)	0.562	0.576
iBuyer price reduction (%)	3.266	1.545
Mean pricing error	0.143	0.148
d(iBuyer share)/d(P(sells in 90 days)	0.017	0.011
d(iBuyer share)/d(Mean pricing error)	-0.017	-0.016

Panel B: Parameters Calibrated Externally / Normalizations

Parameter	Description	Value	Source
ρ	Discount rate	0.050	Anenberg and Bayer (2020)
μ	Unmatching rate	0.152	Census
p_f	Probability needs renovation	0.109	Fraction of listings mentioning renovation
τ	iBuyer closing time (days)	1.000	Industry reports

Panel C: Parameters Calibrated by Method of Moments

Parameter	Description	Value
$\bar{u}$	Matched flow utility	15.1
$\underline{u}$	Unmatched flow utility	-3.1
λ	Matching rate	148
λ_{ib}	Relative iBuyer matching rate	0.97
c	Cost increase from iBuyer possession	2.61
η	Likelihood of additional cost from iBuyer possession	5.10
ξ	Noise in iBuyer signal	0.18
R	Repair cost	5.24
σ_m	Variance on T1EV selling shock	0.21
σ_i	Variance on T1EV iBuyer transaction shock	0.56

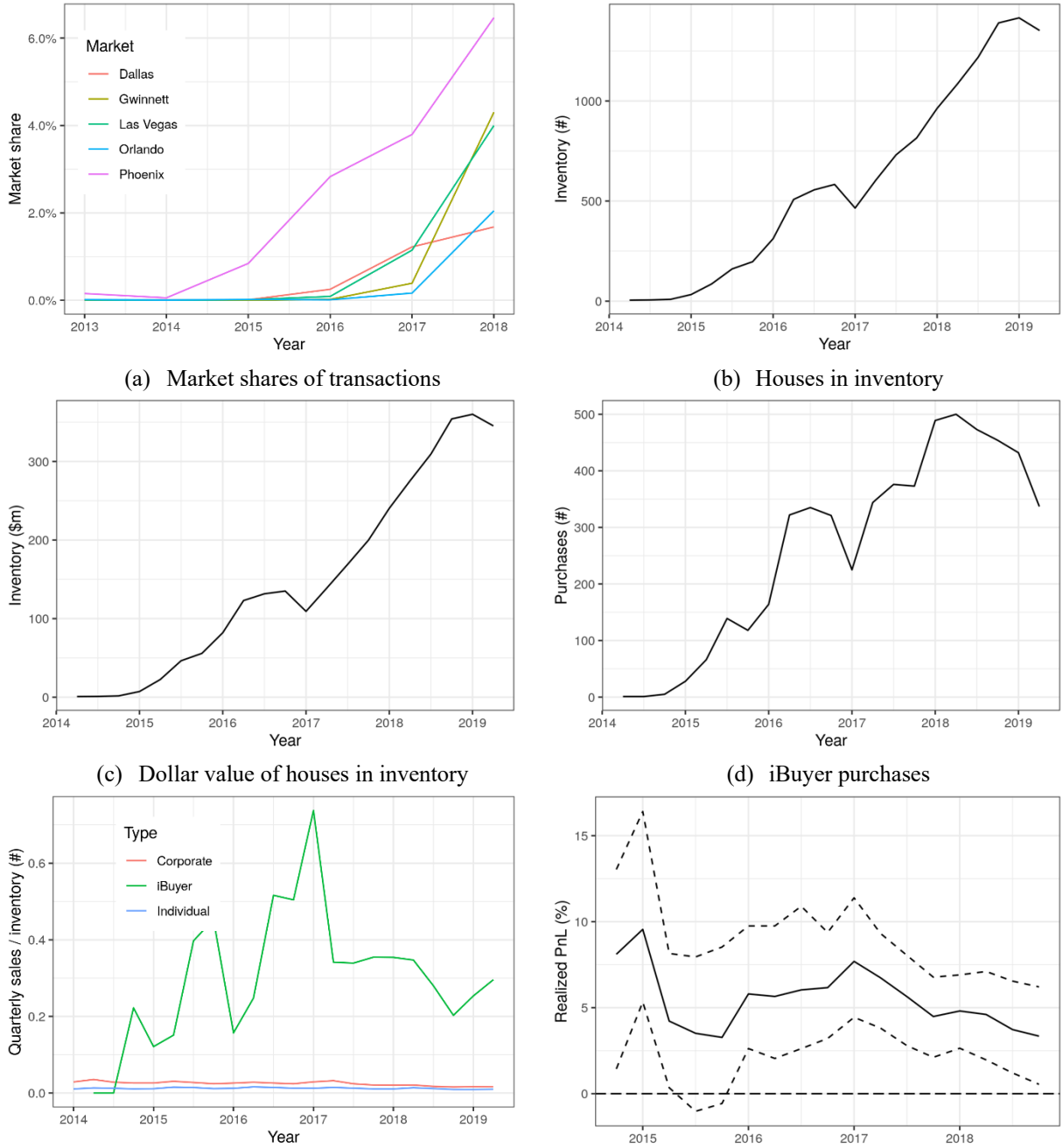
Table 8: Model Validation: iBuyer Entry and Regional Outcome Variables

This table shows the regional effects of iBuyer entry using the zip-code level data from Redfin. iBuyer Share is the change in iBuyer share between pre (2013) and post (2018), omitting Phoenix, which is used to fit the hedonic model. *% Sold in 2 weeks* is the change in houses sold within two weeks of listing. *Sale Price* is the percentage change in sale prices. % $\widehat{\imath Buyer}$ is the predicted market share of iBuyer based on the physical characteristics of houses sold in the zip code from 2011 to 2014 using a model based on 2018 Phoenix transactions. iBuyer share is the actual iBuyer share. $\widehat{\imath Buyer\ Share}$ is the instrumented iBuyer share, using the % $\widehat{\imath Buyer_z}$ based on physical characteristics as an instrument. All columns include the zip level regional controls utilized earlier. 1st Stage is the first stage regression of iBuyer share on iBuyer propensity based on house characteristics. OLS is the regression of outcomes on actual iBuyer share. RF is the reduced form regression of outcomes on iBuyer propensity. IV is the instrumental variables regression using propensity to instrument for iBuyer share. Standard errors are shown in parentheses.

	iBuyer Share	ΔLiquidity: % Sold in 2 weeks			Sale Price		
	(1st Stage)	(OLS)	(RF)	(IV)	(OLS)	(RF)	(IV)
% $\widehat{\imath Buyer_z}$	0.653	-	1.424	-	-	-1.491	-
	(0.079)	-	(0.243)	-	-	(0.312)	-
$iBuyer$	-	0.140	-	-	-0.887	-	-
$share$	-	(0.163)	-	-	(0.201)	-	-
$\widehat{\imath Buyer}$	-	-	-	2.180	-	-	-2.282
$share$	-	-	-	(0.498)	-	-	(0.515)
Zip controls	Y	Y	Y	Y	Y	Y	Y
Observations	330	330	330	330	330	330	330
R^2	0.455	0.182	0.259	0.330	0.657	0.661	0.606

Figure 1: iBuyer Market Shares, Inventory, and Realized Gross Return (Spread)

Panel (a) shows iBuyer market share in buying or selling transactions across five large markets: Dallas, Texas, Gwinnett County, Georgia, Las Vegas, Nevada, Orlando, Florida, and Phoenix, Arizona using Corelogic data. Panel (b) shows the stock of houses owned by iBuyers at the end of every quarter. Panel (c) shows the dollar value of this inventory (based on purchase price, in millions of dollars). Panel (d) shows the number of purchases by quarter. Panel (e) shows the inventory turnover, defined as sales per inventory, for iBuyers, individuals, and other corporate owners, which are tagged in the Corelogic data and shown here to contrast with iBuyer corporate owners. Panel (f) shows the median realized spread, that iBuyers earn on purchased and sold homes with 25% and 75% bands shown.



(a) Market shares of transactions

(b) Houses in inventory

(c) Dollar value of houses in inventory

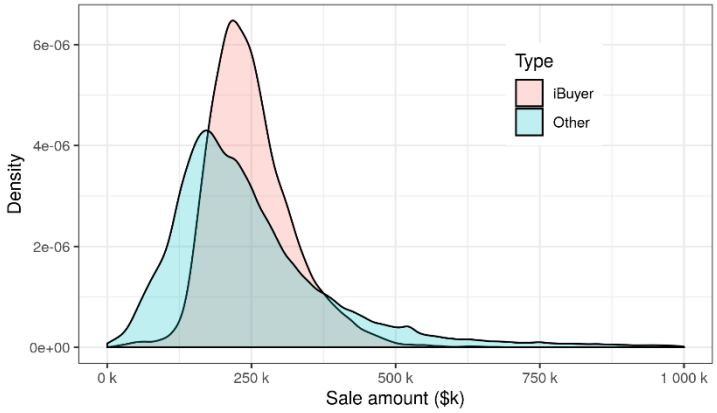
(d) iBuyer purchases

(e) iBuyer inventory turnover

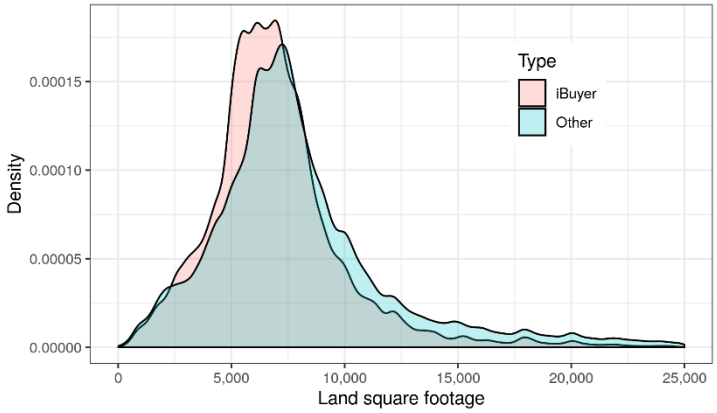
(f) iBuyer realized gross return

Figure 2: Characteristics of iBuyer Houses

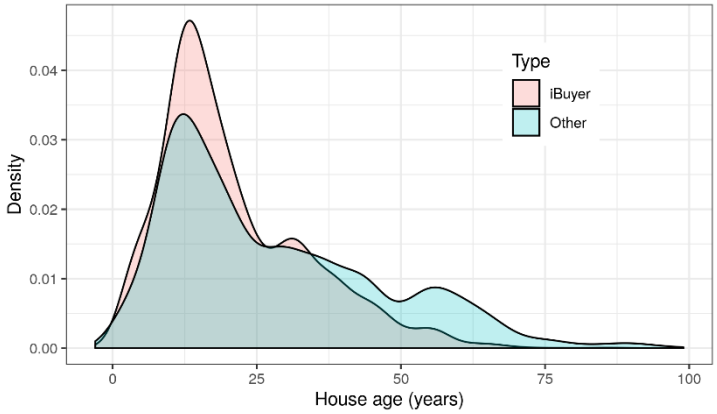
This figure shows the distribution of house prices (panel a), land square footage (panel b), and house age (panel c) for iBuyer (red) versus non-iBuyer (blue) house purchases. The figure uses transaction-level data from Corelogic between 2013 and 2018.



(a) House sale price



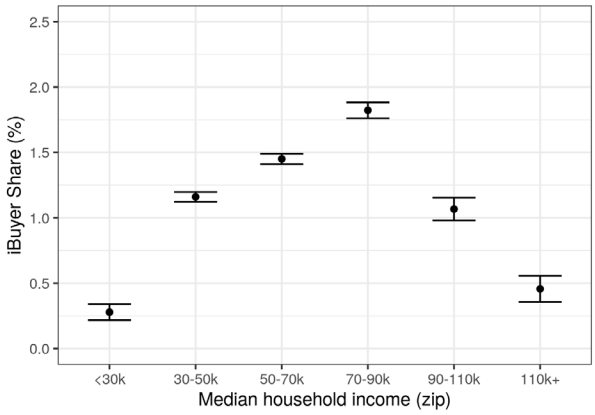
(b) Land square footage



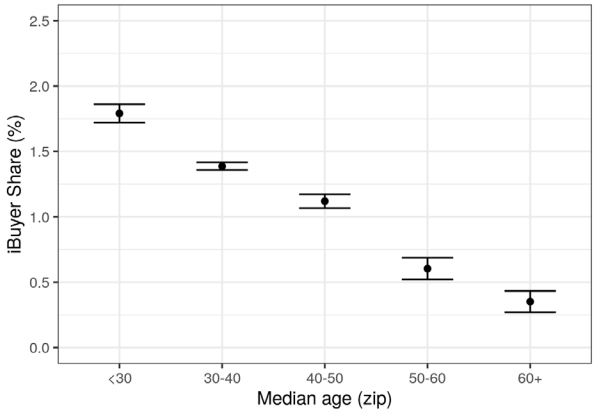
(c) House age

Figure 3: Demographics of iBuyer Markets

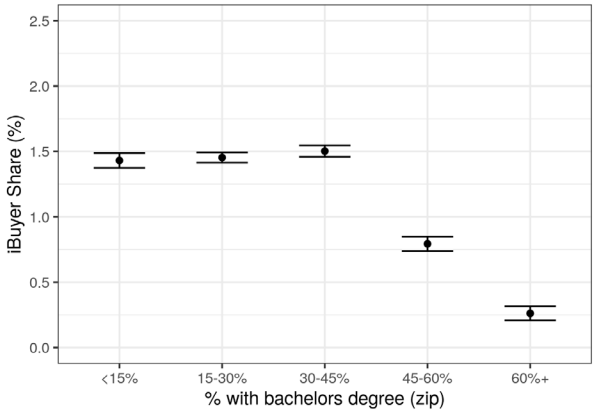
This figure shows average iBuyer market share in a zip code using Corelogic transaction data in 2013-2018 versus binned demographics at the zip code level. iBuyer market share is defined as the fraction of houses that are sold to iBuyers. Panel (a) shows iBuyer share for median household income; (b) for median age, and (c) for college education. Bars represent 95% confidence intervals.



(a) iBuyer share versus household income



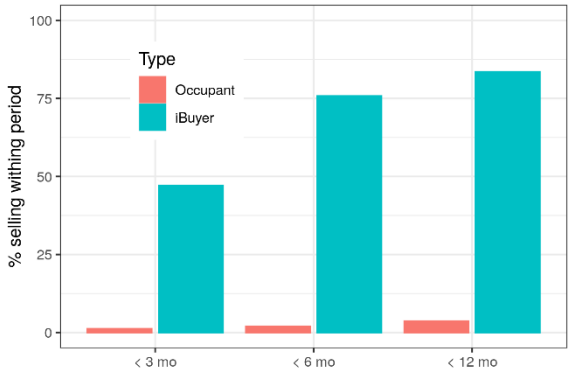
(b) iBuyer share versus age



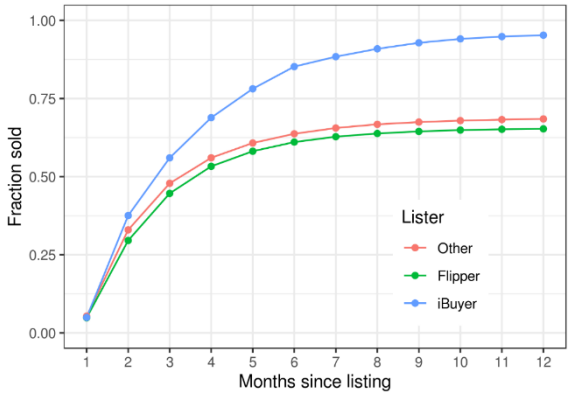
(c) iBuyer share versus % with college education

Figure 4: Inventory Holding Times and Hazard Rates of Sales of iBuyers

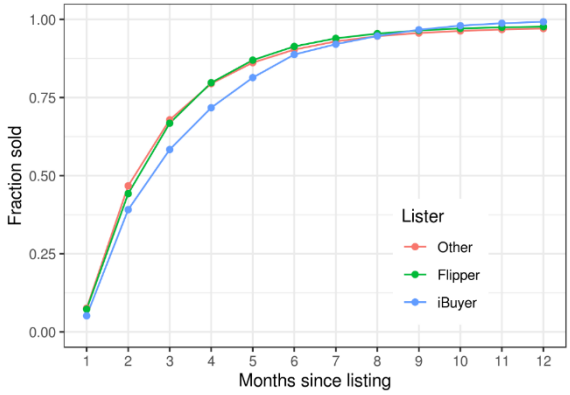
Panel (a) shows the fraction of houses purchased and sold within one year using housing transaction data between 2013 and 2017, omitting 2018 to have a full year of data for the most recent transactions. Panel (b) shows the cumulative fraction of houses sold by month after listing for all listings in MLS data between 2013 and 2017. Panel (c) shows the fraction of houses sold by month for all listings in MLS between 2013 and 2017 that eventually lead to a sale. In Panel (a), an "Occupant" is a non-iBuyer owner-occupier of the house. In Panels (b) and (c), a "flipper" is a "non-occupant" owner who lists the house within one year of purchasing it, and "other" is other (non-iBuyer, non-flipper) listers.



(a) Fraction sold by owner type



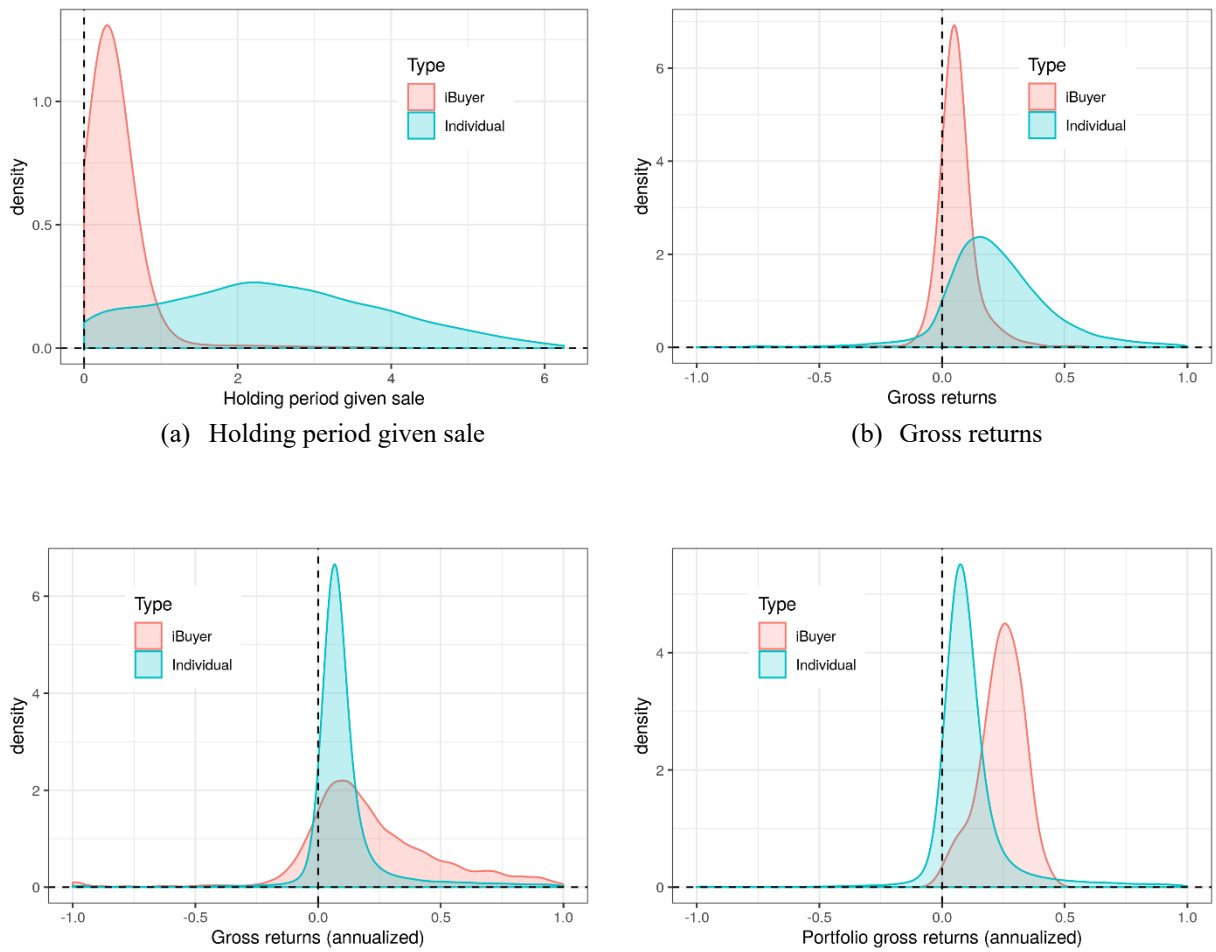
(b) Fraction sold by month among all listings



(c) Fraction sold by month among sold listings

Figure 5: Gross Returns: iBuyers versus Individuals

This figure shows holding periods and realized gross returns for iBuyers (red) versus non-corporate individuals (blue) from Corelogic transaction data. Observations are all purchases in Corelogic between 2013 and 2018 where the buyer sells the house during the sample period. Panel (a) shows the distribution of holding periods in years. Panel (b) shows raw gross returns, calculated as the change in price (in percentage terms) between purchase and sale. Panel (c) shows annualized returns, calculated by annualizing the gross returns. Panel (d) shows annualized portfolio returns, calculated by examining the average annualized return for all houses purchased in a given quarter by a single purchaser.



(a) Holding period given sale

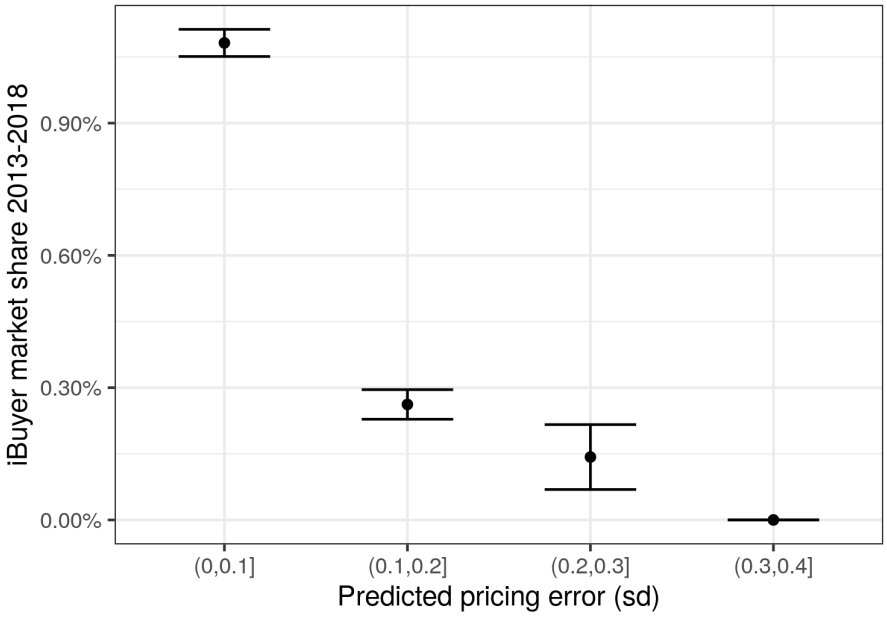
(b) Gross returns

(c) Gross returns (annualized)

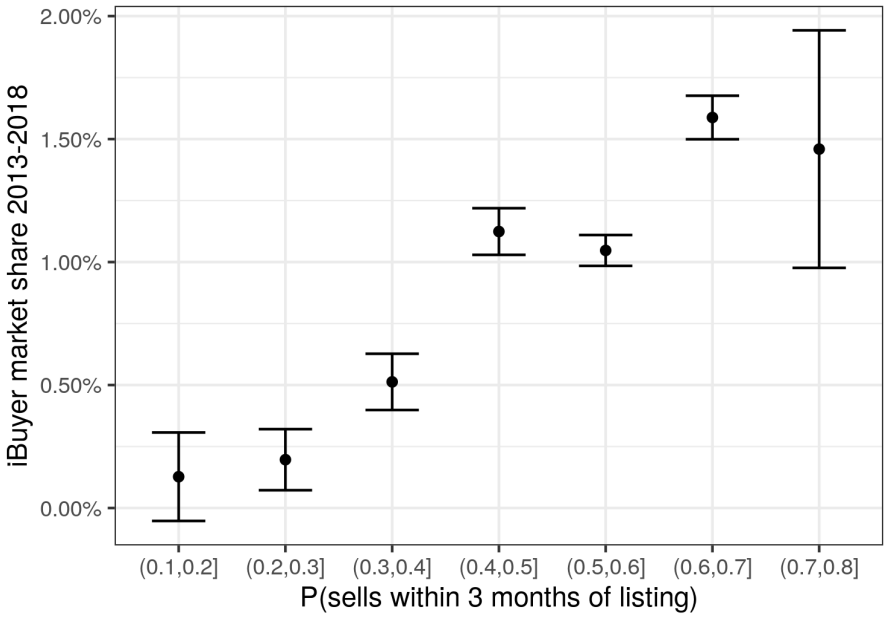
(d) Portfolio gross returns (annualized)

Figure 6: iBuyer Market Share and Pricing Errors and Liquidity

This figure shows iBuyer market share (Y axis) and various proxies for pricing errors and liquidity. We use Corelogic transaction data between 2013 and 2018. iBuyer market share is the fraction of homes purchased by iBuyer. In Panel (a), predicted pricing error is the predicted squared residual based on house hedonics, normalized by the standard deviation of the residuals. In Panel (b), predicted liquidity is the predicted probability that a house sells within 90 days of listing based on house hedonics and MLS data. Bars indicate 95% confidence intervals of standard errors of the estimates.



(a) Pricing error and iBuyer market share



(b) House liquidity and iBuyer market share

Figure 7: Equilibrium Housing Trading Model with iBuyers: Transition Paths

This figure illustrates graphically the transition paths in our equilibrium housing trading model with iBuyers. Once the homeowner becomes unmatched she wants to sell a house. She needs to decide whether to sell to an iBuyer or to list a house using a traditional selling channel. This decision will depend on her mismatch shock, the cost of accessing iBuyer, and the house repair shock. If she decides to list she needs to repair the house if it needs repairs and decide the listing price p_{hh}. She will be matched with potential buyers at the rate $\lambda F(m_s, m_b)$. Once she sells she will transition into a buyer while the buyer will transition into a matched homeowner. Alternatively she can sell to an iBuyer that offers a (quick) closing time, τ, and an acquisition price p_{ib}^b that also depends on the noisy signal the iBuyer receives regarding the house quality. If iBuyer buys a home then it subsequently lists it and decides the listing price strategy p_{ib}^s. The iBuyer will be matched with potential buyers will be matched with potential buyers at the rate $\lambda_{ib} F(m_s, m_{ib})$.

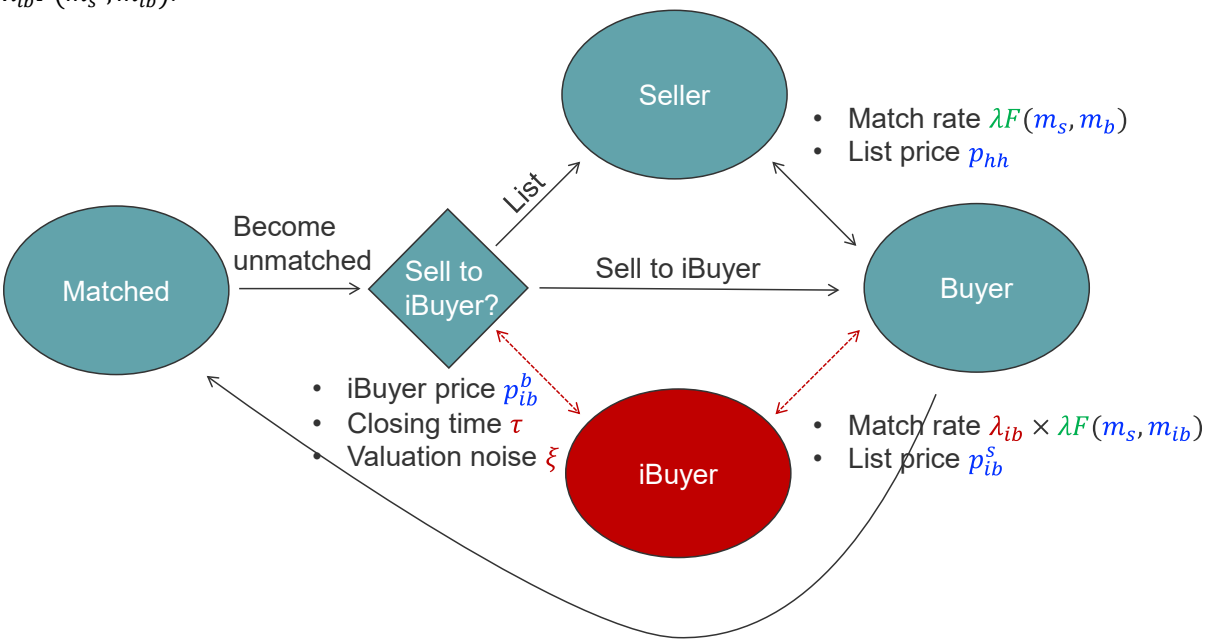
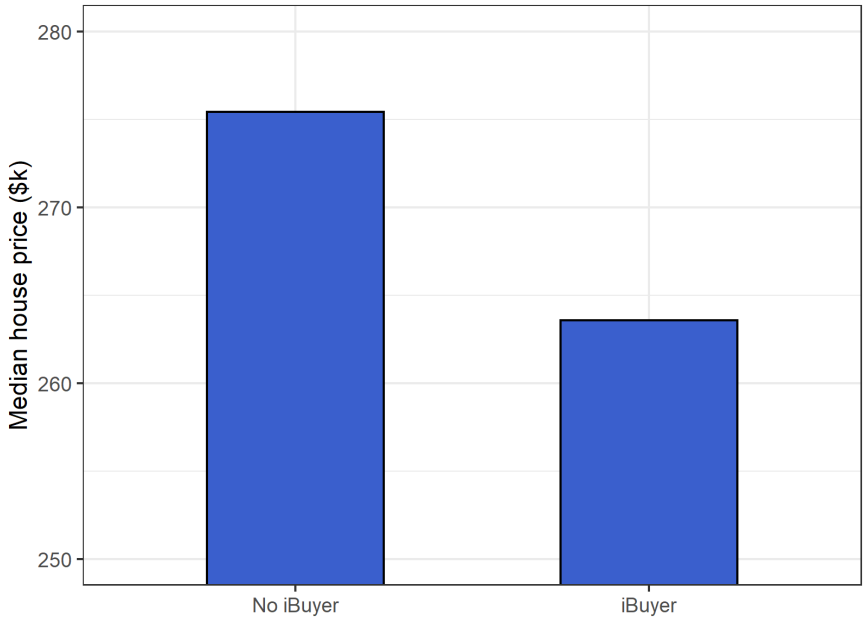
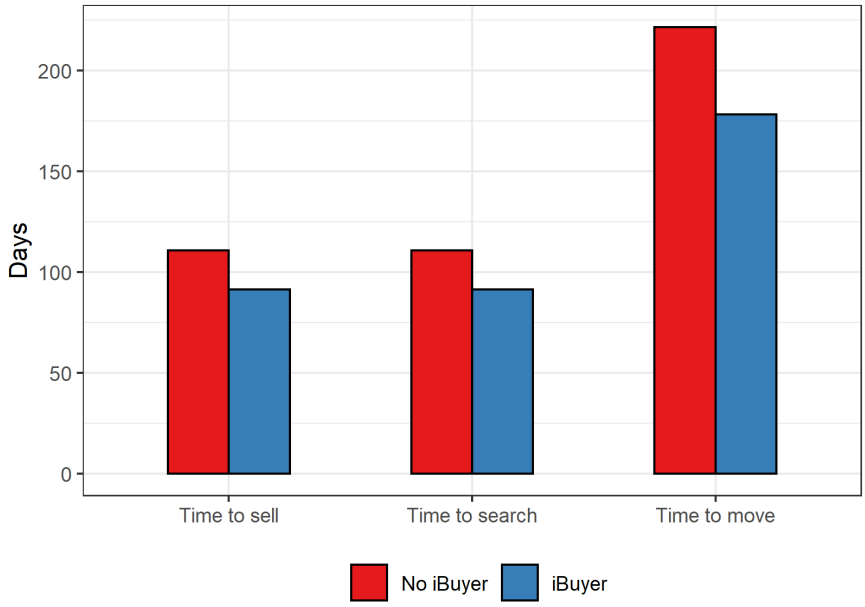


Figure 8: Equilibrium Housing Trading Model with iBuyers: Transition Paths

This figure shows counterfactual simulations from the model with and without iBuyers. Panel (a) shows median house prices in a word without iBuyers (left) and in a world without iBuyers (right). Panel (b) shows relocation times in a world without iBuyers (red bars) and a world without iBuyers (blue bars). Time to sell is the average time for a property to sell (including both iBuyer and household listers). Time to search is the average time a buyer spends looking for a new house. Time to move is the sum.



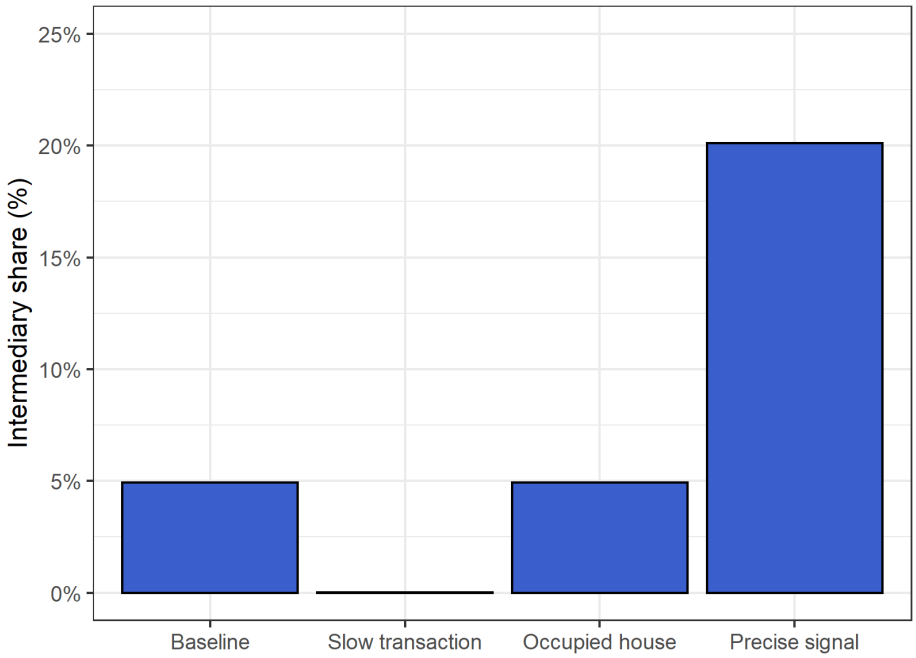
(a) iBuyer presence and median house prices



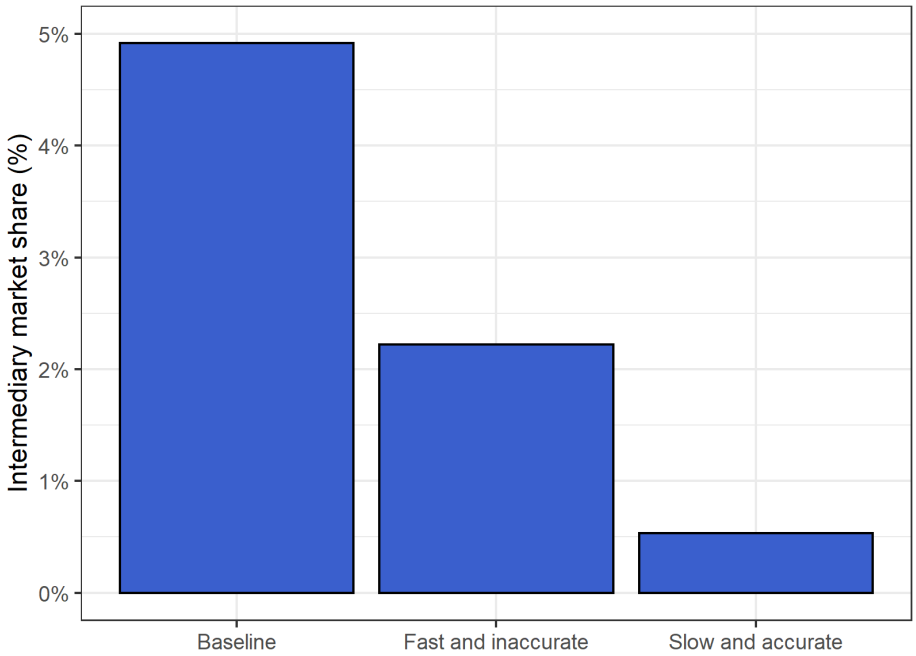
(b) iBuyer presence and relocation times

Figure 9: Intermediation Economics and the role of Technology

This figure shows counterfactual simulations from the model with and without iBuyers. Panel (a) shows median house prices in a word without iBuyers (left) and in a world without iBuyers (right). Panel (b) shows relocation times in a world without iBuyers (red bars) and a world without iBuyers (blue bars). Time to sell is the average time for a property to sell (including both iBuyer and household listers). Time to search is the average time a buyer spends looking for a new house. Time to move is the sum.



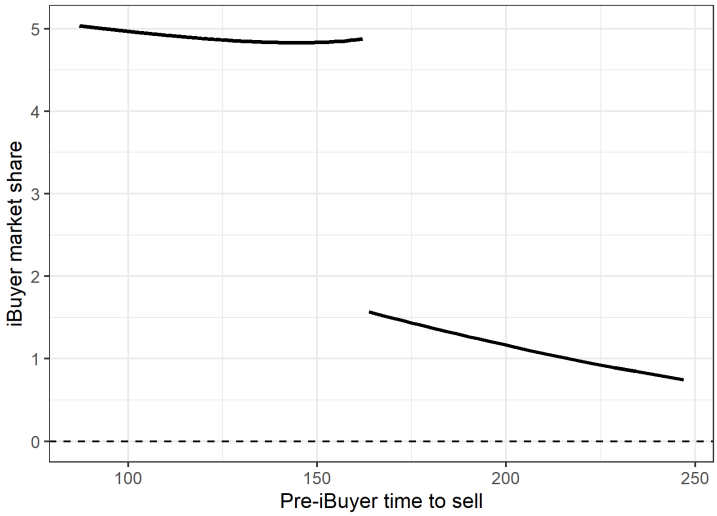
(a) Importance of intermediary characteristics



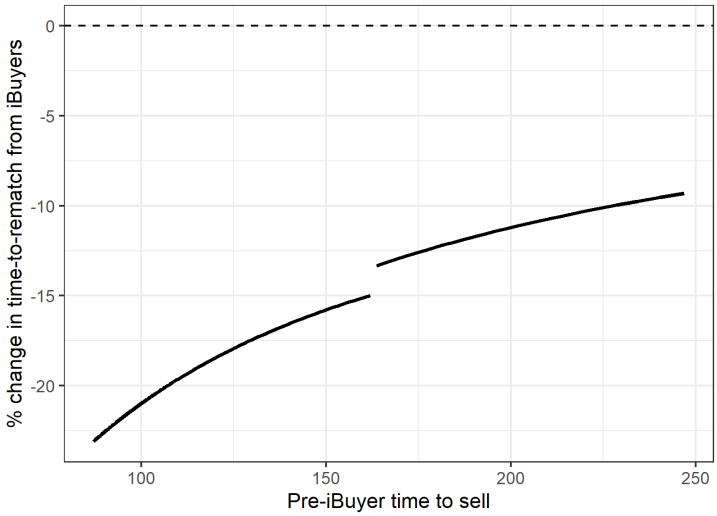
(b) Technological alternatives

Figure 10: Limits of iBuyer Market Penetration and Liquidity Provision

Panel (a) shows the association of pre-iBuyer entry housing liquidity (time to sell) and iBuyer market share implied by our calibrated model. Panel (b) shows the relative change in time to move (re-matching time) in percentage terms as a function of pre-iBuyer entry housing liquidity implied by our model.



(a) iBuyer market share

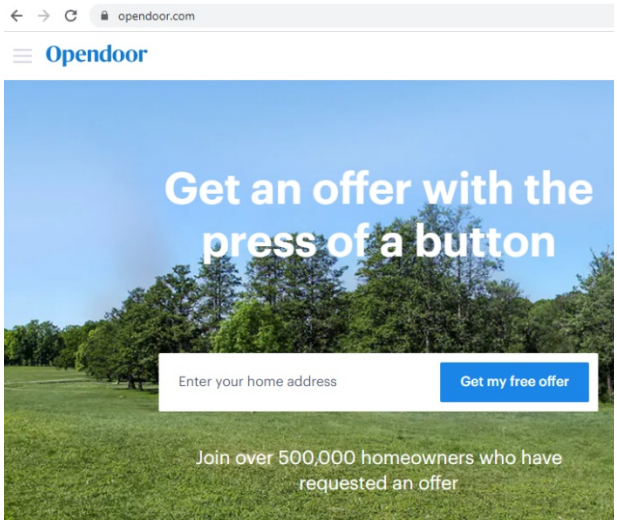


(b) Rematching time

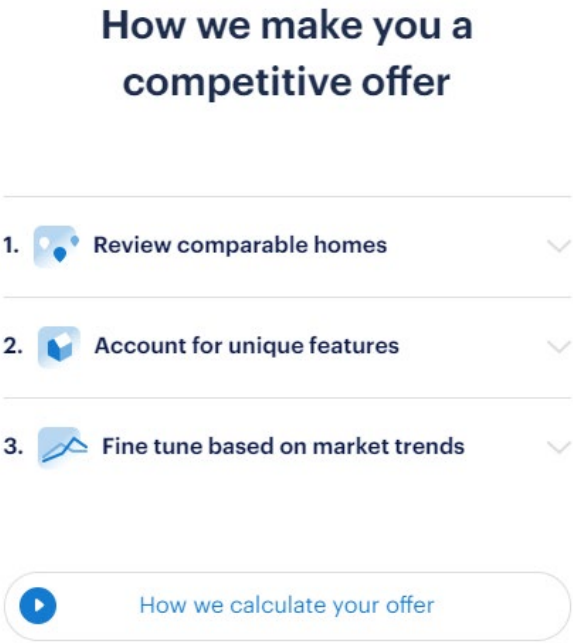
Online Appendix

Appendix A.1: Opendoor.com

This figure shows screenshots from Opendoor's website. The website was visited on January 21, 2020.



(a) Opendoor's main page



(b) Opendoor's process

(c) Opendoor's value proposition

Appendix A.2: iBuyer Classification

This section documents the classification procedure for iBuyers in the Corelogic and MLS data. The companies we consider are Opendoor,[19] Offerpad,[20] Knock,[21] Zillow Offers,[22] and RedfinNow.[23] We identify buyers and sellers in Corelogic and MLS as follows.

Corelogic: Corelogic identifies the owner name (which corresponds to the buyer in a recorded sale transaction) and the seller name. In the case of corporate owners, these are often the names of one-off legal entities with ties to the "main" iBuyer, e.g., "`OFFERPAD SPVBORROWER5 LLC`." In both cases, we identify a buyer or a seller as an iBuyer if the match one of the following regular expressions:

Company	Regular Expression
Opendoor	`opendoor`
	`open door`
	`\\<od [a-z].*`
Offerpad	`offerpad`
	`offer pad`
Redfin	`redfin`
	`red fin`
Zillow	`zillow`
Knock	`knock`

The match counts buyer or seller names that contain the string. For example, "`offerpad`" matches with the corporate entity "`OFFERPAD SPVBORROWER5 LLC`." The expression "`\\<od [a-z].*`" captures cases such as "`OD ARIZONA BORROWER 2 LLC`," which can be traced as a corporation registered at Opendoor's San Francisco headquarters. Manual inspection shows that our search strings do not leave out any common buyers or sellers, but there is the possibility of being underinclusive of transactions with unusual corporate entity names. A transaction has an iBuyer seller if we find a match in the seller's name. A transaction has an iBuyer buyer if we find a match in the owner's name.

MLS: We use the same set of regular expressions as above. A listing has an iBuyer seller if we find a match in the listing agent's name or the owner's name. A listing has an iBuyer buyer if we find a match in the buyer office name or the buyer agent name. As above, manual inspection shows

[19] https://www.opendoor.com/
[20] https://www.offerpad.com/
[21] https://www.knock.com/
[22] https://www.zillow.com/offers/
[23] https://www.redfin.com/now

that our search strings do not leave out common buyers or listers, but there is the possibility that our search is underinclusive of iBuyer transactions with unusual corporate entity names.

Appendix A.3: Corelogic and MLS Matching and Tie-out

This table shows the matching rate and consistency between Corelogic transactions and MLS listings. Panel A presents the fraction of single family, arms-length transactions in Corelogic with a match in MLS. *Day match* is the fraction of Corelogic transactions with a sale in MLS where the property ID and sale date matches exactly. *Week*, *month*, and *quarter match* is the fraction of Corelogic transactions with a sale in MLS where the property ID matches exactly and the sale date is within seven days, in the same calendar month, or in the same quarter, respectively. Panel B shows the consistency of reported MLS and Corelogic sale prices by various match windows and buyer/seller types: *Cor(log(MLS),log(Corelogic))* is the correlation between the log MLS sale price and the log Corelogic sale price. *Exact price match* is the fraction of matches where the sale price matches exactly; *|Deviation| < 1%* is the fraction of matches where the absolute deviation is within 1%. *|Deviation| < 5%* is the fraction of matches where the absolute deviation is within 5%. *Mean(|Deviation|)* is the mean of the absolute value of the deviation, e.g., 0.120 means that the mean of the absolute value of Corelogic price divided by MLS price minus one is 0.120.

Panel A: Corelogic transactions with an MLS listing

Year	N	# iBuyer Buys	# iBuyer Sales	Day match	Week match	Month match	Quarter match
All	182,486	6,555	9,922	0.145	0.343	0.411	0.537
2010	95,146	2	0	0.069	0.139	0.166	0.218
2011	101,780	1	1	0.090	0.189	0.227	0.292
2012	109,205	0	2	0.125	0.266	0.315	0.408
2013	106,800	3	0	0.158	0.345	0.417	0.548
2014	132,406	9	1	0.145	0.370	0.448	0.587
2015	148,971	447	333	0.149	0.342	0.411	0.538
2016	159,916	1,407	1,185	0.160	0.399	0.479	0.634
2017	145,648	1,583	2,050	0.170	0.427	0.509	0.667
2018	128,340	2,241	3,964	0.186	0.447	0.531	0.685
2019	54,274	862	2,386	0.176	0.445	0.536	0.681

Panel B: Corelogic and MLS sale price consistency

Match Window	Cor(log(MLS),log(Corelogic)	Exact price match	|Deviation| < 1%	|Deviation| < 5%	Mean(|Deviation|)
All Transactions					
Day	0.956	0.651	0.760	0.808	0.120
Week	0.967	0.780	0.851	0.891	0.068
Month	0.967	0.807	0.870	0.907	0.065
Quarter	0.966	0.830	0.885	0.921	0.061
iBuyer Buys					
Day	0.948	0.881	0.952	0.976	0.021
Week	0.988	0.916	0.980	0.992	0.004
Month	0.989	0.907	0.978	0.992	0.004
Quarter	0.859	0.355	0.417	0.624	0.076
iBuyer Sells					
Day	0.773	0.708	0.776	0.808	0.057
Week	0.797	0.856	0.891	0.913	0.048
Month	0.819	0.868	0.902	0.923	0.043
Quarter	0.837	0.892	0.919	0.939	0.040

Appendix A.4: Listing Dynamics Robustness

This table shows a robustness check on the pricing dynamics of listings where an iBuyer is the seller. In particular, allows for the possibility that listers withdrawal unsuccessful listings and relist shortly thereafter. Therefore, in contrast to the main Table in the body of the paper, the outcome variables (total listings, whether a sale occurs, days between first listing and sale, and sale-to-first listing price) are augmented with outcomes from subsequent relistings that occur within 30 days of the time that the first listing is withdrawn. Data are from MLS provided by ATTOM Data between 2013 and 2018 at the combined listing level. *Log first price* is the log of the first listing price. *Mentions renovations* is an indicator for whether the listing description describes the house as being renovated, i.e., includes "renovation," "refurbish," or "remodel." *Total listings* is the number of price adjustments in a given listing spell. *Leads to sale* is an indicator for whether the listing leads to a sale. *Days on market* is the number of days between the first listing and the sale. *Sale-to-list* is the sale price divided by the initial listing price. In Panel (a), Columns (1)-(5) and (8) are linear models; (6) and (7) are Cox Proportional Hazard Rate models. Columns (1)-(4) and (6) consider all listings. Columns (5), (7), and (8) consider only listings leading to sales. A *Flipper* is an absentee owner who lists the house within one year of purchasing it. The *iBuyer* and *Flipper* categories are measured relative to the base category of other lister. All columns include house controls including square footage, whether the house is multistory, and house age. The linear models include zip times quarter fixed effects. Standard errors are shown in parentheses.

	Dependent variable:						
Model	Linear	Linear	Linear	Linear	Hazard	Hazard	Linear
Outcome	Relists within 30 days	Total listings	Leads to sale	Days-on-market	Days-on-market	Days-on-market	Sale-to-list
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Flipper	0.065	0.387	-0.009	4.862	-0.101	-0.099	-0.002
	(0.003)	(0.010)	(0.002)	(0.391)	(0.005)	(0.005)	(0.0003)
iBuyer	-0.021	1.444	0.149	29.733	0.240	-0.095	-0.007
	(0.066)	(0.064)	(0.011)	(2.141)	(0.028)	(0.028)	(0.001)
Hedonic controls	Y	Y	Y	Y	Y	Y	Y
Zip-Quarter FE	Y	Y	Y	Y	N	N	Y
Sample	Failed first list	All		Sales only	All	Sales only	
Observations	265,805	887,208	887,208	653,385	791,798	653,385	615,834
R^2	0.157	0.341	0.251	0.385	0.070	0.057	0.142

Appendix A.5: Instrument Regional Analysis

This figure shows the first stage relationship between predicted iBuyer market share and actual market share at the zip code level used as an instrument in the regional analysis. Predicted propensity, the x-variable, is 25 equally-sized bins of predicted iBuyer market share at the zip-code level. The y-axis is the average realized iBuyer market share in each zip code falling within the predicted market share bin. The dashed line is a best-fit linear regression, with the shaded region showing the 95% confidence interval.

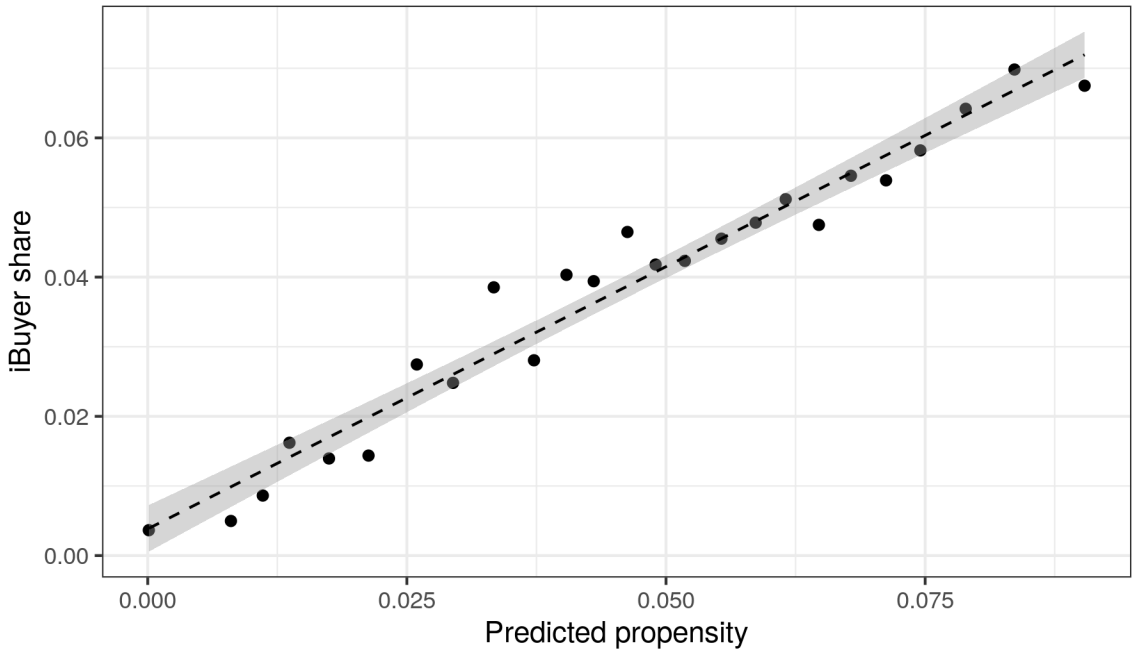


Table A.6: Spread regressions with alternate price index

This table shows the regressions and decomposition of the iBuyer spread using an alternate index measure. In particular, we use Zillow single family house price indices as opposed to median transaction prices, which account for changes in quality of the housing stock.

Spread Regressions with Redfin Price Indices

	Dependent variable:					
	Gross Return (ann)		IndexSpread (ann)		NonIndexSpread (ann)	
	(1)	(2)	(3)	(4)	(5)	(6)
iBuyer	0.087	0.075	0.010	0.020	0.072	0.049
	(0.002)	(0.002)	(0.001)	(0.0005)	(0.002)	(0.002)
House Controls	N	Y	N	Y	N	Y
Zip x Quarter FE	N	Y	N	Y	N	Y
Observations	102,140	94,499	102,140	94,499	102,140	94,499
R^2	0.014	0.225	0.0004	0.581	0.011	0.231