

NBER WORKING PAPER SERIES

WHY IS INTERMEDIATING HOUSES SO DIFFICULT?
EVIDENCE FROM IBUYERS

Greg Buchak
Gregor Matvos
Tomasz Piskorski
Amit Seru

Working Paper 28252
<http://www.nber.org/papers/w28252>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2020, Revised July 2022

We thank Sumit Agarwal, Eric Budish, Peter DeMarzo, Rebecca Diamond, Darrell Duffie, Claudia Robles-Garcia, Sophia Gilbukh, Adam Guren, Bob Hall, Tim McQuade, Jim Poterba, Adriano Rampini, Martin Schneider, Johannes Stroebel, Pierre-Olivier Weill, Randall Wright, Jeff Zwiebel and participants at Berkeley, Chicago, Kellogg, Stanford, Columbia Workshop in New Empirical Finance, Duke Fuqua, Georgia State, National University of Singapore, Johns Hopkins, HKUST, SMU, UCLA, NABE, NBER Summer Institute (Real Estate and Household), NBER Corporate Finance, Search and Matching in Macro and Finance Virtual Seminar Series, Stanford Institute of Theoretical Economics, and Western Finance Association 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 The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Greg Buchak, Gregor Matvos, Tomasz Piskorski, and Amit Seru. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Why is Intermediating Houses so Difficult? Evidence from iBuyers
Greg Buchak, Gregor Matvos, Tomasz Piskorski, and Amit Seru
NBER Working Paper No. 28252
December 2020, Revised July 2022
JEL No. G0,G2,G5,L0,R20

ABSTRACT

We study the frictions in dealer-intermediation in residential real estate through the lens of “iBuyers,” technology entrants, who purchase and sell residential real estate through online platforms. iBuyers supply liquidity to households by allowing them to avoid a lengthy sale process. They sell houses quickly and earn a 5% spread. Their prices are well explained by a simple hedonic model, consistent with their use of algorithmic pricing. iBuyers choose to intermediate in markets that are liquid, and in which automated valuation models have low pricing error. These facts suggest that iBuyers’ speedy offers may come at the cost of information loss concerning house attributes that are difficult to capture in an algorithm, resulting in adverse selection. We calibrate a dynamic structural search model with adverse selection to understand and quantify the economic forces underlying the tradeoffs of dealer intermediation in this market. The model reveals the central tradeoff to intermediating in residential real estate. To provide valuable liquidity service, transactions must be closed quickly. Yet, the intermediary must also be able to price houses accurately to avoid adverse selection, which is difficult to accomplish quickly. We find that low underlying liquidity exacerbates adverse selection. Our analysis suggests that iBuyers’ technology provides a middle ground: they can transact quickly limiting information loss. Even with this technology, intermediation is only profitable in the most liquid and easy to value houses. Therefore, iBuyers’ technology allows them to supply liquidity, but only in pockets where it is relatively least valuable. We also find limited scope for dealer intermediation even with improved pricing technology, suggesting that underlying liquidity will be an impediment for intermediation in the future.

Greg Buchak
Graduate School of Business
Stanford University
655 Knight Way
Stanford, CA 94305b
buchak@stanford.edu

Tomasz Piskorski
Columbia Business School
3022 Broadway
Uris Hall 810
New York, NY 10027
and NBER
tp2252@columbia.edu

Gregor Matvos
Kellogg School of Management
Northwestern University
2211 Campus Drive
Global Hub 4361
Evanston IL, 60208
and NBER
gregor.matvos@kellogg.northwestern.edu

Amit Seru
Stanford Graduate School of Business
Stanford University
655 Knight Way
and NBER
aseru@stanford.edu

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 The substantial frictions in the buying and selling of homes make housing quite illiquid. These frictions in the housing market can affect households' matching with the appropriate houses and affect household mobility. Imagine a household, which wants to sell its house to pursue a new job. To purchase the house in the new location, they may have to first sell their current house. If the process takes a while, they may rent a suboptimal house in the new location during this period, which could take over a year. In fact, the difficulty in selling and purchasing a new home may force the household to abandon the new job entirely. These difficulties in homeowner-to-homeowner sales suggest a natural role for dealer intermediation: a homeowner could sell their house directly to an intermediary, which would resell it once it finds a buyer. The homeowner could then purchase a new house without delay. An immediacy discount on the sale to the intermediary would split the gains from trade. Yet, despite seemingly large demand for such intermediation, until recently, such transactions were rare, suggesting that dealer intermediation in this market is subject to substantial frictions.

This has changed with the entry of iBuyers, a recent technological disruption in the real estate market. iBuyers, such as Opendoor, Offerpad, RedfinNow, and Zillow Offers², are real estate companies that use an automated valuation model and other technology to make cash offers on homes quickly, within hours, through their on-line acquisition platforms.³ In other words, iBuyers offer exactly the kind of dealer intermediation that has been absent from this market until now. In this paper, we use iBuyers as a window to study frictions in dealer intermediation in residential real estate, and the role that their technology has played in making dealer intermediation viable.

This paper proceeds in two main steps. We first use rich micro data to document where and how iBuyers' intermediate in real estate. These facts generate insights on the main economic frictions that limit intermediation in real estate. We then build a calibrated structural search model that explicitly incorporates the economic forces we document. The model reveals the central tradeoff to intermediating in residential real estate. Intermediation is only valuable if intermediaries can purchase houses quickly. However, with the current technology, speed comes at the cost of information loss concerning house attributes. Because this information is likely available to the property seller, intermediaries are vulnerable to adverse selection. The intermediation problem is exacerbated in homes with low underlying liquidity, which expose the intermediary to cost shocks, and reinforce adverse selection. A slower or less precise "low-tech" dealer would not be able to

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

² Zillow decided to exit the iBuying market in November 2021 (see Section V.C for more discussion).

³ Appendix A.1 shows screenshots from Opendoor's website as of 2020.

survive in this market. Even with perfect technology, the extent of intermediation is limited to 20% of transactions. Finally, we use the model to study the limits to liquidity provision by intermediaries in durable goods markets more broadly.

We begin 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. We first document that iBuyers act as a dealer intermediary. They allow the seller to forgo listings and sell quickly. The average time for a listed house to sell in iBuyer markets is roughly 90 days. Sellers to iBuyers are almost 30pp more likely to forgo listings and sell to iBuyers directly, saving this time. Once they acquire a property, they 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 use traditional selling channels relying on multiple listing services to dispose their inventory. They earn a positive spread (gross return) of about 5% on the houses they buy and sell.

A large portion of this spread is due to the 3.6pp discount of home value at which iBuyers purchase homes. 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 observable characteristics. This estimate suggests that an average seller to iBuyers is willing to forgo \$9,000 to sell their house immediately. We note that iBuyer price discount does not reflect fees charged by iBuyers. iBuyers charge comparable if somewhat higher fees than the typical real estate agent fees of about 6% or less associated with a traditional house sale. Hence, selling to iBuyers involves broadly similar if not higher fees compared to a traditional sale and a substantial price discount (on average). This substantial iBuyer price discount suggests a significant demand for dealer intermediation by households. Yet, until the entry of iBuyers, such transactions were rare, indicating large frictions in supplying liquidity. We learn about these frictions from observing which segments of the housing market iBuyers choose to intermediate.

We document that iBuyers choose to intermediate in markets where automated valuation models have low pricing error. These facts suggest that iBuyers are concerned that their choice of providing speedy offers comes at the cost of information loss, specifically, information that is difficult to capture in an algorithm. We first confirm iBuyers' claims that they use automated valuation models, which allow them to offer speedy transactions. We follow the strategy in Buchak et al. (2018) and show that observable property characteristics and zip times quarter fixed effects explain over 80% of the variation in prices for iBuyer transactions, versus 68% of the variation in non-iBuyer transactions. The lower R^2 of non-iBuyers suggests that they use information that is not captured by standard hedonics when pricing houses. For example, it is difficult for an algorithm

to capture the views, the quality of the neighboring park, or if the neighbors' house is poorly maintained. These aspects affect the value of a property but are difficult to code and capture with an algorithm. Because such information is likely also available to the property seller, iBuyers are vulnerable to adverse selection.

If iBuyers' algorithmic pricing exposes them to adverse selection, then they should intermediate in segments, in which their informational disadvantage is smallest: those in which their algorithmic valuation models employing hard hedonic information works well, leaving little room for other information to affect prices. Indeed, we find that iBuyers' market share declines substantially in segments in which pricing errors using hedonic models are largest. Even when they are present in a specific geographic market, they choose not to purchase any houses in the quartile of houses that are most difficult to price with hedonic models. Consistent with adverse selection, we also find that iBuyers earn the lowest returns when they transact difficult to price houses. These facts suggest that dealer intermediation in real estate, which requires providing speedy offers comes at the cost of potential adverse selection.

We also find that iBuyers choose to intermediate in houses with the highest underlying liquidity. Even if they are present in a geographic market, iBuyers do not transact in houses whose probability of selling in less than 3 months is below 20%. This may seem surprising. One would expect more illiquid houses have higher demand for liquidity provision. These results suggest that despite higher demand for liquidity, it may be difficult to supply liquidity for these houses, and the latter effect dominates. Our evidence suggests that liquidity provision is efficient only when houses are already relatively easy to value and liquid—i.e., when additional liquidity is least valuable.

Next, we develop a search-based equilibrium model of house trading with iBuyers. The model has several goals. In the data we find evidence consistent with several frictions that may impede intermediation. We model these frictions to understand how they affect house choices, trading, and prices in equilibrium. The reduced form evidence we present suggest the presence of frictions, but without our model it does not allow us to understand which frictions are central in limiting intermediation, or how technology plays a role. We calibrate the model to the data to identify the main tradeoff faced by iBuyers when they intermediate in this market. We also quantify which aspects of technology were central for these types of intermediaries to succeed. Last, we use the model to better understand the overall limitations to intermediation in this market, even if technology improves beyond its current level.

We build on a standard continuous time search and matching equilibrium model of homeownership, into which we introduce a dealer intermediary: iBuyers. 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 allow for homeowner heterogeneity in impatience to move, captured through differential costs of being matched to the wrong house, which captures more or less pressing reasons for moving such as changing a job or the merely the desire to upsize the home. Critically, the homeowner can only own one house at a time, and in consequence, she must move sequentially: she first needs to find a buyer for her current house using the listings process, and then look to buy a new house. Once she has bought her new house, she once again becomes a matched homeowner. Instead of listing her house and waiting for it to sell, a homebuyer can sell the house to iBuyers who lists it for resale. iBuyers are not constrained to hold only one house at a time, which allows them to intermediate. They possess a pricing technology that allows them to purchase almost immediately—the speed at which they can undertake a transaction is one of the features we study. This speed comes at the cost of information loss, which results in adverse selection. This technology may also allow them to find buyers faster—they may have a different matching function than other sellers. Last, iBuyers do not obtain utility flows from homeownership although they still must pay maintenance costs.

We calibrate the model to the post iBuyer entry period, using the model to match the features of the data such as iBuyer penetration, the discount iBuyers pay for houses, the time their houses stay in inventory relative to other households' time on the market, as well as the fact that iBuyer penetration is lower in less liquid market and markets in which houses are difficult to price. The model can replicate the patterns we document, suggesting that the frictions we built into the model are consistent with the data. To provide external validation of the model, we examine the predictions of the model with respect to iBuyer entry. Intuitively, we perform a difference in difference exercise in the model and in the data, where the entry of iBuyer is the “natural experiment”. The model estimated on post-iBuyer entry data does a good job matching the reduced form difference in difference estimates of increases in equilibrium prices and transaction speeds, even in transactions where iBuyers are not directly involved. Thus, this validation exercise provides further comfort in the model estimates.

In the data we find evidence consistent with several frictions that may impede intermediation. The calibration reveals that the central tradeoff to intermediating in residential real estate markets comes down to three forces: speed, information quality, and liquidity. First, the speed of transaction is critical. Consider a slow, but precise “low-tech” intermediary, which does not do algorithmic pricing and therefore cannot close the transaction quickly. Even with perfect information, their ability to provide intermediation would essentially vanish, and their share of transactions would fall nearly to 0% in the current iBuyer markets. The model tells us that almost all iBuyer market transactions are with impatient sellers, a prediction for which we find reduced form support in the fact that sellers to iBuyers are more likely to leave markets post sale, which corresponds to our interpretation that impatient sellers are motivated to move for reasons such as

a new job or other forced geographic relocation. With this in mind, the extreme sensitivity of iBuyer market share to execution speed is intuitive, because the ability to close transactions quickly is the main benefit of liquidity provision through dealer intermediation.

This speed comes at the cost of information loss about homes, which potentially results in adverse selection. Our model finds a role for adverse selection, showing that decreases in the precision of their pricing signal reduces market share. Eliminating iBuyers' ability to distinguish low- and high-quality homes (aside from features which can be contracted on) would reduce iBuyer market share from roughly 5% to just above 1%.

The model also explains why iBuyers only provide liquidity in already liquid markets. A longer time on the market exposes iBuyers to maintenance shocks. Because the house is unoccupied, these shocks loom large for iBuyers—for example, being arm's length, they may notice a leaking roof only after substantial damage. This already lowers the profits from dealer intermediation, and thus its viability. Lower profits from intermediating in low liquidity houses also force iBuyers to lower purchase prices. Through this secondary channel, low underlying liquidity also exacerbates adverse selection: as liquidity decreases and holding costs increase in expectation, iBuyers are forced to offer less favorable prices. Owners of (unobservable) good homes become less willing to transact, leaving more owners of (unobservable) bad homes in the iBuyer market. The share of homes on which iBuyers are adversely selected thus increases.

While iBuyers have difficulty competing in illiquid markets, our model also shows that they also struggle to compete in very liquid “hot” markets. Intuitively, when we simulate a market where houses trade much faster than historical average sale times, iBuyer offer is not sufficiently fast relative to an ordinary listed sale. Because iBuyers need to buy at a discount, good types are for the most part no longer willing to sell to them. Even among these bad types, iBuyer market share is extremely low because the marginal benefit of selling quickly to the intermediary (at a discount) relative to a (sped up) listing becomes very small if traditional listing channel becomes very fast.

We find that a major quantitative drawback of the iBuyer business model is that the houses are left unoccupied and valuable service flows are lost. Being able to frictionlessly rent the home to an (even unmatched) occupant during the sale process would increase iBuyer market share by nearly three times: they would be willing to pay more for homes and make up the difference through rental payments. We also find that iBuyers are no better at selling houses than other households. Though they sell their house through a listing process, one could have thought that they could have used their own and related party websites (Zillow, RedFin), which are heavily trafficked, to improve matching. Data suggests limited advantage from those in terms of speed of selling homes they intermediate in.

Our analysis points to why dealer intermediation in real estate has been so limited up to this point. To provide valuable liquidity service, transactions must be closed quickly. At the same time, the intermediary must be able to price houses precisely, which is difficult to accomplish quickly. Low underlying liquidity exacerbates this problem. Our analysis suggests that iBuyers technology has found a middle ground, which allows them to transact quickly with limited information loss. Even with this technology, they have been limited to the most liquid and easy to value houses, particularly focusing on markets where the cost of valuation mistakes are likely to be low. In other words, iBuyers comparative advantage allows them to add liquidity to the market, but only in pockets of the market where liquidity was already very high or very low. We also find limited scope for dealer intermediation even with improved pricing technology, and strong reasons for iBuyers to avoid markets where valuation will be difficult, such as older homes, suggesting that underlying liquidity will be an impediment for intermediation in the future. The model suggests that one large avenue for technology to improve intermediation would be to improve matching. Matching frequency could increase in the market overall, because of technology improvements in listings. iBuyers could also find better matches through related party websites. Either of these improvements could increase the scope for dealer intermediation in this market, and improve liquidity in one of the most important asset markets for households.

Related Literature

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 online fintech lenders (Buchak et al. 2018; Fuster et al. 2019) and the impact of the sharing economy on the housing market (e.g., Calder-Wang 2019). We add to this literature by studying 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-banks in the housing market (e.g., Buchak et al. 2020; Jiang et al. 2020).

We contribute to the literature on frictions in matching households to houses. We build on a large literature of quantitative models search and matching models (Wheaton 1990, Genesove and Mayer 1997, Genesove and Han 2012; Head et al. 2014; Landvoigt et al. 2015; Anenberg 2016; Guren 2018; Anenberg and Bayer 2020; Piazzesi, Schneider, and Stroebel 2020; Guren and McQuade 2020; Andersen et al. 2020; Rekkas, Wright, and Zhu 2020). The presence of search frictions and household balance sheet constraints provides a natural demand for intermediation. The existing literature has focused on the matchmaking role of intermediaries, i.e., real estate brokers (Levitt and Syverson 2008; Hendel, Nevo, and Ortalo-Magné, 2009; Barwick and Pathak 2015; Barwick, Pathak, and Wong 2017; Gilbukh and Goldsmith-Pinkam 2019; Agarwal et al. 2021), and speculators (Fu, Qian, and Yeung, 2015; Chinco and Mayer 2016; DeFusco, Nathanson, and Zwick 2017; Nathanson and Zwick 2018; Mian and Sufi 2022). In contrast, our focus is instead

on trying to understand the lack of balance sheet (dealer) intermediation, which should naturally occur in a market in which consumers' demand immediacy. We contribute by introducing balance sheet intermediaries into the existing models and analyze the frictions that limit dealer intermediation in residential real estate markets.

Our analysis is also related to a large body of work focusing on dealer intermediation and trading frictions in decentralized asset markets. Perhaps the closest are to our work are the structural and quantitative models of Gavazza, Lizzeri, and Roketskiy (2014) and especially Gavazza (2016). Similar to Gavazza (2016), intermediaries provide the benefits of faster trade, as in Grossman and Miller (1988), and direct trade can occur between consumers in addition to trading with intermediaries. While Gavazza (2016) focuses on the role of bargaining and the entry / exit decisions of intermediaries, we instead focus on the role of asymmetric information faced by the intermediary. In Gavazza (2016), frictions can lead to excess intermediation, while in our setting, frictions limit the extent of valuable intermediation.

More broadly, our paper is also related to the large theoretical search and matching models in the presence of dealer intermediaries (see Weill 2020 for a survey). Relative to the canonical models such as Rubinstein and Wolinsky (1987), Grossman and Miller (1988) or Duffie, Garleanu, and Pedersen (2005), our central focus is on the intermediary that allows for immediate trade but the cost of immediacy is that the intermediary is exposed to asymmetric information relative to other market participants. In this way, we contribute to the literature analyzing the impact of asymmetric information in these markets, similar to Hendel and Lizzeri (1999), Guerrieri, Shimer, and Wright (2010), Guerrieri and Shimer (2014, 2018), Eisfeldt (2014), Chang (2018) and Lester et al. (2019) who study the interaction between asymmetric information and other frictions. We differ from these papers with our focus on intermediaries, who are subject to asymmetric information, which we assume is not present in the rest of the market, and study how this information wedge and other forces limits the scope of intermediation. Our tractable model matched to the data allows us to provide quantitative answers to these questions.

Our work is also related to the work on the housing market including recent work that studies the role of asymmetric information in real estate markets (e.g., Kurlat and Stroebel 2015; Stroebel 2016; Indarte 2021; Gupta and Hansman 2022); what differs is that we focus on impact of asymmetric information in limiting intermediation. Our paper is also broadly related to recent quantitative studies of housing and mortgage markets (e.g., Benetton 2021; Berger et al. 2017, Favilukis, Ludvigson, and Van Nieuwerburgh, 2017; Beraja et al. 2019; Greenwald 2018; Ganong and Noel 2020; Kaplan, Mitman, and Violante, 2020; Buchak et al. 2020; Gorback and Keys 2020; Wong 2021), which are interested in the setting of house prices as we show that presence of dealer intermediaries could affect the overall house price level.

Last, our discussion of iBuyers response to Covid and Zillow exit also relates our work to models of dealer intermediation, which study how intermediaries respond to market level shocks. For example, the model of Weill (2007) shows that intermediaries may respond to temporary shocks by providing liquidity, i.e., they absorb shocks. In contrast, our model suggests that intermediaries in durable goods markets subject to asymmetric information may provide less liquidity once exposed to such shocks, consistent with the data.⁴

II. Data and Institutional Background

II.A. Data Sources

Transactions Data: We use Corelogic deeds records data on housing transactions from five markets with a 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 square 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,

⁴ Lagos, Rocheteau, and Weill (2011) show that, when trading frictions are large, well-capitalized dealers may be unwilling to provide liquidity during crises, and even under circumstances when it would be socially efficient for them to do so. This echoes the behavior of iBuyers in very illiquid markets, though the economic mechanism is different from the one considered by us.

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 these summary statistics.

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 and absentee owners, and impose a series of cleaning and filtering steps.⁵

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

⁵ For example, “Corp,” “INC,” “LLC,” and so on. Moreover, we remove white space, names where the first or last name has one or fewer characters (e.g., is an initial or is blank), and names that appear on more than ten unique purchase transactions in the data set.

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.⁶ 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, Dallas, and Gwinnet County, Georgia, (Figure 1 Panel (a)). iBuyers had roughly 1% market share in Phoenix in 2015; by 2018 this had grown to roughly 6%. Similar striking growth has occurred in Gwinnet (4%), Las Vegas (4%), 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 fewer 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

⁶ Before this final step, 0.8% of homeowners appear to be living in more than one house in a given year.

by their total inventory, is typically between 0.3 and 0.6 per quarter (Panel (e)). iBuyers typically hold inventory for a 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% gross 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)). In addition, iBuyers charge home sellers a service fee that during our sample period was somewhat higher (in the order of 6.5-7.5%) than the typical real estate agent fees of 6% or less associated with a traditional house sale as reported by iBuyers. Hence homeowners selling to iBuyers face, on average, somewhat higher fees compared to traditional home sale. Moreover, as we show below, they sell their homes to iBuyers at a discount relative to prices of similar homes sold through a traditional listing channel.

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 through 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.⁷ 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, iBuyers follow such a model: they purchase houses, hold them in inventory, and sell them, earning a spread. Here, we delve deeper into iBuyers' business model, and generate a set of facts.

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 iBuyer? With the merged transaction-listing dataset, we

⁷ 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.

examine whether houses sold to or by iBuyers were more or less likely to be listed on MLS prior to purchase by estimating the following specification:

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

An observation is at the deeds-records transaction level where i indexes a house in zip code z at quarter t , and each deeds records transaction record may or may not have an associated listing in MIS. $Buyer_is_iBuyer_{izt}$ is an indicator for whether an iBuyer buys the house, and $Listing_{izt}$ 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.⁸ 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 suggests that selling to iBuyers provides a 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.⁹

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, equivalent to equation (1) above:

$$Listing_{izt} = \beta Seller_is_iBuyer_{izt} + H_i' \mathbf{B} + \mu_{zt} + \epsilon_{izt} \quad (2)$$

⁸ Because the match between MLS listings and Corelogic transactions is imperfect, we perform sensitivity analysis around this matching window to confirm that the results hold.

⁹ Finally, we examine whether iBuyers are more likely to purchase houses that correspond to previously “failed” listings: in other words, do homeowners tend to sell to iBuyers as a second chance after their initial listing fails to produce a sale? The results, in Appendix A.5, suggest that this is not the case: Fewer than 0.5% of iBuyer purchases correspond to properties where there was a listing in the previous 365 days that did not correspond to a sale. In contrast, we find that 2% of non-iBuyer purchases follow a previously failed listing. This suggests that iBuyers’ business models do not revolve around attracting properties with previously unsuccessful listings.

An observation is at the deeds-records transaction level where i indexes a house in zip code 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 Column 6). iBuyers do not appear to possess a different technology for selling houses, relying on the standard listing process instead of targeting buyers through a website.

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.¹⁰ 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.¹¹ Flippers are a useful comparison group because they share some similarities with iBuyers: They are absentee owners, who likely purchase houses as an investment.¹² With this in mind, we estimate whether iBuyers listing prices differ from other listers using the following:

$$\log(List\ price_{izt}) = \beta Lister_{izt} + H_i' \mathbf{B} + \mu_{zt} + \epsilon_{izt} \quad (3)$$

Here i indexes a house in zip code 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 2.3% higher than ordinary sellers' listings on comparable properties (Table 2 Panel B Column 1). Flippers also list more aggressively than ordinary buyers, with a markup of 0.8%. In other words, both flippers and iBuyers appear to follow the strategy of listing at a high price.

¹⁰ 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

¹¹ Note that with this definition we can only identify flippers in the MLS dataset (and not in the Corelogic dataset). Thus, results based on MLS will tend to differentiate flippers and other individuals, while results based on Corelogic will necessarily include flippers with other individuals. Based on their selling behavior, this will mildly reduce observed differences between iBuyers and other sellers in the Corelogic data.

¹² 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.

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 5 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 13.6pp more likely to have a “failed” sale once the house is listed (Table 2, Panel B Column 4). 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 when making comparisons above 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 account for “failed” sales and censoring we estimate a Cox proportional hazard model on sales propensity:

$$\lambda(t|X_i) = \lambda_0(t)\exp(\beta Buyer_i + H_i' B) \quad (4)$$

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 36% greater rate than non-iBuyer, (Table 2 Panel B Column 7). 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 2% lower rate.¹³

Finally, using the text of iBuyer listings, we rule out that iBuyers are making significant renovations. Table 2 Panel B Column (2) shows that iBuyers are less likely to advertise making renovations, while flippers are significantly more likely to do so. Appendix A.6 performs a sensitivity analysis around these results by allowing for the possibility that sellers withdraw listings and relist, treating withdrawals and relists in close proximity as a single listing. The results are

¹³ These rates are calculated as $\exp(\text{hazard coefficient})-1$. Columns (8) and (10) add initial listing price to control for listing aggressiveness, and the addition of these controls do not meaningfully alter the results.

qualitatively similar. Additionally, in Appendix A.6, we examine seasonality, which has been documented (e.g., Ngai and Tenreyo 2014) to play an important role in residential real estate transactions. We find that iBuyers tend to post more listings specifically in off-season times, consistent with the above ideas that they can hold inventory at a lower cost and list more strategically compared to regular homeowners.

III.B. Returns for Liquidity Provision

We showed that 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 and decomposes the sources of this return.

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

We now 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,¹⁴ as

$$Gross\ Return_{iztt'}^{Ann} = \left(\frac{Price_{izt'}}{Price_{izt}} \right)^{\left[\frac{1}{t'-t} \right]} - 1 \quad (5)$$

The subscript i denotes a house, z the zip code 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:

$$Gross\ Return_{iztt'}^{Ann} = H_i' B + \mu_{zt} + \epsilon_{iztt'} \quad (6)$$

$Gross\ Return_{iztt'}^{Ann}$ 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

¹⁴ 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.

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, iBuyers' annualized gross return is roughly 6.6% pp higher than those of typical individuals (Table 3 Panel B Column 2). We separate the gross return into a component that is attributable purely to overall house price appreciation and the bid/ask spread—the decomposition is described in detail in Appendix A.8. 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 the overall house price appreciation (Table 3 Panel B Column 6).¹⁵

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, and/or selling them for more than an average household. We investigate the purchase discount and the selling premium using the following hedonic specification:

$$\begin{aligned} \log(\text{Sale Price}_{izt}) &= \beta(\text{Buyer is iBuyer}_{izt}) + H_i' \mathbf{B} + \mu_{zt} + \epsilon_{izt} \\ \log(\text{Sale Price}_{izt}) &= \beta(\text{Seller is iBuyer}_{izt}) + H_i' \mathbf{B} + \mu_{zt} + \epsilon_{izt} \end{aligned} \quad (7)$$

An observation is a house transaction, where i indexes a house in zip code z at quarter t . Sale Price_{it} is the transaction price. $\text{Buyer is iBuyer}_{izt}$ and $\text{Seller is iBuyer}_{izt}$ 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. We compare the price of properties acquired (or sold) 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 Column 1). 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. (Table 4 Panel A Column 4). For the average iBuyer house with a price of approximately \$250,000, this premium amounts to

¹⁵ 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 (Figure 4 Panel (d))

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 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 earlier listing results that iBuyers are unlikely to renovate homes.

As we discussed above, iBuyer price discount also does not seem to reflect potentially lower 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 during our sample period, which is substantially larger than the typical real estate agent fees of 6% or less associated with a traditional house sale.¹⁶

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. These results indicate that like fintech lenders in the mortgage market (see Buchak et al. 2018), iBuyers provide consumers with non-price attributes like convenience rather than simple cost savings.

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

We have documented that iBuyers act as liquidity providers, buying and 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 note that housing markets are quite segmented (Piazzesi, Schneider, and Stroebel 2020) and 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.

¹⁶ See <https://www.opendoor.com/w/pricing> (accessed on January, 2020).

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 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:

$$\log(\text{Sale Price}_{izt}) = H_i' \mathbf{B} + \mu_k + \epsilon_{izt} \quad (8)$$

An observation is a house transaction, where i indexes a house in zip code 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 Panel B). 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, 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.¹⁷

¹⁷ The reliance on easy to assemble hard information also differentiates iBuyers from several other fintech participants. For example, some fintech lenders in the mortgage market rely less on standard hard information than other market participants, when pricing mortgages (Buchak et al, 2018). In brief, fintech mortgage lenders primarily originate government-backed mortgages with pre-arranged fee schedules. They could potentially use additional information about borrowers that is available from various data sources (e.g., cell phone or electrical bills) in their pricing.

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

We saw that iBuyers entry is very selective, both geographically, and, conditional on geography, which types of homes they purchase. We now use iBuyers' revealed preferences to study why intermediation in real estate markets has been so limited. 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, since it requires maintenance and upkeep that iBuyers may not be willing to provide. We look at iBuyer entry to better understand which force dominates.

We first estimate which houses can be easily priced with an algorithm using hard information, i.e., which types of houses have a small pricing error when priced with hedonics. To ensure that iBuyers' pricing decisions do not mechanically affect our classification, we estimate a hedonic pricing model using 2006-2012 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 (8), 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 residuals in absolute value, $|\hat{e}_{izt}|$, from this specification are those that are not priced well by hedonics. We then predict which house characteristics make them difficult to price:

$$|\hat{e}_{izt}| = H_i' \Delta + \zeta_{izt} \quad (9)$$

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.¹⁸ Intuitively differences in land characteristics can result 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 covering 2008-2012 period to estimate a hedonic model to

¹⁸ As a robustness check, we estimate the same analysis for a narrower training sample spanning 2008-2012. The remarkable stability in the coefficient estimates provide confidence that iBuyers using different sets of data would come to largely similar conclusions.

predict whether a given listing sells within 90 days from the listing date, our measure of a liquid market segment:

$$Sells\ Within\ 90\ Days_{izt} = H_i' \mathbf{B} + \epsilon_{izt} \quad (10)$$

$Sells\ Within\ 90\ Days_{izt}$ is a zero-one indicator for whether house i in zip z listed at time t sells within ninety days of its listing. Similar to the results on pricing errors, in unreported tests, we find that cheaper (measured by *previous* sale price), smaller, and single-story houses are more likely to sell quickly. Using the estimates from (9) and (10), we construct for every house its predicted standardized pricing error $|\hat{e}_{izt}|$ 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:

$$iBuyer_{izt} = \beta |\widehat{e}|_{izt} + \gamma \widehat{Sells\ Within\ 90\ Days}_{izt} + \mu_{zt} + \epsilon_{izt} \quad (11)$$

As with the earlier specification, an observation is a house transaction, where i indexes a house in zipcode z at quarter t . $iBuyer_{izt}$ is a zero-one indicator for whether the buyer is an iBuyer. $|\widehat{e}|_{izt}$ is the predicted pricing error, 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} , i.e., we study which houses iBuyers choose to intermediate, conditional on having entered a geographic market.

A house with a 10pp greater absolute predicted pricing error is .79% 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%, (which by 2018 had increased to roughly 5%) 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 0.37%. 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. In Appendix Section A.9, we confirm this is indeed the case.

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 may possess more information and 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 next, these results are 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

We now 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. 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. Homeowners differ in their flow cost of being unmatched, which captures the fact that some have an urgent motivation to move (such as relocating for a job), and others' motives are less urgent (such as downsizing). Critically, we assume that the homeowner's balance sheet is constrained, so owning two homes at the same time is prohibitively costly. This can arise, for example, if households have limited wealth and borrowing against a house comes with LTV constraints imposed by the lender. We approximate costly balance sheets by assuming that households 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, 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.¹⁹

¹⁹ Analogous to, e.g., job seekers and job postings as in Diamond (1982). We believe that our main qualitative insights would be robust to alternative modelling assumptions regarding the search process (e.g., directed search). See Piazzesi, Schneider, and Stroebel (2020) for a recent analysis of implications of various modeling assumptions regarding the search behavior for the housing market equilibrium.

We depart from the standard setting by introducing an intermediary that can provide liquidity in the market: iBuyers. Instead of listing houses and waiting for a buyer, sellers can sell them directly to iBuyers who 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 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 at which they can purchase the house from a willing seller. The reason why intermediaries are valuable to sellers in the first place is because they can execute the transaction without waiting for a buyer, but such transactions may not be instantaneous. Time to transact can depend on the speed at which the intermediary can process the paperwork, potentially inspect the property and other sources of delay. It also could mean cutting out the need for staging before the property can be sold. The second is information. As we document above, the speed at which iBuyers can close transactions puts them at an information disadvantage relative to homeowners, and other potential sellers, who can take time to thoroughly screen a purchase, even if iBuyers can use algorithmic pricing. The third is matching technology. While we see little evidence of differences in matching homeowners with houses, we want to study the potential for intermediation with if iBuyers improved matching technology in the future. We study the model by varying how changes in these characteristics affect the equilibrium outcomes in the market, and the profits from intermediation, to understand the key factors enabling and preventing viable intermediation.

IV.A.1. Market and information structure

Homeowners: At any instant, a homeowner is one of four states, between which she transitions over time (e.g., Anenberg and Bayer (2020)). These states are denoted $\{h, s_p, s_I, b\}$. h denotes a matched homeowner, who is happy with the house in which she currently resides. s_p and s_I denote patient and impatient selling homeowners, respectively: 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 not own a house and is looking for one. The total homeowner population has an exogenous mass $M = 1$, with $\{m_h, m_{s_p}, m_{s_I}, m_b\}$ denoting respectively the endogenous mass of h -types, s_p/s_I -types, and b -types. For notational convenience we denote the total mass of sellers as $m_s = m_{s_p} + m_{s_I}$. Homeowners become unmatched at rate μ , and conditional on becoming unmatched, becomes patient or impatient with probability distribution denoted by $dS(s)$.

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} captures the flow component common across homeowners; $\tilde{\epsilon}_i$ allows for

idiosyncratic differences in homeowners utility flows from their current property representing household-specific preferences over build or location. In effect, this allows for the possibility that sellers may have some market power since houses are horizontally differentiated. When homeowners become unmatched, patient sellers and buyers receive utility flow $\underline{u} < \bar{u}$. Impatient sellers receive $\underline{\underline{u}} < \underline{u}$, which captures their additional disutility from their current unmatched house. This represents the idea that while unmatched homeowners still obtain some utility benefits from occupying the house, the house is no longer a good match. 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. All agents discount the future at rate ρ .

Listings: Selling households can list their house and wait to be randomly matched with potential buyers. Buyers and sellers 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. 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. Households can choose to sell the house to iBuyers instead of listing it. The iBuyer transaction closes in $\tau \geq 0$ days. Closing delays can arise because of the time it takes to gather and process the documentation, but also because of some 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 impatient households are technologically less savvy than others, so they do not transact with iBuyers even at the expense of speed. Alternatively, some patient households simply like the convenience of transacting online, relative to a listing process. We capture these preferences with an idiosyncratic utility shock, ϵ_{ib}^i . $\epsilon_{ib}^i \sim dE^{ib}(\epsilon_{ib}^i)$ is distributed type-1 extreme value distribution with scale parameter σ_{ib} .

Upon purchasing the house, iBuyers sell the house through a listing. We allow iBuyers to 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 flow 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. At rate η , the flow costs become high, m^H .

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 iBuyers face as asymmetric information in “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 $r = R$. Repair costs can represent poorly laid insulation, but also capture features, which lower the price of the house that are difficult to notice directly, such as noise in the neighborhood, school quality or the quality of light at different times of day. 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 (or an implied price discount) is paid by the seller.²⁰ 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:

$$\begin{aligned}\phi(G|r = R) &\equiv \phi_{G|R} = \xi \\ \phi(G|r = 0) &\equiv \phi_{G|\sim R} = 1 - \xi\end{aligned}\tag{M.1}$$

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 purchasing the house, iBuyers find out whether the house will actually require repairs before listing, and pay the required cost. Additionally, we assume that iBuyers do not directly observe whether the seller is patient or impatient (although the iBuyer will rationally consider which types of sellers are more likely to accept its offer).

IV.A.2. Homeowners’ problem

Homeowners choose their actions to maximize their expected utility. Let $\{v_h, v_{sp}, v_{sI}, v_b\}$ denote the value functions of matched households, both patient and impatient sellers, and homeowners who are in the market to purchase a home, 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,

²⁰ This assumption is without loss of generality, but allows for listed houses to be homogenous in quality, increasing model tractability. We could also allow for some degree of informational asymmetry between regular homeowners which would not change our qualitative results as long as iBuyers face on average more informational asymmetry than regular buyers due to their use of fast algorithm-based pricing. See Kurlat and Stroebel (2015) for recent analysis on informational asymmetries in real estate market.

how likely their house is to require repairs, the utility of selling the house on their own or to an iBuyer. Formally, a matched homeowner i has the following value function:

$$(\rho + \mu)v_h^i = \bar{u} + \tilde{\epsilon}_i + \mu \int_{s_*, r, v, \epsilon_{ib}^i} \max\{v_{s_*} - r, \delta(\tau)(p_{ib}^b(v) + v_b) + \epsilon_{ib}^i\} dG(\epsilon_{ib}^i, v, r, s_*) \quad (\text{M.2})$$

Recall, ρ is the subjective discount rate, μ is the unmatched rate, r is the (random variable) repair cost, $\delta(\tau)$ is the time discount factor for closing an iBuyer transaction, $p_{ib}^b(v)$ is the iBuyer offer price conditional on the signal, and ϵ_{ib}^i is the consumer's idiosyncratic preference shock over an iBuyer Transaction. $v_{s_*} - r$ is the household's value by becoming a seller and paying any necessary Repair costs. $\delta(\tau)(p_{ib}^b(v) + v_b) + \epsilon_{ib}^i$ is the discounted value of selling immediately to an iBuyer and becoming a buyer. The integration is over whether the seller is impatient or patient, 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 $dG(\epsilon_{ib}^i, v, r, s_*) = dE^{ib}(\epsilon_{ib}^i)dG(v, r)dS(s_*)$. Given the above decision problem, the probability of selling to an iBuyer given the price offered, and the repairs needed is $\pi_{ib}(p, r) = 1 - E^{ib}[v_s - r - \delta(\tau)(p + v_b)]$.

The value a matched homeowner obtains from the house, 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. For the remainder of the paper, we focus on v_h and $\epsilon_i \sim E(\epsilon_i)$, with $v_h^i = v_h + \frac{\tilde{\epsilon}_i}{\rho + \mu} \equiv v_h + \epsilon_i$. We interpret $\epsilon_i \equiv \frac{\tilde{\epsilon}_i}{\rho + \mu}$ as the capitalized idiosyncratic flow utility from the house, in line with the earlier separation of flow utility into a common component \bar{u} and idiosyncratic component $\tilde{\epsilon}_i$. 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 with an expected utility of v_{s_*} . 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 accepts the offer $\pi(p)$ declines in price. After she sells her house she becomes a buyer, and her utility increases by $v_b - v_s$. Recall that the sellers can be either patient or impatient, so they face a different tradeoff with respect to the speed of sale. Formally, the value functions for patient and impatient sellers if they choose to list the house at price p are:

$$\begin{aligned}\rho v_{s_P} &= \underline{u} + F_s^{hh}(m_s, m_b) \max_{p_p} \pi(p_p)(p_p + v_b - v_{s_P}) \\ \rho v_{s_I} &= \underline{\underline{u}} + F_s^{hh}(m_s, m_b) \max_{p_I} \pi(p_I)(p_I + v_b - v_{s_I})\end{aligned}\tag{M.3}$$

Recall, \underline{u} and $\underline{\underline{u}}$ are the flow utilities for patient and impatient sellers, respectively, m_b and m_s are the masses of buyers and sellers, and p_p and p_I are the endogenous listing prices of patient and impatient sellers. The difference in the sellers' patience results in differential listing prices: intuitively, impatient sellers are willing to set lower listing prices to speed up the sale.

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 list price, accounting for the fact that they may encounter a patient or impatient seller or match with an iBuyer. Let j index the seller type. Upon encountering the seller and seeing the house and the price it is listed at, the buyer's idiosyncratic valuation, ϵ_i , realizes, she pays a viewing cost κ , and she chooses whether to accept or to continue looking. The buyer accepts the offer if her utility from homeownership exceeds that of remaining a buyer by the sale price, if $v_h + \epsilon_i - p_j > v_b$. Thus, the probability of purchase as a function of price can be expressed as $\pi(p) = 1 - E[v_b + p - v_h]$. Her value function is given by:

$$\rho v_b = \underline{u} + \sum_j F_b^j(m_j, m_b) E[\max\{v_h + \epsilon_i - p_j, v_b\} - v_b - \kappa]\tag{M.4}$$

IV.A.3. iBuyers' problem

When a homeowner becomes unmatched and offers to sell the house to iBuyers, iBuyers inspect the house for repair costs. They offer a price $p_{ib}^b(v)$ that depends on their signal of whether repairs are needed. A homeowner, who knows the actual 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 takes possession, realizes the true extent of repair costs, and makes any repairs, if necessary. Let v_{ib}^L denote the value iBuyers place on the house whose flow maintenance costs are low, and v_{ib}^H the value of a house whose flow maintenance costs have become high. The iBuyer's expected profit when making an offer to an unmatched homeowner given signal v is as follows:

$$v_{ib}^{offer}(v) = \max_{p_{ib}} \int_{s_*, r} \pi_{ib}(p_{ib}, r, s_*) (v_{ib}^L - p_{ib} - r) dF(r|v) dS(s_*)\tag{M.5}$$

When setting prices, iBuyers face a similar problem to other sellers holding the quality composition of sellers fixed. Decreasing the price which they offer to purchase houses, increases

the profits per house, $(v_{ib}^L - p_{ib} - r)$ but decreases the probability that they acquire the house $\pi_{ib}(p_{ib}, r, s_*)$. In addition, for iBuyers, the type of sellers who take them up on their offer, $dF(r|v)$, changes, resulting in adverse selection. Intuitively, if iBuyers lower their purchase prices, then the sellers with high quality houses are less likely to sell to them, and to list instead. There are two forces that prevent the complete breakdown of the market despite adverse selection. First, sellers' idiosyncratic value of selling to an iBuyer, such as preferences for technology imply that at least some homeowners with high quality houses are willing to sell to iBuyers even as they lower prices. Moreover, impatient sellers have a higher incentive to sell to iBuyers even when prices decline, dampening the effect of adverse selection.

After buying the house, the iBuyer's decision resembles that of 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:

$$\rho v_{ib}^L = m^L + \eta(v_{ib}^H - v_{ib}^L) + F_s^{ib}(m_s, m_b) \max_{p_{ib}^L} \pi(p_{ib}^L)(p_{ib}^L - v_{ib}^L) \quad (\text{M.6})$$

If iBuyer's maintenance costs increase to m^H , she lowers the price to sell it faster given the increased cost of holding it:

$$\rho v_{ib}^H = m^H + F_s^{ib}(m_s, m_b) \max_{p_{ib}^H} \pi(p_{ib}^H)(p_{ib}^H - v_{ib}^H) \quad (\text{M.7})$$

IV.A.4. Population dynamics

Having described the decision problems of individual participants in the market, we can turn to population dynamics. There are six prices in the market at any point in time: the price at which patient and impatient households list their houses, p_{hh_p} , and p_{hh_I} , the prices at which iBuyers buy high and low quality signal 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 $\mathbf{P} \equiv \{p_{ib}^b(v), p_{hh_p}, p_{hh_I}, p_{ib}^L, p_{ib}^H\}$, we define $\pi_{ib_P}(\mathbf{P})$ as the unconditional probability that a seller, patient or impatient, sells to an iBuyer, given by integrating jointly over repair costs and signals:

$$\pi_{ib}(\mathbf{P}) \equiv \int_{s,v,r} \pi_{ib}(p_{ib}^b(v), r) dG(v, r) dS(s) \quad (\text{M.8})$$

We define these probabilities conditional on being patient or impatient, and denote them as $\pi_{ib_P}(\mathbf{P})$ and $\pi_{ib_I}(\mathbf{P})$. Then, first three equations describe how the population of different types of

households change over time. The population of matched households' changes as a function of the exogenous unmatching rate, and the endogenous rate at which buyers' rematch with new houses:

$$\frac{dm_h}{dt} = -\mu m_h + \sum_{j \in \{hh, ib^L, ib^H\}} F^j(m_j, m_b) \pi(p_j) \quad (\text{M.9})$$

The population of individual sellers is a function of the exogenous unmatching rate, the share of the unmatched population choosing iBuyers $\pi_{ib_p}(\mathbf{P})$, and $\pi_{ib_l}(\mathbf{P})$ for patient and impatient, the probability that unmatched become patient or impatient, P_p or P_l , and the speed at which households can sell their houses to become buyers:

$$\begin{aligned} \frac{dm_{sp}}{dt} &= \mu P_p m_h (1 - \pi_{ib_p}(\mathbf{P})) - F^{hh}(m_{sp}, m_b) \pi(p_{hh_p}) \\ \frac{dm_{sl}}{dt} &= \mu P_l m_h (1 - \pi_{ib_l}(\mathbf{P})) - F^{hh}(m_{sl}, m_b) \pi(p_{hh_l}) \end{aligned} \quad (\text{M.10})$$

The share of household buyers' houses evolves as a function of sale speeds and purchase speeds from both households and iBuyers:

$$\begin{aligned} \frac{dm_b}{dt} &= \mu m_h \sum_{s \in \{P, I\}} P_s \pi_{ib_s}(\mathbf{P}) + \sum_{s \in \{P, I\}} F^{hh}(m_{s_s}, m_b) \pi(p_{hh_s}) \\ &\quad - \sum_{j \in \{hh_p, hh_l, ib^L, ib^H\}} F^j(m_j, m_b) \pi(p_j) \end{aligned} \quad (\text{M.11})$$

Further, the population of iBuyer houses can be split by their maintenance costs:

$$\frac{dm_{ib}^L}{dt} = \mu m_h \sum_{s \in \{P, I\}} P_s \pi_{ib_s}(\mathbf{P}) - \eta m_{ib}^L - F^{ib}(m_{ib}^L, m_b) \pi(p_{ib}^L) \quad (\text{M.12})$$

$$\frac{dm_{ib}^H}{dt} = \eta m_{ib}^L - F^{ib}(m_{ib}^H, m_b) \pi(p_{ib}^H) \quad (\text{M.13})$$

To close the model, since the housing stock is fixed, for every house on the market, there is exactly one potential buyer:

$$m_s + m_{ib}^L + m_{ib}^H = m_b \quad (\text{M.14})$$

IV.A.5. Equilibrium

We look for a stationary equilibrium. The equilibrium is a set of prices \mathbf{P} such that:

- 1) iBuyers maximize profits when setting purchase prices (M.5) and listing prices (M.6, M.7),
- 2) Households maximize utility when purchasing houses (M.2, M.4) and when listing houses and selling them to iBuyers, (M.2,M.3).
- 3) Stationarity: State variables $\{m_h, m_s, m_b, m_{ib}^L, m_{ib}^H\}$ are constant as determined by their laws of motions, $\frac{dm_h}{dt} = \frac{dm_{sP}}{dt} = \frac{dm_{sI}}{dt} = \frac{dm_b}{dt} = \frac{dm_{ib}^L}{dt} = \frac{dm_{ib}^H}{dt} = 0$ (M.9—M.14),
- 4) Beliefs: iBuyers beliefs about housing repairs are consistent as described above (M.1).

IV.B Model Calibration

We calibrate several parameters externally in relation to existing literature. We then calibrate the remaining parameters by matching model-implied moments to moments we observe in the data. Table 7 Panel A describes these moments and our model’s fit.

IV.B.1 Externally calibrated parameters

We follow Anenberg and Bayer (2020) and set the discount rate ρ to 0.05. The census 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.²¹ Agents in our model move after becoming unmatched, and consequently this number corresponds to the unmatching rate μ in our model. We set the probability that a house needs serious repairs, ϕ_R , to the fraction of listings mentioning renovation, which is 0.109 in the MLS data. We define half of the households to be impatient and half to be patient.²² Finally, we assume for the baseline analysis that iBuyers’ time to close, τ , is 15 days—we explore how changing this parameter affects iBuyers’ ability to intermediate in Section VI. Finally, as a normalization, we set the meeting cost κ to 0.577, which is the mean of a type-one extreme value distribution, to prevent households from searching for a better matched house for too long.

IV.B.2 Parameters calibrated to the data: Identification

We calibrate the remaining 11 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 quantitatively. Discussing the identification of the remaining parameters also presents an opportunity to provide exposition on the economics underlying the model.

²¹ <https://www.census.gov/topics/population/migration/guidance/calculating-migration-expectancy.html>

²² In Appendix A.11, we examine whether our counterfactual conclusions are sensitive to this assumption by estimating the model with an assumption of 15% and 25% impatient sellers. The results are quantitatively and qualitatively similar.

IV.B.2.1 Household preferences and shocks

Match utilities: Since only relative utilities across potential outcomes are relevant for household choices, the level of utility is a normalization, and we normalize \bar{u} , the utility flow for a matched homeowner as a baseline.²³ The gap between the utilities of housing between matched and unmatched households, \bar{u} , \underline{u} , and $\underline{\underline{u}}$ are reflected in listing prices, differences in listing prices between patient and impatient sellers, and iBuyer discounts. A buyer is willing to pay more for a house if rematching is more beneficial—when this gap is larger. In response, sellers will list at a higher price. By selling directly to an iBuyer, sellers become rematched sooner and obtain a utility increase faster. When rematching is more valuable, the iBuyer purchase discount grows.

To identify the different flow utilities for patient and impatient sellers, we exploit the fact that impatience changes the preference between price and selling speed. Lower listing prices result in faster sales and impatient households are therefore willing to list at lower prices. We identify the utility differences in the data by imposing that the bottom 50% of list prices, controlling for observables, are made by impatient sellers.²⁴ We then examine differences in median list prices for these groups to aid in the identification of $\underline{\underline{u}}$: a lower $\underline{\underline{u}}$ generates a larger gap in list prices between patient and impatient sellers. Observe that listing in the presence of iBuyers is an equilibrium outcome: to make sure the model and data are consistent, we match listing prices from the data (households which have endogenously not sold to iBuyers) to the non-iBuyer listing prices in the model (also households which have endogenously not sold to iBuyers).

Variance of preference: Households' preferences over houses and selling to iBuyers differ, with σ_m and σ_i capturing the extent of these differences across households. If differences are small, then the average utility difference across choices should predict actions well. Intuitively, iBuyers purchase houses at discounts, so they are a “bad deal” for the average seller. Therefore, dispersion in preferences increases sellers' likelihood of selling 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 imply that buyers view houses as differentiated, thereby increasing sellers' market power and in turn 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

²³ We normalize $\bar{u} = 24$ for numerical convenience; evaluation of the value function requires exponentiating the difference the other flow utilities, and this normalization keeps them reasonably close to zero.

²⁴ In Appendix A.10, we examine the robustness of our key outcomes to the 50% assumption by re-estimating the model and rerunning the counterfactuals assuming that 25% and 15% of the sellers are impatient. Both qualitative and quantitative results are largely unaffected by different assumptions.

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.

IV.B.2.2 Matching and Intermediation

Match rates: A popular explanation of iBuyers' advantage is their matching technology. Our calibration, on the other hand, suggests that iBuyers possess a slightly worse matching technology than other sellers, with a match rate that is 94% as high as other sellers, $\lambda_{ib} = 0.94$. The match intensity, λ , is related to how long a house stays on the market. There are two main forces that determine duration in equilibrium. With a higher λ , buyers can reject sellers and see more offers per unit time, which reduces seller markups and tends to reduce times on the 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 with approximately equal list prices in 2018. Our model suggests that with better matching technology, for a given price iBuyers should sell more quickly, which we do not observe in the data, implying a worse matching technology.

iBuyer holding costs: iBuyer market share and selling behavior helps to identify their cost of holding a home relative to the cost of unmatched sellers. When iBuyer holding cost is low relative to sellers their ability to intermediate increases. They are able to offer higher prices and sell at lower prices, leading to a higher market share. On the other hand, when their holding cost is relatively high, sellers are better off listing their homes themselves. Additionally, a lower iBuyer holding cost means iBuyer can be more patient in selling, leading to longer listing times. Our estimates put iBuyer holding costs roughly between those of patient and impatient sellers: Their annual net flow utility from holding the house is roughly \$1,400 higher than impatient sellers, which gives iBuyers an advantage in intermediating in this segment in addition to balance sheet capacity. On the other hand, their annual holding costs are roughly \$8,500 lower than patient sellers who still obtain some utility from living in their current house. We interpret this holding cost differential of roughly \$700 per month as the approximate rental value for a "second-best" renter as compared to empty iBuyer-owned home. This large differential gives iBuyers a significant disadvantage among sellers with less pressure to move quickly.

Because iBuyers houses are vacant, they face the probability η that maintenance costs increase by $m^H - m^L$, if issues are undetected for a while. When iBuyers experience a cost increase, their urgency to sell the house grows, and they lower their listing price. We match the probability that an iBuyer reduces prices before successfully finding a seller to match η . The cost increase $m^H -$

m^L to a first order governs the size of the price reduction. Additionally, these parameters broadly impact iBuyer market share, through the fact that a higher c or η increase expected iBuyer costs and therefore their market share.

These parameters also influence the relationship between iBuyer market share and the liquidity of the house. The adverse cost shock is more likely to realize the longer the iBuyer possesses the house. Thus, when selling a particular house is expected to take longer for any seller, iBuyers are at a comparative disadvantage relative to homeowners who live in the house and are thus protected from increases in maintenance costs. Therefore, our model predicts that iBuyers' market share declines in market liquidity, which we find in the data. Quantitatively, we measure the predicted derivative, $d(\text{iBuyer share})/d(P(\text{sells in 90 days}))$, and use it as another moment we match in the data to discipline the calibration.

iBuyers Information Disadvantage and Adverse Selection: iBuyers are subject to asymmetric information, which results in adverse selection. Houses may need “repairs” costing R , and iBuyers receive signals regarding whether they need repairs with false positive and true negative probabilities ξ . These repair costs are a stand-in for house features that are difficult to capture in a remote pricing algorithm—true repair costs, but also whether the house has a bad odor, or whether the neighbors are noisy. Homeowners can observe whether the house needs repairs but iBuyers cannot. iBuyers offer a single pooling price conditional on a noisy signal of house quality, which results in adverse selection. On average, homeowners with houses needing repairs attempt to pool with homeowners with houses not needing repairs. This incentivizes homeowners not needing repairs to list themselves, while homeowners needing repairs will attempt to sell through iBuyers, in a classic lemons problem. The adverse selection problem is alleviated though two forces, which prevent full market breakdown. Some sellers' have a high idiosyncratic preference for selling to iBuyers, for example because they like transacting through low hassle technology. Second, impatient sellers have a higher incentive to sell to iBuyers who can close a transaction quickly. Both forces imply that at some homeowners with high quality houses are willing to sell to iBuyers even as they lower prices, dampening the effect of adverse selection.

Within the model, repair costs are a source of unexplained house price variation and adverse selection, and we use both of these ideas to calibrate the values of R and ξ . In Section III we show that iBuyers' pricing technology can be well approximated with a hedonic regression. We interpret the residual mean square error of this regression as unexplained price variation due to unobserved repair cost R (or the cost of improving the house quality to sell it). The noisier the iBuyer's signal is, the more severe the adverse selection becomes, decreasing its ability to intermediate. That is, 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, $d(\text{iBuyer share})/d(\text{Mean pricing error})$. We map this

moment to its empirical analog obtained 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. We estimate that the repair cost R is \$1,490, consistent with the idea that iBuyers are able to detect or contract around the most severe issues with the house but are exposed to smaller, harder-to-quantify errors in valuation. We estimate a low but positive iBuyer valuation noise, ξ , at 0.003. The relatively low cost and valuation noise jointly suggest that iBuyer technology is remarkably effective in valuing homes.²⁵ However, consistent with our data, the model shows that these relatively small errors have large consequences for iBuyers' ability to intermediate houses, given the narrow margins they operate with.

IV.C The Equilibrium Consequences of iBuyer Entry (External Validation)

We calibrate the model to post iBuyer entry period. Because our data also contain the period prior to iBuyer entry, we can examine the predictions of the calibrated model with and without iBuyers. Because pre-entry quantities were not used in calibrating the model, we use this exercise to further validate our models' ability to fit data.

Our model predicts that the average time on market decreases from 91 to 87 days once iBuyers enter. Impatient households are the predominant sellers to iBuyers and gain substantially in terms of times to sale when iBuyers enter. Even prior to iBuyer entry, they experience shorter times to sale, because they choose to sell their houses quickly at lower prices. Once iBuyers enter, roughly 10% of impatient households choose to sell to iBuyers instead of listing (Table 8). Patient households choose to list their houses exclusively. The average time to sale declines significantly, by 8 days for impatient households, though even patient households, face slightly shorter times to sale in the new equilibrium. We also predict that equilibrium house prices rise slightly after iBuyer entry, by 0.8% (Figure 8 Panels (a)). Intuitively, better intermediation increases the consumption value that households derive from a house, increasing the present value of the utility flow.

To validate the model, we compare elasticities generated across markets after iBuyers enter using a difference in differences style analysis to those from the model. We describe the details of the empirical exercise in Appendix A.11. Briefly, we need to account for several features to compare the model implied estimates to the data. First, as we describe in Section II.C, the extent of iBuyer entry differs across markets depending on market characteristics. On the one hand, this is an impediment, since it is difficult to choose the average market. On the other hand, this allows us to exploit differences across markets in the extent of iBuyer entry and use this variation to identify the equilibrium impact of iBuyers in a difference in differences style exercise. Second, iBuyer

²⁵ Note that as we discussed in Section III iBuyers focus on market segments with relatively easy-to-price homes. The valuation errors could be much higher in markets where iBuyers do not endogenously enter.

entry may be endogenous to time on the market and prices. We therefore also instrument for iBuyer entry using the physical characteristics of the housing stock transacting before iBuyer entry. Third, we measure time on the market using Redfin data at the zip code level, which report the fraction of listed homes in a zip code that sell within two weeks of listing. Last, we must transform model implied quantities to be comparable to the coefficients from the empirical estimation.

Consistent with the model's predictions, results in Appendix A.11 show that we find a positive elasticity of price to iBuyer market share, typically close to the model's prediction within standard error bounds, and a similarly positive relationship for fraction of houses selling within two weeks of listing. For example, to measure changes in (log) house prices in the data after iBuyer entry, we exploit repeated (non-iBuyer) sales of the same property in zip codes within the differences in differences setting. The regression coefficient is a semi-elasticity of prices to iBuyer share, with an instrumental variable coefficient of 1.4, and a standard error of 0.27. The model equivalent semi-elasticity given the iBuyer share of 4.88% and a change in prices of 0.8%, is $0.8\%/4.88\%=1.6$. Assuming an exponential distribution of sale times our model predicts that the probability of selling in two weeks increases by 0.14% after iBuyer entry, compared to the positive but somewhat larger increase of 1.57% in the reduced form analysis. Broadly, our model makes predictions that are qualitatively and quantitatively consistent with this reduced form exercise.

The second part of Appendix Section A.11 performs a more qualitative model validation exercise. Our model predicts that impatient sellers are those that sell to iBuyers and who benefit most from iBuyer entry though swifter closing times. 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.

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

As our model illustrates, there are three characteristics of financial intermediaries which affect their ability to intermediate in this market. The first force is speed: A listing homeowner has to go through the lengthy listing process of matching with a buyer, whereas selling to the dealer-iBuyer takes place in τ days it take for an iBuyer to formally close the transaction. The reason why intermediaries are valuable to sellers in the first place is because they can execute the transaction without waiting for a buyer, with τ measuring the speed at which iBuyers can transact. The second force is occupancy: A listing homeowner (typically) remains in her house during the listing process, allowing her to derive some utility and perform routine maintenance. In contrast, a dealer-iBuyer leaves the house unoccupied, foregoing the utility and potentially failing to perform routine maintenance. Our model reflects this through differences in flow utilities, \underline{u} , and $\underline{\underline{u}}$ 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 iBuyers, partially due to the speed at which the latter transact, and who therefore value the home remotely. Our model reflects this in two ways: the precision of the signal, specialized in our calibration as ξ , and the “repair” cost for a mis-valued home, R , i.e., the severity of the valuation error conditional on an error.

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 the ability of iBuyers to intermediate in the housing market. 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.

V.A. The Economics of Intermediating Houses: Sources of iBuyer Technological Advantage

V.A.1. Speed

We first investigate the quantitative importance of iBuyers’ transaction speeds by varying the parameter governing time to close, τ . The baseline value is $\tau = 15$ 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 that speed is critical to intermediation in this market. The ability to hold more than one house on the balance sheet does not confer a significant advantage if iBuyers then take a long time to accomplish the transaction. The market share of iBuyers shrinks to 0%. Without their speed advantage, iBuyers offer essentially nothing of benefit to households that they could not affect better themselves.²⁶

V.A.2. Occupancy: Intermediating Real Assets

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 intermediation by 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

²⁶ An alternate specification of the model could use differences in risk aversion to motivate iBuyer activity. This assumption would generate qualitatively similar results. For instance, a slow iBuyer would offer little advantage to a risk-averse homeowner who would still be exposed to house price fluctuations while the deal closed.

they engage in routine maintenance, which protects from potential adverse events such as a leak. 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 in order to produce the utility flow of an unmatched but patient homeowner, \underline{u} , and set the rate of maintenance cost increases η to zero. We chose the unmatched, patient homeowner as the benchmark occupant with the intuition that a potential short-term renter would not be fully matched (as compared to a long-term homeowner) but also willing to rent short term and not rush to a new house (as compared to an impatient seller).

We find that iBuyer market share rises to roughly 14% (Figure 9 Panel (a)). This large increase quantifies the intuition that a significant obstacle to intermediating houses is that iBuyers do not get to live in the house during the sale period. This is an economically inefficient use of resources that are already in place. This counterfactual also suggests that there are large incentives for iBuyers to rent houses during the sale period. Doing so, would significantly increase their ability to intermediate in this market.

This feature differentiates the balance sheet intermediation of real assets from intermediation of financial assets. When a dealer holds intermediates a financial asset, they collect the dividends from the asset. iBuyers, on the other hand, do not collect the “house dividend” which decreases the economic value of balance sheet intermediation. It may explain the large extent of balance sheet intermediation of financial asset and commodities relative to balance sheet intermediation of real assets where intermediaries such as brokers try to improve matching without holding the asset.

V.A.3. Information

Finally, we investigate the role of information asymmetry in limiting residential real estate intermediation. As we show above, iBuyers ability to swiftly execute transactions is critical for their ability to intermediate. The tradeoff is that swift transactions require a swift and remote valuation, which leaves iBuyers at an information disadvantage relative to other parties in the market. To quantify the extent of adverse selection, we compute counterfactuals as iBuyers information quality deteriorates from perfectly informative to completely uninformative. Formally, we increase the probability that a home is needs repairs conditional on a good signal from 0 to 0.5, as captured by parameter ξ . The counterfactual illustrates the extent of adverse selection as valuation quality declines (Figure 10).

Because iBuyers screening technology is imperfect, they sometimes purchase houses requiring repair even conditional on their assessment of the house being in good condition (good signal). As valuation noise increases, iBuyers are forced to offer steeper discounts in purchases prices when receiving the good signal, because there are more houses requiring repairs in the pool. Figure 10,

Panel B shows that discounts increase by over 15% as the quality decreases. These lower prices are unappealing to homeowners whose houses are in good condition and do not require repairs. They are therefore less willing to sell houses to iBuyers, decreasing the quality of the pool iBuyers face. In other words, the share of good types willing to sell their house to iBuyers declines steeply, by over one third (Panel c). Figure 8, Panel (d) shows that adverse selection increases as information quality deteriorates: the fraction of iBuyer purchases that need repairs conditional on receiving a good signal rises steeply, by over 40%.

This counterfactual suggests that adverse selection limits real estate intermediation overall. iBuyers ability to intermediate is substantially smaller as pricing errors decrease, shrinking their market share (Panel a). The results are also consistent with the empirical pattern that iBuyer penetration is lowest in markets, in which iBuyer's have large valuation errors, which we as we document in Section III.F. In these markets, iBuyers are exposed to more adverse selection, so their ability to intermediate is lowest.

V.B. Fast and Accurate: The essential role of technology in intermediation of houses

Balance sheet intermediation in real estate market was rare prior to the iBuyer entry. We show that balance sheet capacity, speed of transaction, and valuation accuracy are all important determinants of their ability to intermediate houses. All of these dimensions of intermediation could have been provided by intermediaries without a technological advantage. Balance sheet capacity does not require technology per se. Nor is technology essential for house valuation or the ability to transact quickly. Here, we use the model to illustrate that iBuyers' ability to intermediate is contingent on doing all three dimensions (somewhat) effectively. In other words, to intermediate in the housing market, and intermediary has to be able to both provide speed of transaction and, at the same time, be able to price houses fairly accurately. iBuyers' technology allows them to be "fast and sufficiently accurate."

To illustrate that point in Figure 9 Panel (b), we consider an intermediary with a balance sheet capacity to purchase houses, but unsophisticated valuation technology. Such an intermediary can either choose to be "slow and accurate" or "fast and inaccurate." We consider a "slow and accurate" intermediary, which is as accurate in valuation as iBuyer, but achieves that though slower transacting, extending closing time by 10 days. Intuitively, this intermediary chooses to delay closing in order to obtain a more accurate valuation. Our model suggest that such intermediaries would achieve very small market shares, below 1%. In essence, extending closing times by 10 days drives away most of the appeal of iBuyers for their customers, impatient sellers. We also compare iBuyer to a "fast and inaccurate" intermediary. It keeps the short closing time of iBuyers, but then suffers a loss of information, with $\xi = 0.25$, because of its unsophisticated valuation

technology. This intermediary's share is below 2% market share, in contrast to iBuyers' share of 5%. These findings indicate that the ability to be both fast and sufficiently accurate at the same time is the key elements of technological advantage of iBuyers over other intermediaries, and why intermediation was so rare prior to the iBuyer entry.

V.C. Liquidity Provision in Somewhat Liquid Markets: The Limits of iBuyer Liquidity Provision

Finally, we examine the extent to which, and in what contexts, iBuyers are able to provide liquidity to households. A naïve prediction might be that the emergence of iBuyers is most likely to in markets with most illiquid houses, in which households have the highest willingness to pay for an intermediary who can directly buy houses. However, the data show that iBuyers focus on the most liquid markets. To understand the economic mechanism behind this finding, we study markets with different underlying liquidity of housing, by varying the parameter governing the matching rate, λ , and see how the underlying liquidity affects the provision of iBuyer intermediation. To interpret λ more directly, we compute the transaction speed in each market prior to iBuyer entry.

Our model replicates the finding empirical pattern that greater ex-ante liquidity is associated with greater iBuyer market share (Figure 11, Panel (a)). Our model provides two related mechanisms that lead to the positive association between ex-ante liquidity and iBuyer market share. First, even without information frictions, expected holding costs are larger in less liquid markets, making liquidity provision more difficult. iBuyers maintain empty houses and when maintaining an empty house, there is the possibility that an unobserved negative shock (e.g., a leaking roof) occurs which significantly increases maintenance costs. In 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 in markets reduces the iBuyers profits thereby limiting their ability to intermediate.

Second, this lower profitability exacerbates adverse selection. Lower profitability due to holding costs forces iBuyers to offer less favorable purchase prices. Given these low prices, good types would rather list houses themselves, and thus only bad types sell their houses to iBuyers. This endogenous reaction forces iBuyers to reduce prices even further, leading to a partial unraveling of the market. This force is shown clearly in Figure 11 Panel (c): the fraction of iBuyer purchases from bad types increases as liquidity in the market declines. Additionally, Panel (d) shows directly how iBuyers' ability to avoid adverse selection deteriorates with market liquidity. The fraction of iBuyer purchases need repairs conditional on the iBuyer receiving a signal that the house does not need repairs increase as liquidity deteriorates. Overall, these two forces imply that iBuyers have a larger presence in more liquid markets.

The larger presence of iBuyers in more liquid markets also implies that these are the markets in which iBuyer entry leads to the largest liquidity gains. In Figure 11, Panel (b). that iBuyers have

much more pronounced relative effect on the housing turnover (time-to-rematch) of already liquid homes. Thus, iBuyer benefits are smallest in the markets that need liquidity provision most.

iBuyers have difficulty competing in illiquid markets, they also struggle to compete in extremely liquid markets where houses though listings sell far faster than the typical 90 days. We simulate such hot markets by dramatically increasing the liquidity in a market by increasing λ in order to generate a typical sale time of roughly 30 days even without the presence of iBuyers. Figure 12 shows that iBuyer market share in the “hot” market drops to close to zero. Too much liquidity also exacerbates adverse selection: Panel (b) shows that that in hot markets, roughly 80% of iBuyer purchases that receive a “good” signal nevertheless are of low quality. Intuitively, when houses trade in a market extremely quickly, the 15-day transaction time that iBuyers offer is not sufficiently fast relative to an ordinary listed sale. Because iBuyers need to buy at a discount, good types are for the most part no longer willing to sell to them. The only sellers remaining in the market are bad types. Even among these bad types, iBuyer market share is extremely low because the marginal benefit of selling quickly to the intermediary (at a discount) relative to a (sped up) listing becomes very small.

Taken together our findings highlight why liquidity provision in real estate markets has been limited, despite the high demand for liquidity by households. 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 holding costs. Additionally, when markets are highly liquid (i.e., without iBuyers), the market for iBuyer intermediation collapses entirely. 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.

Model Performance out of Sample: COVID-19 Epidemic and the Exit of Zillow Offers

We conclude this section by discussing two recent out of sample events through the lenses of our model. First, is the COVID-19 pandemic, which is not included in our empirical analysis. During the early stage of the COVID-19 pandemic, most iBuyers temporarily suspended their operations. This may seem surprising at first. The extensive social distancing measures dramatically decreased transactions in the traditional listing market. On first instinct, this situation may be a boon for iBuyers whose business model limits physical contact with potential seller, who do not have to list their homes. Instead, as our analysis illustrates, a because social distancing reduces overall market

liquidity, this considerably limits the ability of iBuyers to intermediate. As transaction speed increased after the early part of the pandemic, iBuyers resumed their activity.

Second, our analysis *predates* a period when one of the main iBuyers, Zillow, decided to exit the iBuying market in November 2021 due to the losses incurred on the house transactions. This may also seem surprising because this was a period of liquid house markets. Moreover, house prices were trending upward in most markets during 2021, so a balance sheet of owning homes would naturally generate a profit. We do not want to speculate on the precise reason for why Zillow chose to pursue a strategy of expanding more aggressively than other iBuyers that lead to losses. Instead, we argue that an expansion of activity during the very hot and unpredictable housing market of 2021 was expected to be unprofitable for iBuyers from the perspective of our model, because the conditions were ripe for extensive adverse selection. First, due to increased demand for homes in 2021 and limited supply, the traditional selling channel became significantly faster in 2021. According to Zillow data,²⁷ in June 2021 roughly 67% of US metro areas saw the median days-to-pending fall below 30, in contrast to only 5% of metro areas in June 2020. Appendix A.12 shows these statistics between 2018 and 2022.

As we explained in Section V.C, when houses trade in a market sufficiently quickly, iBuyers lose most of their comparative speed advantage. The most sellers remaining in the iBuying market are the ones exploiting iBuyer valuation errors, resulting in adverse selection. Our counterfactual suggests that iBuyers need to shrink market shares to avoid losses, something Zillow did not do. Second, in late 2021 many market observers became increasingly concerned regarding the persistence of pandemic area housing boom and growing uncertainty regarding the future state of the housing market. Again, our framework shows that challenges in accurately pricing homes during rapidly changing and uncertain economic environment lead to increased adverse selection, which should result in iBuyers shrinking market shares. Consistent with this view, unlike Zillow, other iBuyers appear to reduce their acquisitions during late 2021.²⁸ Through the lens of our model, the conditions in late 2021 were not favorable for iBuyers, so the continued expansion of Zillow's iBuying activity resulted in losses, and ultimately in its exit.

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

²⁷ <https://www.zillow.com/research/data/>

²⁸ <https://www.latimes.com/business/story/2021-11-03/ibuyers-zillow-opendoor-home-sales-southern-california-housing-institutional-investors>

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.²⁹ On the other hand, the economic downturn could increase the share of homeowners that value the convenience of a quick sale, making liquidity provision by iBuyers

²⁹ 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.

more valuable. We leave the analysis of viability of iBuyer business model across various economic environments for future research.

Third, we note that there could be other alternatives to iBuying that could serve similar purpose. For example, intermediaries could provide households with bridge financing during the period of home sale and subsequent acquisition of new home. However, we note that such alternatives are not a simple substitute for iBuying. If households are financially constrained (due to adverse selection or moral hazard concerns), providers of bridge financing may not be able to provide such borrowers with more debt financing compared to existing traditional lenders. We leave analysis of other forms of liquidity provision in the housing market for future research.

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. Additionally, because cars are mobile, limited liquidity in local markets, which can hamper iBuyers, is not likely to be a large barrier for auto intermediation.

References

- Anenberg, E., 2016, Information frictions and housing market dynamics, *International Economic Review* 57, 1449-1479.
- Anenberg, E., and P. Bayer, 2020, Endogenous Sources of Volatility in Housing Markets: The Joint Buyer-Seller Problem, *International Economic Review* 61, 1195-1228.
- Andersen, S., C. Badarinza, L. Liu, J. Marx, T. Ramadorai, 2020, Reference Dependence in the Housing Market, working paper.
- Agarwal, S. and S.H. Ang, Y. Deng, Y. Wang, 2021, Mortgage Brokers and the Effectiveness of Regulatory Oversight, *Management Science* 67, 5278-5300.
- Barwick, P. J., and P. A. Pathak, 2015, The Costs of Free Entry: An Empirical Study of Real Estate Agents in Greater Boston, *RAND Journal of Economics* 46, 103–145.
- Barwick, P. J., P. A. Pathak, and M. Wong, 2017, Conflicts of Interest and Steering in Residential Brokerage, *American Economic Journal: Applied Economics* 9, 191–222.
- Benetton, M., 2021, Leverage Regulation and Market Structure: A Structural Model of the UK Housing Market, *Journal of Finance* 76, 2997-3053.

- Beraja, M., Fuster, A., Hurst, E., Vavra, J., 2019. Regional Heterogeneity and Monetary Policy. *Quarterly Journal of Economics* 134, 109–183.
- Berger, D., Guerrieri, V., Lorenzoni, G., Vavra, J., 2017. House Prices and Consumer Spending. *Review of Economic Studies* 85, 1502–1542.
- 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, 2020, 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.
- Chang, B., 2018, Adverse Selection and Liquidity Distortion, *Review of Economic Studies* 85, 275-306.
- Chinco, A., and C. Mayer, 2016. Misinformed Speculators and Mispricing in the Housing Market. *Review of Financial Studies*, 29, 486-522.
- DeFusco, A. C. Nathanson, and E. Zwick, 2017, Speculative Dynamics of Prices and Volume, working paper.
- Diamond, P, 1982, Wage Determination and Efficiency in Search Equilibrium, *The Review of Economic Studies* 49, no. 2 (1982): 217-227.
- Duffie, D., N. Garleanu, and L. H. Pedersen, 2005, Over-the-Counter Markets, *Econometrica* 73, 1815–1847.
- Eisfeldt, A., 2014, Endogenous liquidity in asset markets, *Journal of Finance* 59, 1-30.
- Favilukis, J., Ludvigson, S., Van Nieuwerburgh, S., 2017. The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium, *Journal of Political Economy* 125, 140–223.
- Fu Y., W. Qian, B. Yeung, 2015, Speculative investors and transaction tax: Evidence from the housing market, *Management Science*, 62, 3254-327.
- Fuster, A., M. Plosser, P. Schnabl, and J. Vickery, 2019, The Role of Technology in Mortgage Lending, *Review of Financial Studies* 32, 1854–1899.
- Ganong, P., Noel, P., 2020, Liquidity versus Wealth in Household Debt Obligations: Evidence from Housing Policy in the Great Recession, *American Economic Review* 110, 3100–3138.

- Gavazza, A., A. Lizzeri, and N. Roketskiy, 2014, A Quantitative Analysis of the Used-Car Market, *American Economic Review* 104, 3668-3700.
- Gavazza, A., 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.
- Genesove, D., and L. Han, 2012, Search and Matching in the Housing Market, *Journal of Urban Economics* 72, 31-45.
- Gilbukh., S, and P. Goldsmith-Pinkam 2019, Heterogeneous Real Estate Agents and the Housing Market, working paper.
- Gorback, C.S., and B. J. Keys, 2020, Global Capital and Local Assets: House Prices, Quantities, and Elasticities, working paper.
- Grossman, S. J., and M. H. Miller, 1988, Liquidity and Market Structure, *Journal of Finance* 43, 617–633.
- Greenwald, D., 2018. The Mortgage Credit Channel of Macroeconomic Transmission. Working paper. MIT Sloan, Cambridge, Massachusetts.
- Guren, A., 2018, House Price Momentum and Strategic Complementarity, *Journal of Political Economy* 126, 1172-1218.
- Guren, A., and T. J. McQuade, 2020, How Foreclosures Exacerbate Housing Downturns, *Review of Economic Studies* 87, 1331-1364.
- Guerrieri, V., R. Shimer, and R. Wright, 2010, Adverse Selection in Competitive Search Equilibrium, *Econometrica* 78, 1823-862.
- Guerrieri V., and R. Shimer, 2014, Dynamic Adverse Selection: A Theory of Illiquidity, Fire Sales, and Flight to Quality, *American Economic Review* 104, 1875-1908.
- Guerrieri V., and R. Shimer, 2018, Markets with Multidimensional Private Information, *American Economic Journal: Microeconomics* 85, 1502-1542.
- Gupta, A., and C. Hansman, 2022, Selection, Leverage, and Default in the Mortgage Market, *Review of Financial Studies* 35, 720-770.
- Head, A., H. Lloyd-Ellis, and H. Sun, 2014, Search, Liquidity, and the Dynamics of House Prices and Construction, *American Economic Review* 104, 1172–1210.
- Hendel I., and A. Lizzeri, 1999, Adverse Selection in Durable Goods Markets, *American Economic Review* 89, 1097-1115.

- Hendel, I., A. Nevo, and F. Ortalo-Magné, 2009, The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison.com, *American Economic Review* 99, 1878–1898.
- Indarte, S., 2021, Moral Hazard versus Liquidity in Household Bankruptcy, working paper.
- Jiang, E., G. Matvos, T. Piskorski, A. Seru, 2020, Banking without Deposits: Evidence from Shadow Bank Call Reports, Working Paper.
- Kaplan, G., Mitman, K., Violante, G., 2020. The Housing Boom and Bust: Model Meets Evidence. *Journal of Political Economy* 128, 3285-3345.
- Kurlat, P., and J. Stroebe, 2015, Testing for information asymmetries in real estate markets, *Review of Financial Studies*, 28, 2429-2461.
- Lagos, R., G. Rochetau, and P-O. Weil, 2011, Crises and Liquidity in Over-the-Counter Markets, *Journal of Economic Theory* 146, 2169-2205.
- Landvoigt, T., M. Piazzesi, and M. Schneider, 2015, The Housing Market(s) of San Diego, *American Economic Review* 105, 1371-1407.
- 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.
- Lester, B., A. Shourideh, V. Venkateswaran, and A. Zetlin-Jones, 2019, Screening and Adverse Selection in Frictional Markets, *Journal of Political Economy* 127, 338-377.
- Mian, A., and Sufi, A., 2022, Credit Supply and Housing Speculation, *Review of Financial Studies* 35, 680-719.
- Nathanson, C, and E. Zwick, 2018, Arrested Development: Theory and Evidence of Supply-Side Speculation in the Housing Market, *Journal of Finance* 73, 2587-2633.
- 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. Stroebe, 2020, Segmented Housing Search, *American Economic Review* 110, 720-59.
- Rekkas, M, R. Wright, and Y Zhu, 2020, How Well Does Search Theory Explain Housing Prices?, working paper.
- Rubinstein, A., and A. Wolinsky, 1987, Middlemen, *Quarterly Journal of Economics* 102, 581-594.

Stroebel, J., 2016, Asymmetric Information about Collateral Values, *Journal of Finance*, 71, 1071-1112.

Weill, P-O., 2007, Leaning Against the Wind, *Review of Economic Studies* 74, 1329–1354.

Weill, P-O., 2020, The Search Theory of OTC Markets, *Annual Review of Economics* 12, 747-773.

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

Wong, A., 2021, Refinancing and the Transmission of Monetary Policy to Consumption, working paper.

Table 1: Summary Statistics

This table shows summary statistics for the main datasets used in the paper: the transaction deeds records between 2013 and 2018 from Corelogic (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 in years. *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-flipper 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	1,384,235	319,406	215,204	109,000	185,000	260,000	379,000	750,000
iBuyer	2,158	240,151	67,019	165,000	197,000	228,000	267,000	364,150
Flipper	106,714	296,901	206,755	94,900	164,900	244,900	359,000	699,000
Mentions renovations								
Other	1,384,235	0.104	0.305	0	0	0	0	1
iBuyer	2,158	0.023	0.149	0	0	0	0	0
Flipper	106,714	0.288	0.453	0	0	0	1	1
Total listings								
Other	1,384,235	3.169	2.202	1	2	2	4	7
iBuyer	2,158	6.113	3.629	3	4	5	7	13
Flipper	106,714	3.086	2.518	1	2	2	4	8
Has sale								
Other	1,384,235	0.602	0.490	0	0	1	1	1
iBuyer	2,158	0.823	0.381	0	1	1	1	1
Flipper	106,714	0.566	0.496	0	0	1	1	1
Days on market								
Other	833,190	91.239	84.921	25	43	63	107	247
iBuyer	1,777	86.977	64.126	28	42	66	111	221
Flipper	60,440	88.784	77.086	25	43	66	106	224
Discount to list								
Other	832,090	-0.037	0.068	-0.149	-0.062	-0.026	0	0.032
iBuyer	1,777	-0.024	0.041	-0.097	-0.044	-0.017	0	0.028
Flipper	60,344	-0.041	0.067	-0.150	-0.067	-0.030	0	0.025

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 the 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. When not included as an explicit category, *Flippers* are treated as non-iBuyer buyers/sellers, because they cannot be identified in the Corelogic data. Transaction data is from Corelogic and MLS data is provided by ATTOM Data 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 (0.006)	-0.272 (0.006)	-0.275 (0.006)	- -	- -	- -
Seller is iBuyer	- -	- -	- -	0.121 (0.006)	0.135 (0.005)	0.128 (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 (0.007)	-0.057 (0.007)	1.338 (0.037)	0.136 (0.009)	27.115 (1.700)	26.191 (1.639)	0.309 (0.024)	0.309 (0.024)	-0.019 (0.024)	-0.020 (0.024)	-0.005 (0.001)
Flipper	0.008 (0.001)	0.142 (0.001)	0.172 (0.006)	-0.012 (0.001)	2.180 (0.317)	1.394 (0.307)	-0.080 (0.005)	-0.079 (0.005)	-0.069 (0.005)	-0.070 (0.005)	-0.002 (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). *Flippers*, as defined previously, are treated as non-iBuyer individuals because they cannot be identified directly in the Corelogic data. 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 (1.28)	24.66 (19.83)	9.28 (9.09)	1.02 (0.69)	11.21 (15.66)
iBuyer	0.39 (0.38)	4.91 (6.39)	17.78 (19.72)	113 (117)	24.15 (8.67)

Panel B: Gross return (spread) regressions

	<i>Dependent variable:</i>					
	Gross Return (ann)	Index Return (ann)	Index Return (ann)	Index Return (ann)	Non Index Return (ann)	Non Index Return (ann)
	(1)	(2)	(3)	(4)	(5)	(6)
iBuyer	0.074 (0.002)	0.066 (0.002)	0.004 (0.001)	0.015 (0.001)	0.068 (0.002)	0.050 (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 zero otherwise (Column 1 and 2), and the dummy variable taking the value of one if the property seller was an iBuyer and 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 story (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) use 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 R² 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. *Flippers*, as defined previously, are treated as non-iBuyer individuals because they cannot be identified in the Corelogic data alone. 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

	<i>Dependent variable: Log(sale amount)</i>			
	(1)	(2)	(3)	(4)
Buyer_is_iBuyer	-0.036 (0.005)	-0.031 (0.005)	- -	- -
Seller_is_iBuyer	- -	- -	0.016 (0.004)	0.019 (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

		<i>All Markets</i>			<i>High iBuyer Market</i>	<i>Other Markets</i>
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 pre-iBuyer entry Corelogic transaction data between 2006 and ends at the end of 2012. 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 2008 and 2012, with Column (3) corresponding to the pricing model and (4) corresponding to the errors model. Standard errors are shown in parentheses.

	2006-2012 (Main specification)		2008-2012 (Robustness)	
	<i>Dependent variable:</i>			
	Log(house price)	Squared deviation from predicted price	Log(house price)	Squared deviation from predicted price
	(1)	(2)	(3)	(4)
House age (omitted: > 50 years)				
Age < 5 years	0.662 (0.003)	-0.163 (0.001)	0.705 (0.003)	-0.171 (0.001)
Age 5-15 years	0.563 (0.002)	-0.161 (0.001)	0.602 (0.003)	-0.175 (0.001)
Age 15-50 years	0.269 (0.002)	-0.111 (0.001)	0.292 (0.003)	-0.113 (0.001)
Land square footage (omitted: > 25k)				
Square footage < 5k	-0.913 (0.003)	-0.045 (0.002)	-0.936 (0.004)	-0.034 (0.002)
Square footage 5-10k	-0.601 (0.003)	-0.062 (0.002)	-0.612 (0.003)	-0.048 (0.002)
Square footage 10-25k	-0.281 (0.003)	-0.030 (0.002)	-0.274 (0.003)	-0.024 (0.002)
Multistory	0.197 (0.001)	-0.009 (0.001)	0.195 (0.001)	-0.022 (0.001)
Other house characteristics	Y	Y	Y	Y
Observations	889,661	889,242	680,640	680,360
R ²	0.661	0.037	0.623	0.044

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, \widehat{e}_{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. Columns (1)-(3) use the full sample; columns (4)-(6) examine only 2018 transactions. 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) and (4) to (2), (3), (5), and (6). *Flippers*, as defined previously, are treated as non-iBuyer individuals. Standard errors, clustered at the property level, are shown in parentheses.

	<i>Dependent variable:</i>					
	iBuyer buyer (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
\widehat{e}_{izt}	-4.713 (0.242)	- (0.474)	-5.227 (0.840)	-9.360 (0.840)	- (0.549)	-7.962 (1.454)
$\widehat{SellsWithin90Days}$	- (0.192)	1.919 (0.199)	1.582 (0.199)	- (0.549)	4.235 (0.577)	3.713 (0.577)
Sample	2014-2018			2018		
Zip x Quarter FE	Y	Y	Y	Y	Y	Y
Observations	557,172	259,772	259,772	96,273	50,725	50,725
R ²	0.027	0.030	0.031	0.025	0.028	0.028

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
List price (\$k)	262.00	262.00
iBuyer market share (%)	4.88%	4.88%
iBuyer buy discount (%)	-2.67%	-2.90%
iBuyer sale premium (%)	0.50%	-0.01%
HH time on market (days)	91.00	90.99
IB time on market (days)	97.00	96.99
Impatient price delta (%)	3.50%	0.14%
Mean pricing error (%)	14.30%	14.01%
d(iB share)/d(Pricing error)	-7.735	-7.735
d(iB share)/d(P(sells in 90 days))	0.037	0.060
P(iBuyer repair) (%)	5.9%	11.3%
P(iBuyer cut prices) (%)	56%	65%

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
ϕ_R	Probability needs renovation	0.109	Fraction of listings mentioning renovation
p_I	Probability impatient	0.500	Definition
τ	iBuyer closing time (days)	15	Industry reports
\bar{u}	Matched flow utility	24.000	Normalization for numerical performance

Panel C: Parameters Calibrated by Method of Moments

Parameter	Description	Value
\underline{u}	Unmatched patient utility flow (\$k/dt)	0.76
$\underline{\underline{u}}$	Unmatched impatient utility flow (\$k/dt)	-9.23
m_l	iBuyer baseline maintenance cost (\$k/dt)	-7.86
η	iBuyer house depreciation arrival rate (rate/dt)	7.00
$m_h - m_l$	iBuyer flow cost increase from depreciation (\$k/dt)	5.00
λ	Matching technology (rate/dt)	135
λ_{ib}	iBuyer matching scalar (unitless)	0.94
ξ	iBuyer signal noise	0.003
R	Repair cost (\$k)	1.49
σ_i	T1EV variance on iBuyer preference	0.41
σ_m	T1EV variance on house preference	1.00

Table 8: Equilibrium effects of iBuyers

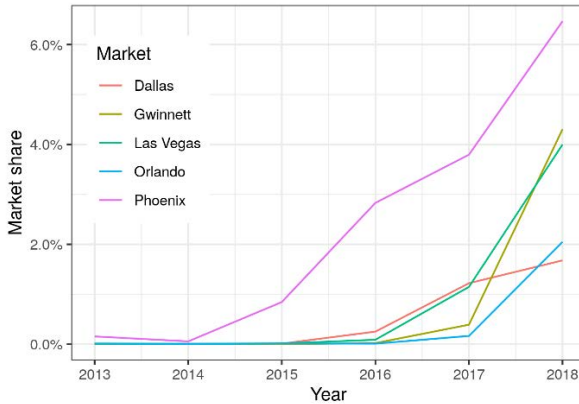
This table shows key equilibrium outcomes from a market with iBuyers and a market without iBuyers. The “without iBuyers” benchmark uses the model calibrated with iBuyers and removes them; the “with iBuyers” benchmark” uses the calibrated model. $E[\text{time to sale}]$ is the mean time to sale in days. $E[\text{time to sale} | \text{patient}]$ is the mean time to sale in days for patient homeowners. $E[\text{time to sale} | \text{impatient}]$ is the mean time to sale for impatient homeowners. Both numbers include the direct effects of iBuyers (e.g., selling to iBuyers) and the indirect effects of iBuyers (e.g., iBuyers increasing liquidity overall).

Panel A: Equilibrium outcomes with/without iBuyers

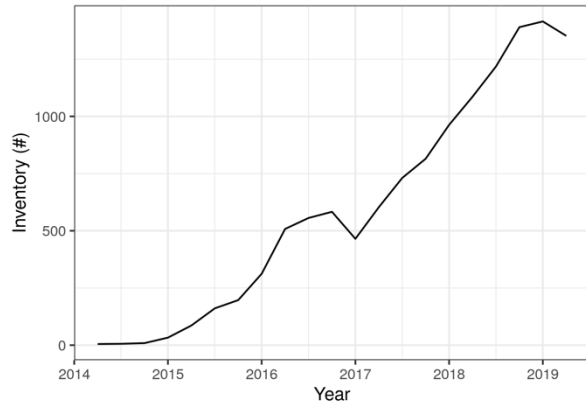
Quantity	Without iBuyers	With iBuyers
Median listing price	260	262
$E[\text{time to sale}]$	91	87
$E[\text{time to sale} \text{patient}]$	94	93
$E[\text{time to sale} \text{impatient}]$	88	80
iBuyer share	-	4.88%
iBuyer share (patient)	-	0.07%
iBuyer share (impatient)	-	9.69%

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

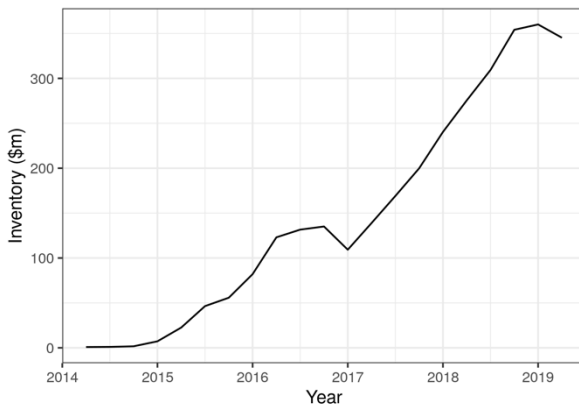
Panel (a) shows iBuyer market share in buying or selling transactions across five large markets: Dallas, Texas, Gwinnett County, Georgia, Las Vegas, Nevada, Orlando, Florida, and Phoenix, Arizona using Corelogic data. Panel (b) shows the stock of houses owned by iBuyers at the end of every quarter. Panel (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 gross return (spread), that iBuyers earn on purchased and sold homes with 25% and 75% bands shown.



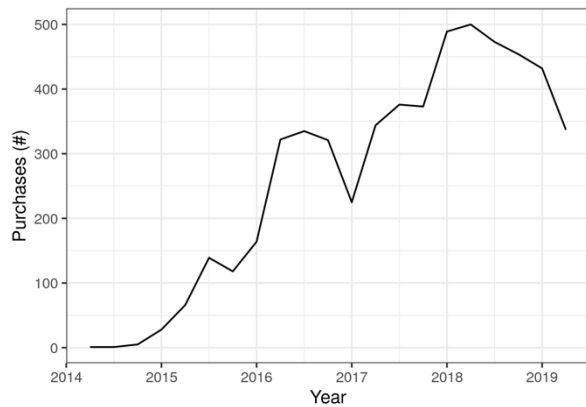
(a) Market shares of transactions



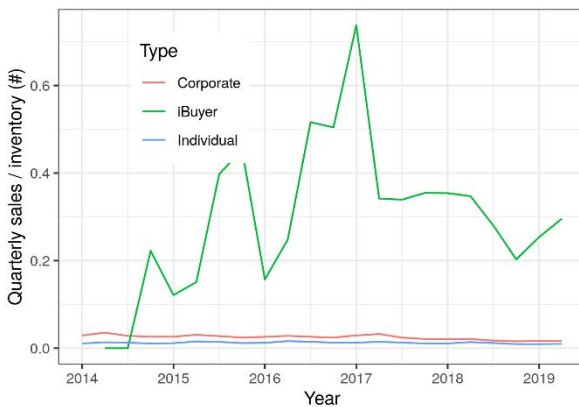
(b) Houses in inventory



(c) Dollar value of houses in inventory



(d) iBuyer purchases



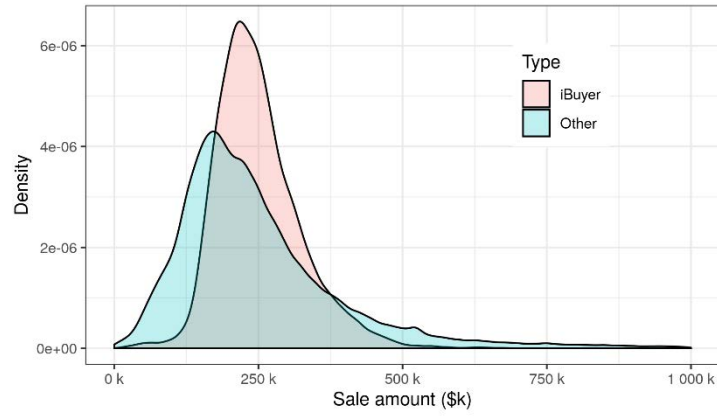
(e) iBuyer inventory turnover



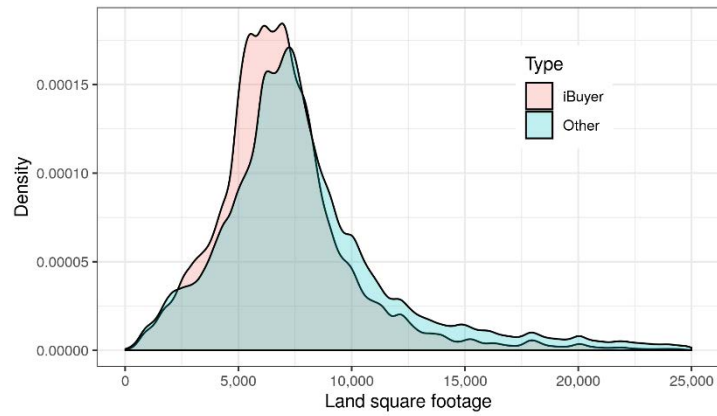
(f) iBuyer realized gross return

Figure 2: Characteristics of iBuyer Houses

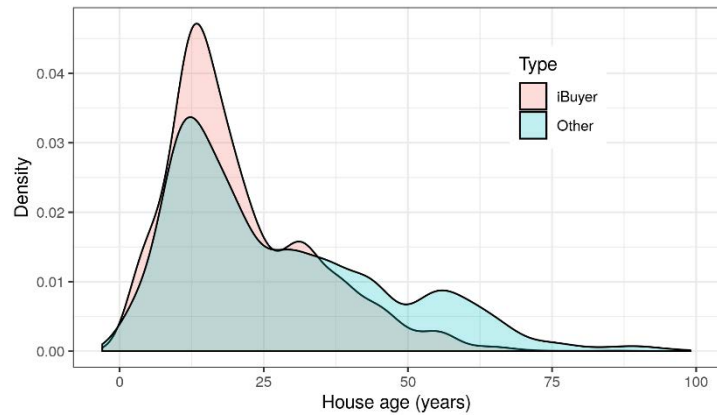
This figure shows the distribution of house prices (panel a), land square footage (panel b), and house age (panel c) for iBuyer (red) versus non-iBuyer (blue) house purchases. The figure uses transaction-level data from Corelogic between 2013 and 2018.



(a) House sale price



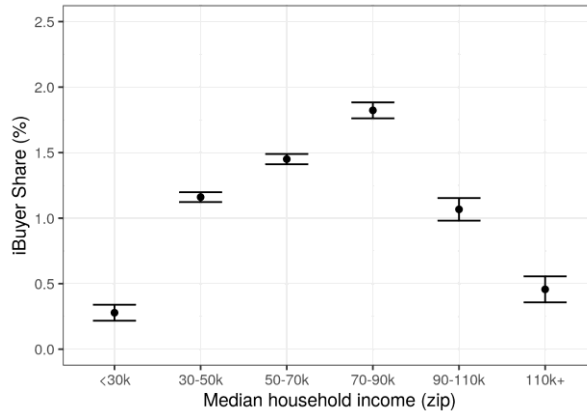
(b) Land square footage



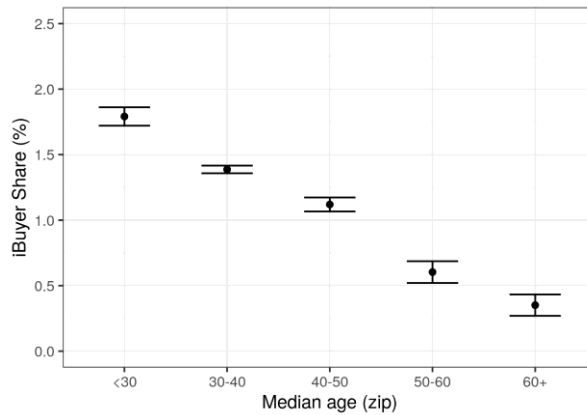
(c) House age

Figure 3: Demographics of iBuyer Markets

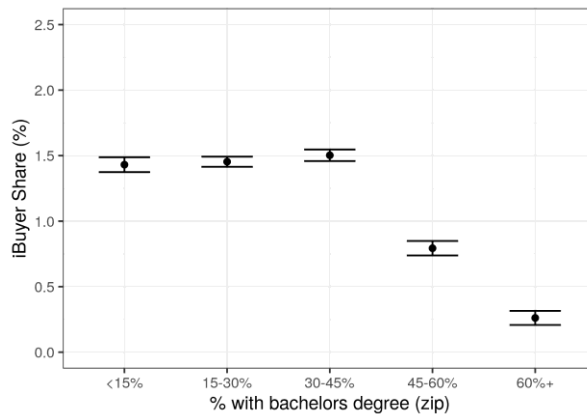
This figure shows average iBuyer market share in a zip code using Corelogic transaction data in 2013-2018 versus binned demographics at the zip code level. We use the American Community Survey (ACS) data from the U.S. Census Bureau to measure several zip-level demographic characteristics. iBuyer market share is defined as the fraction of houses that are sold to iBuyers. Panel (a) shows iBuyer share across median household income; (b) median age, and (c) college education. Bars represent 95% confidence intervals.



(a) iBuyer share versus household income



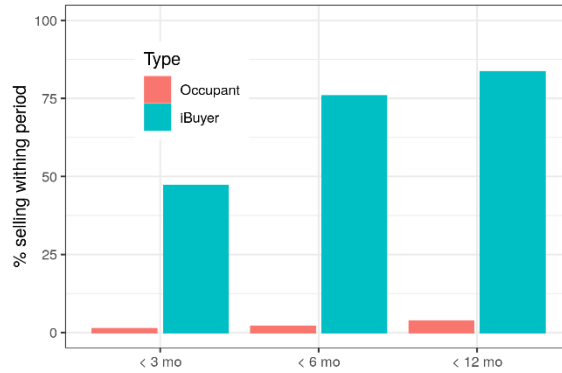
(b) iBuyer share versus age



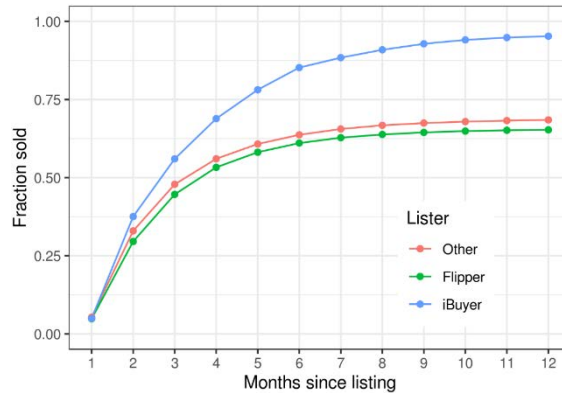
(c) iBuyer share versus % with college education

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

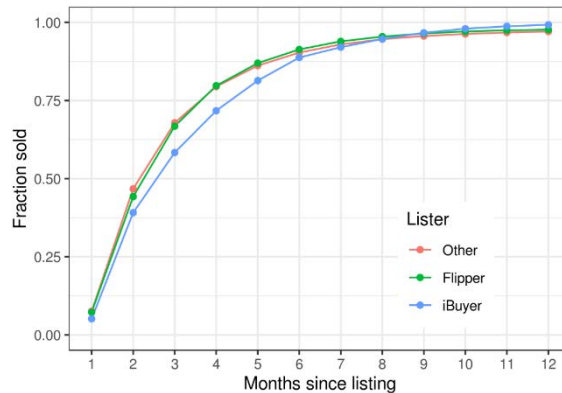
Panel (a) shows the fraction of houses purchased and sold within one year using housing transaction data between 2013 and 2017, omitting 2018 to have a full year of data for the most recent transactions. Panel (b) shows the cumulative fraction of houses sold by month after listing for all listings in 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. Note that flippers are observable only in MLS data. Transaction-level data is from Corelogic and listing data is from ATTOM Data.



(a) Fraction sold by owner type



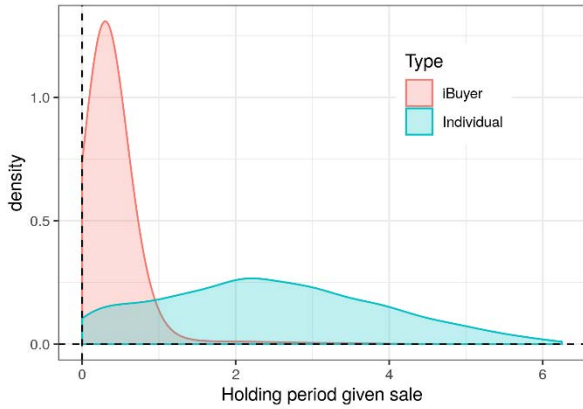
(b) Fraction sold by month among all listings



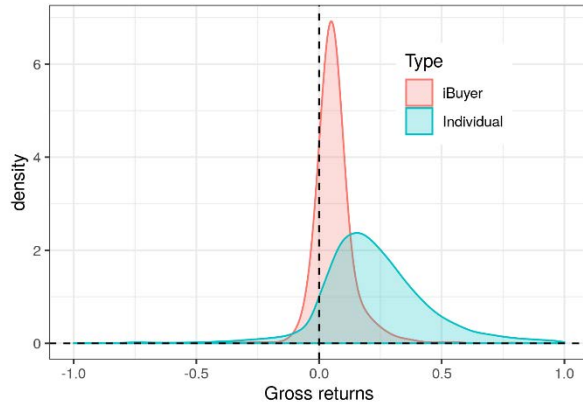
(c) Fraction sold by month among sold listings

Figure 5: Gross Returns: iBuyers versus Individuals

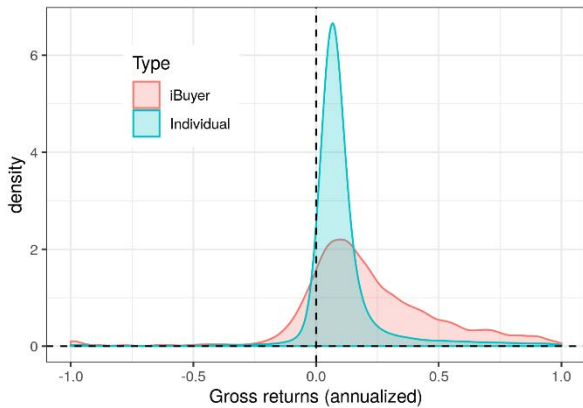
This figure shows holding periods and realized gross returns for iBuyers (red) versus non-corporate individuals (blue) from Corelogic transaction data. Observations are all purchases in Corelogic between 2013 and 2018 where the buyer sells the house during the sample period. Panel (a) shows the distribution 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.



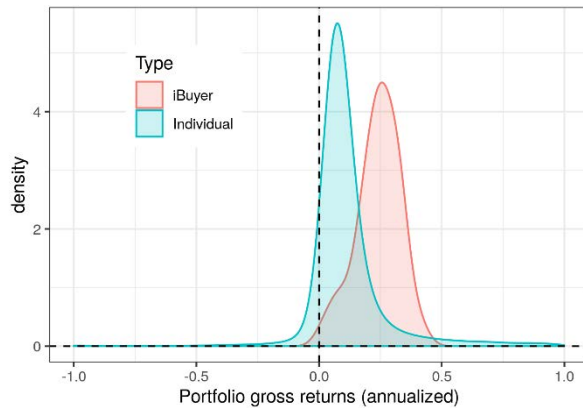
(a) Holding period given sale



(b) Gross returns



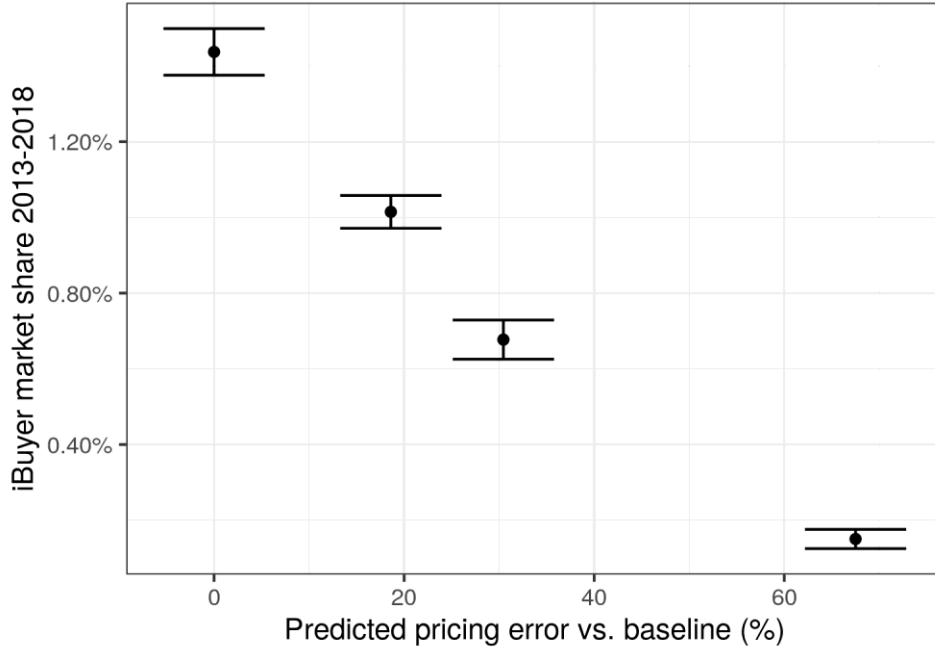
(c) Gross returns (annualized)



(d) Portfolio gross returns (annualized)

Figure 6: iBuyer Market Share and Pricing Errors and Liquidity

This figure shows iBuyer market share (Y axis) across our two proxies for ease of algorithmic pricing and liquidity. iBuyer market share is the fraction of homes purchased by iBuyer. In Panel (a), predicted pricing error is the absolute predicted residual based on house hedonics. 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. We use Corelogic transaction data and MLS data from ATTOM Data covering period between 2013 and 2018.



(a) Pricing error and iBuyer market share

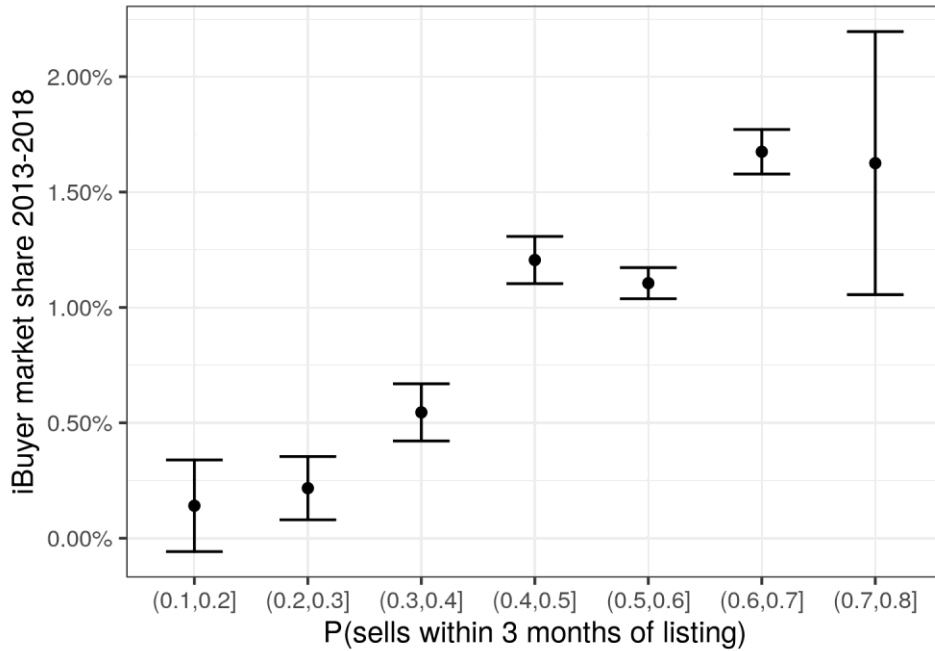


Figure 7: Equilibrium Housing Trading Model with iBuyers: Transition Paths

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 $\lambda_{ib} F(m_s, m_{ib})$.

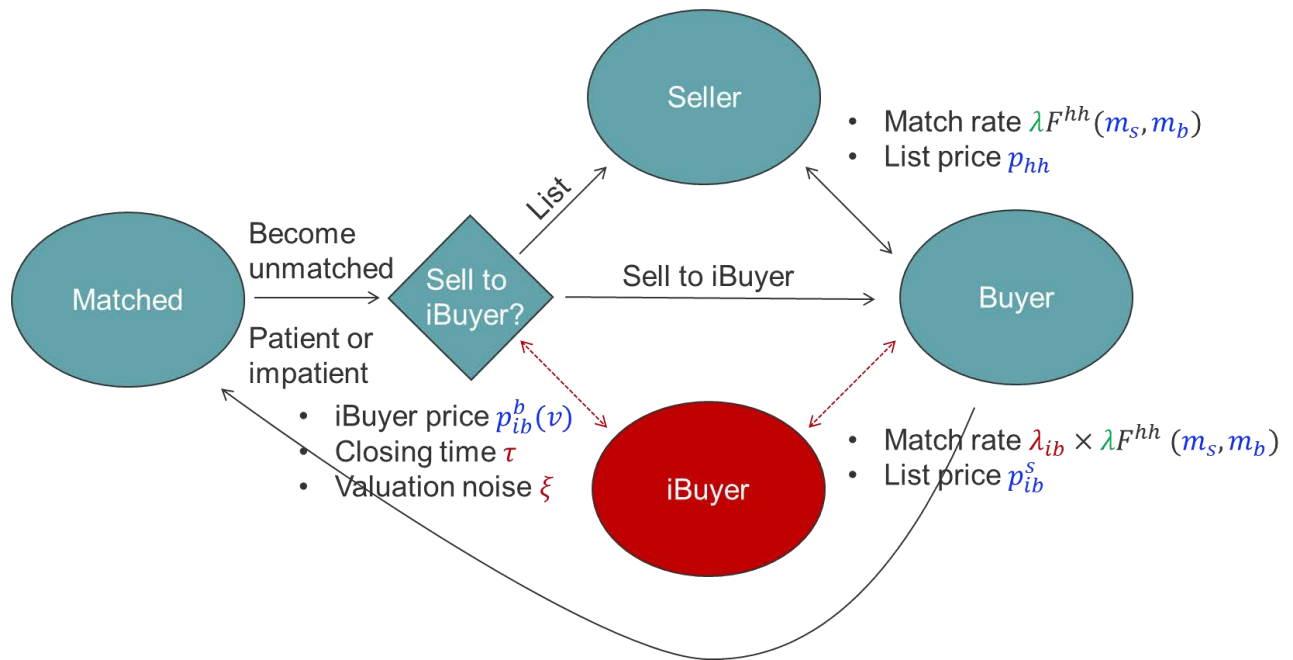
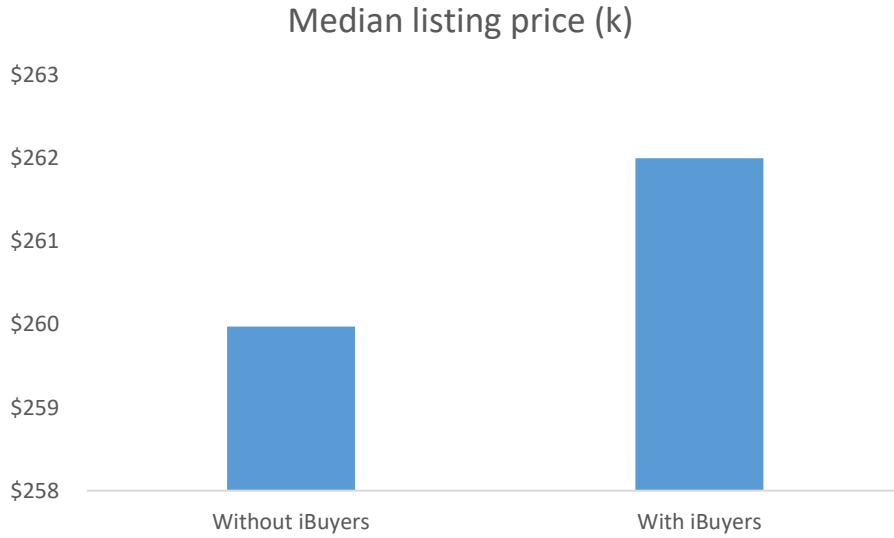
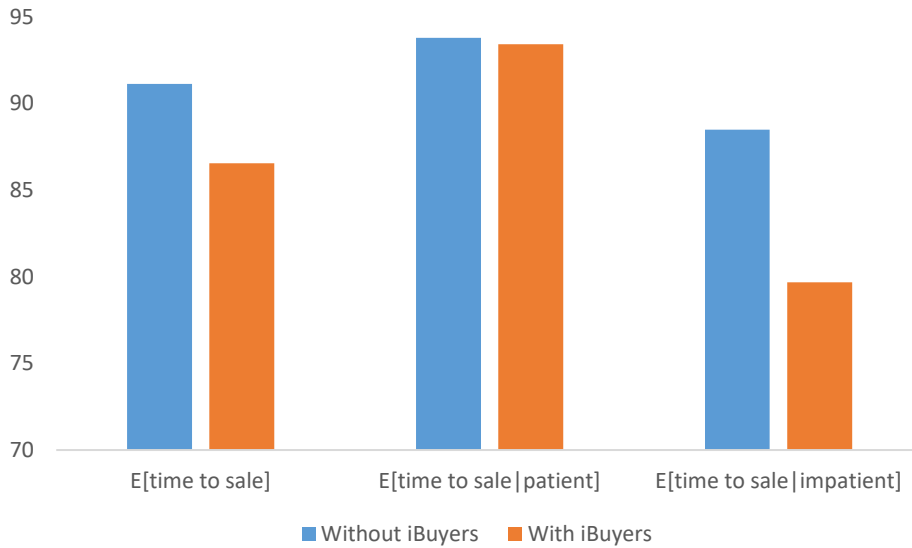


Figure 8: Housing prices and transition speeds

This figure shows equilibrium listing prices (Panel (a)) and expected times to sell (Panel (b)), with and without iBuyers. The expected time to sell includes the expected time when listing for those not selling to iBuyers and the closing time for those that do sell to iBuyers. Patient sellers are those with a higher flow utility while unmatched; impatient sellers are those with a lower flow utility when unmatched.



(a) Equilibrium prices



(b) Expected time to sell

Figure 9: iBuyer market share under different business models

Panel A shows iBuyer market share varying the key economic forces. Baseline is the iBuyer as estimated from the data. The “slow iBuyer” is an iBuyer with the same parameters as the baseline iBuyer, except it takes 90 days to close the sale. The “renting iBuyer” is an iBuyer who rents the house to a (patient) transitioning homeowner at a price to keep the homeowner indifferent between occupying and not occupying the property. The “bad model iBuyer” is an iBuyer where the cost of mispricing the home is \$5k, rather than the roughly \$1k that is estimated. Panel B compares the baseline “high tech” iBuyer to a hypothetical iBuyer which is forced to compromise somewhat in technology: it can transact either relatively slowly (25 days) but with similar valuation technology, or quickly but with worse valuation technology ($\xi = 0.25$).

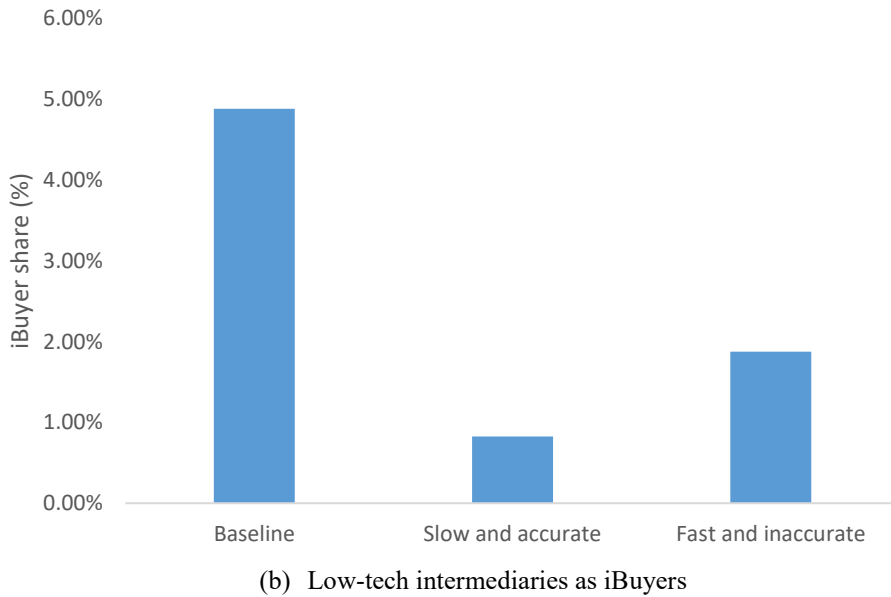
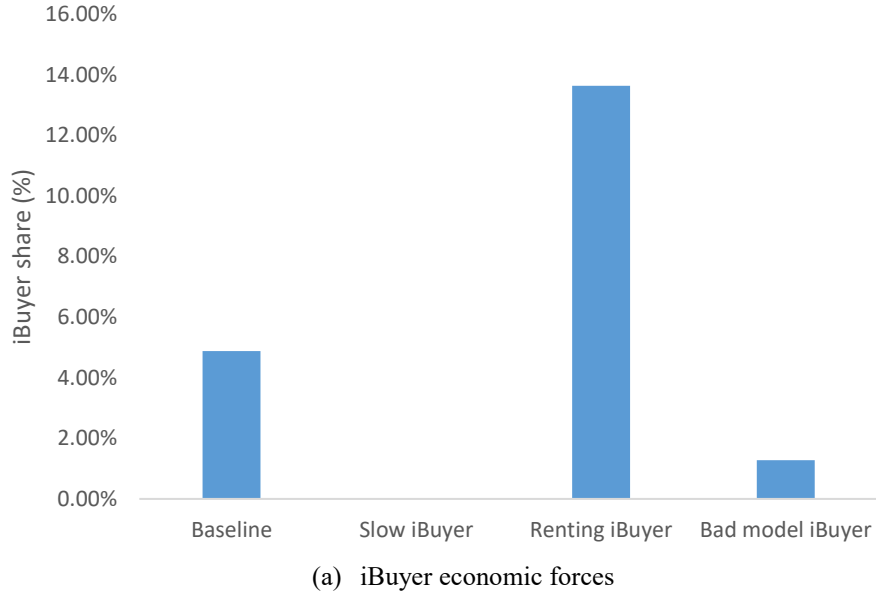
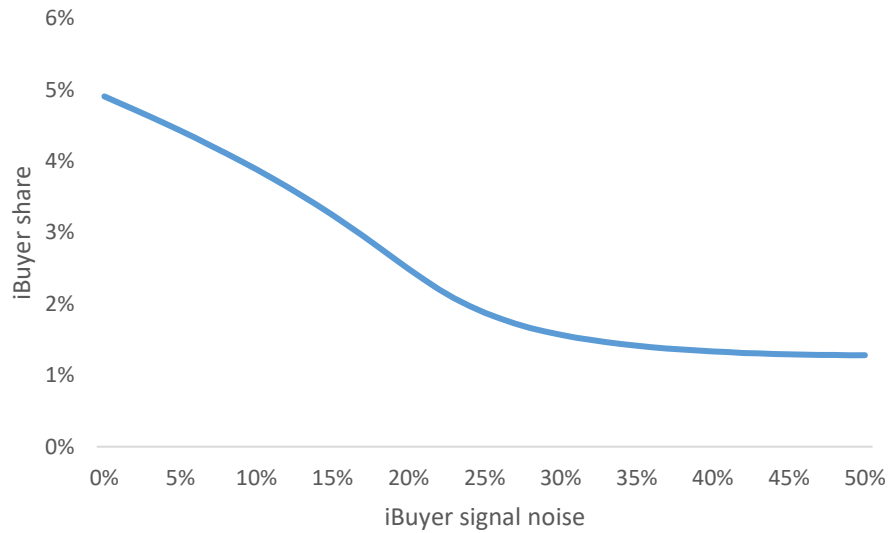
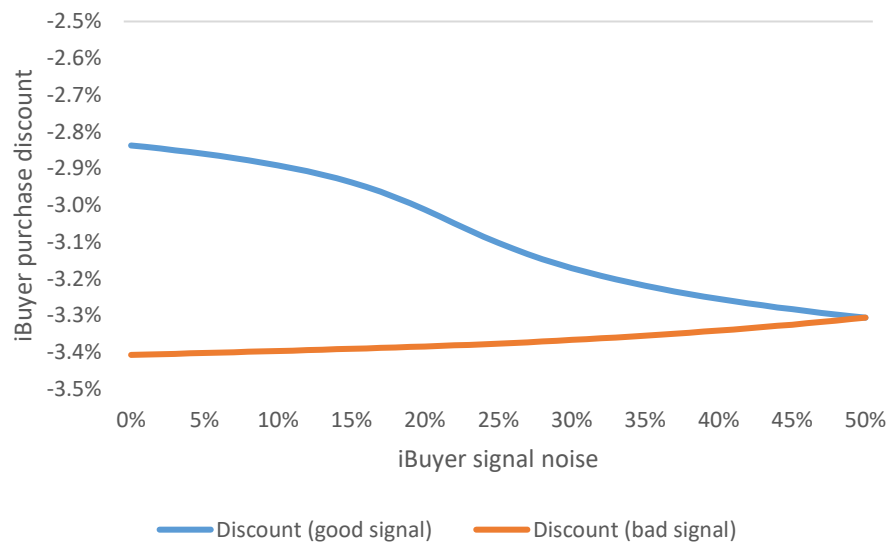


Figure 10: iBuyers and Valuation Accuracy

This figure shows iBuyer market share (Panel (a)) and iBuyer pricing discounts for good and bad signals (in % relative to the non-iBuyer listing price) (Panel (b)) versus the noise of the iBuyer's signal, ξ . Panels (c) and (d) show how iBuyer exposure to adverse selection varies with noise: Panel (c) shows the fraction of iBuyer purchases that need repairs; Panel (d) shows the fraction of iBuyer purchases that need repairs conditional on the iBuyer getting a signal that the house does not need repairs.

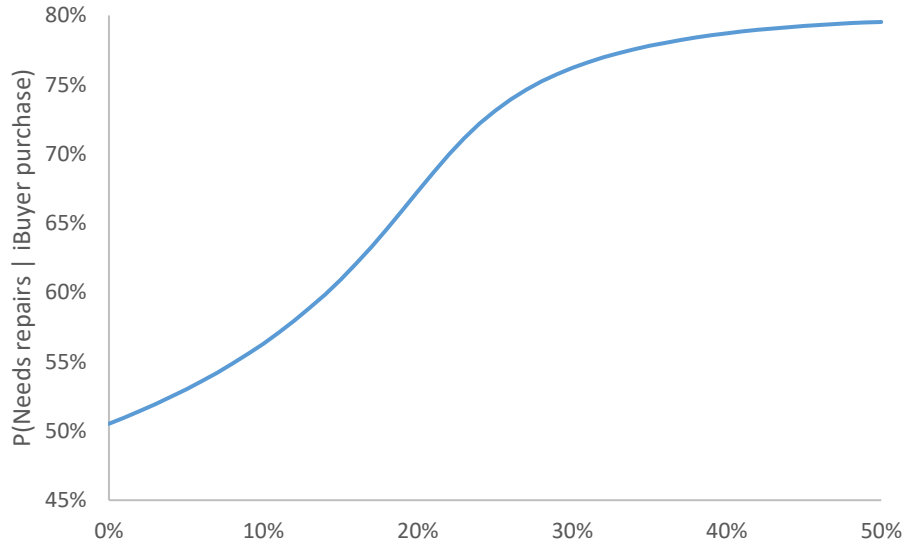


(a) iBuyer Share

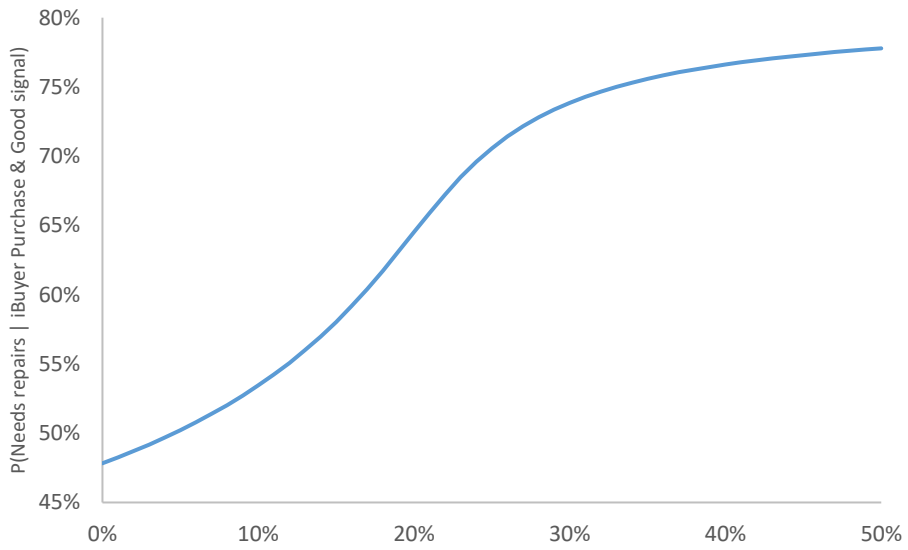


(b) iBuyer offer price discount

Figure 10: iBuyers and Valuation Accuracy, Continued



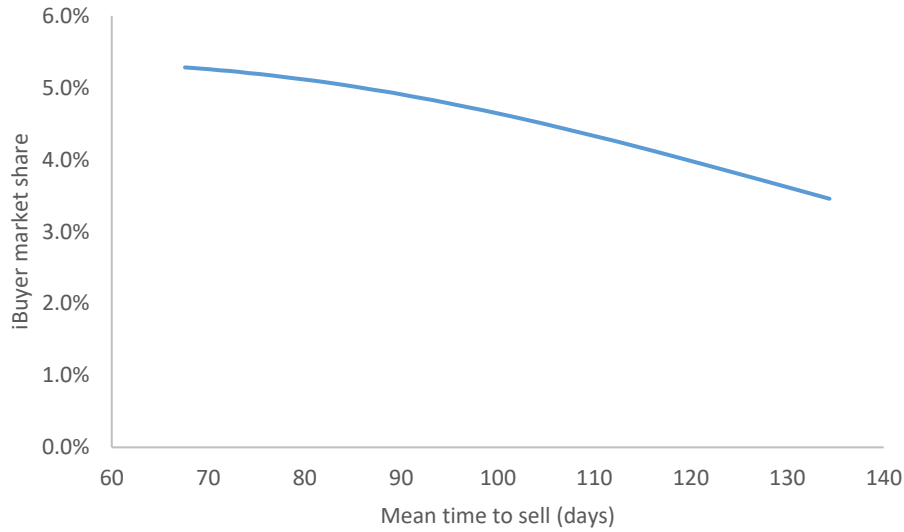
(c) iBuyers buying bad homes



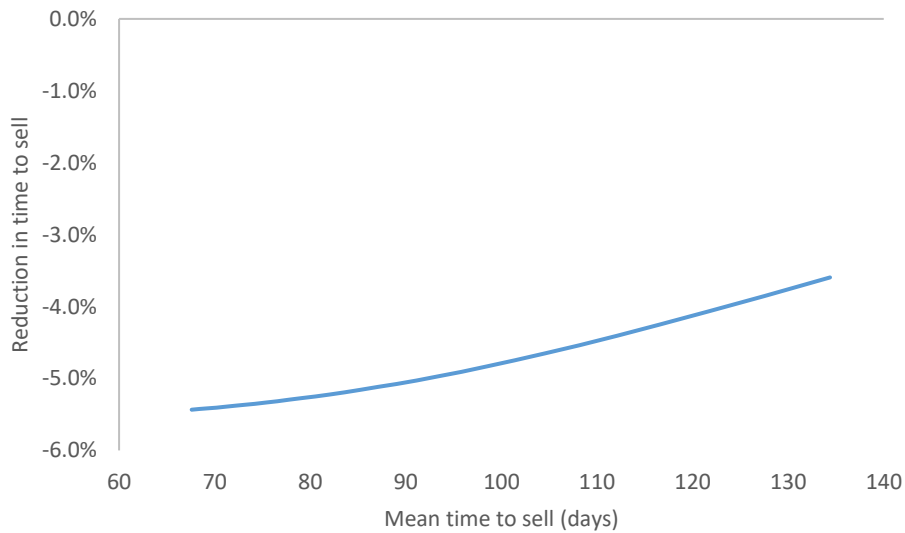
(d) iBuyers and pricing mistakes

Figure 11: iBuyers in Illiquid Markets

This figure shows iBuyer market share (Panel (a)), the reduction in time to sell from introducing iBuyers into markets (in % versus the non-iBuyer baseline) versus baseline market liquidity (Panel (b)). Baseline market liquidity is defined as the pre-iBuyer mean time to sell. Panels (c) and (d) show how iBuyer exposure to adverse selection varies with liquidity: Panel (c) shows the fraction of iBuyer purchases that need repairs; Panel (d) shows the fraction of iBuyer purchases that need repairs conditional on receiving a signal that the house does not need repairs.

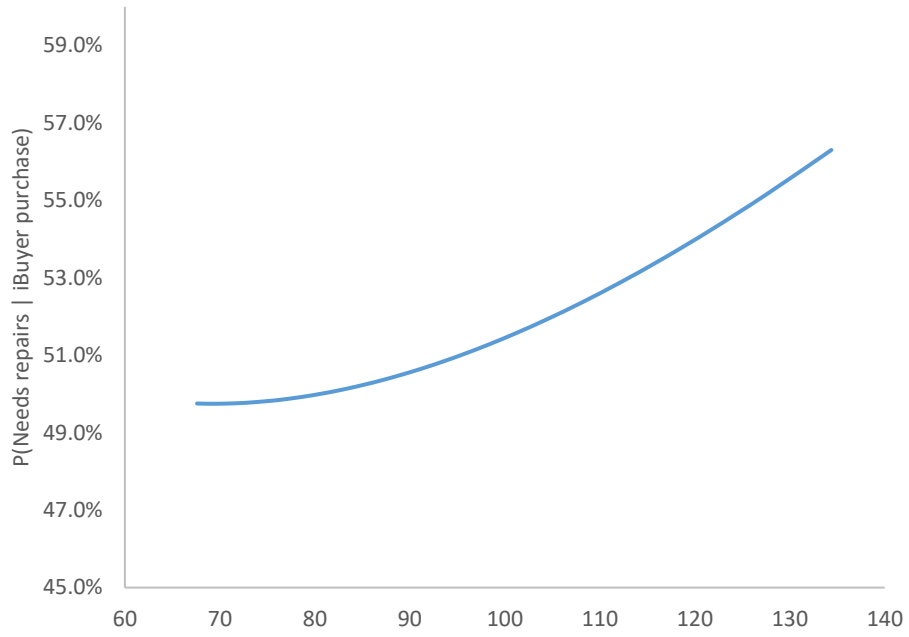


(a) iBuyer market share

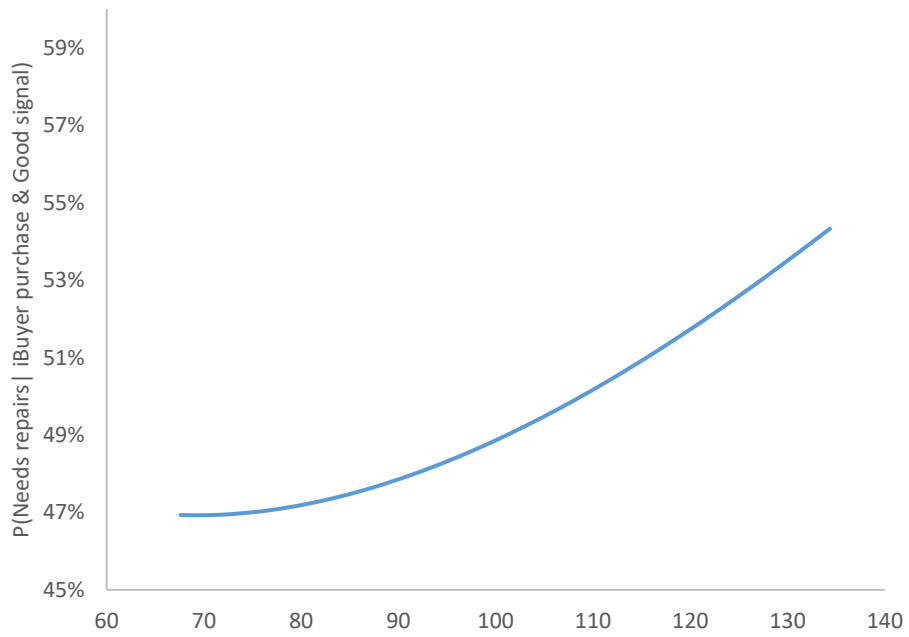


(b) Reduction in time to sell

Figure 11: iBuyers in Illiquid Markets, Continued



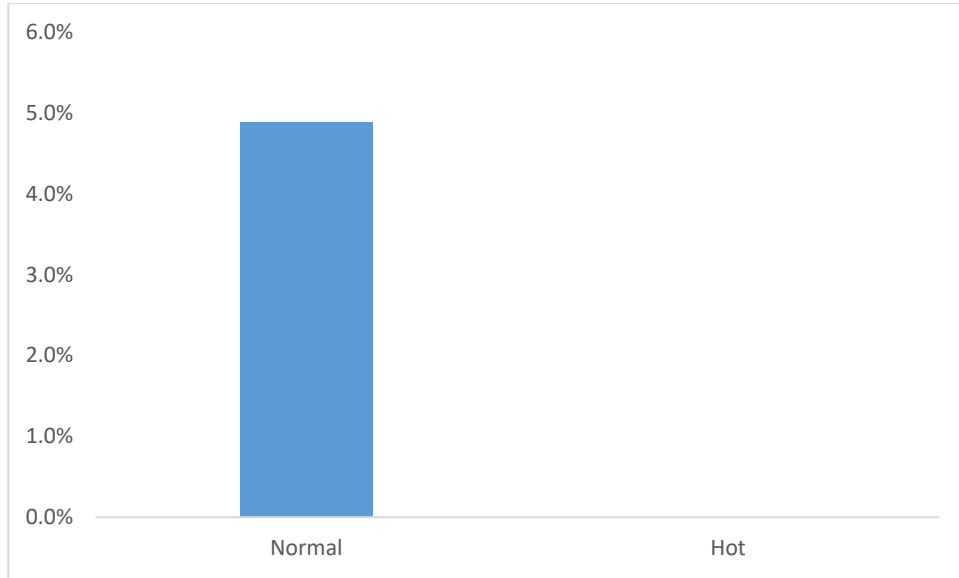
(c) iBuyers buying bad homes



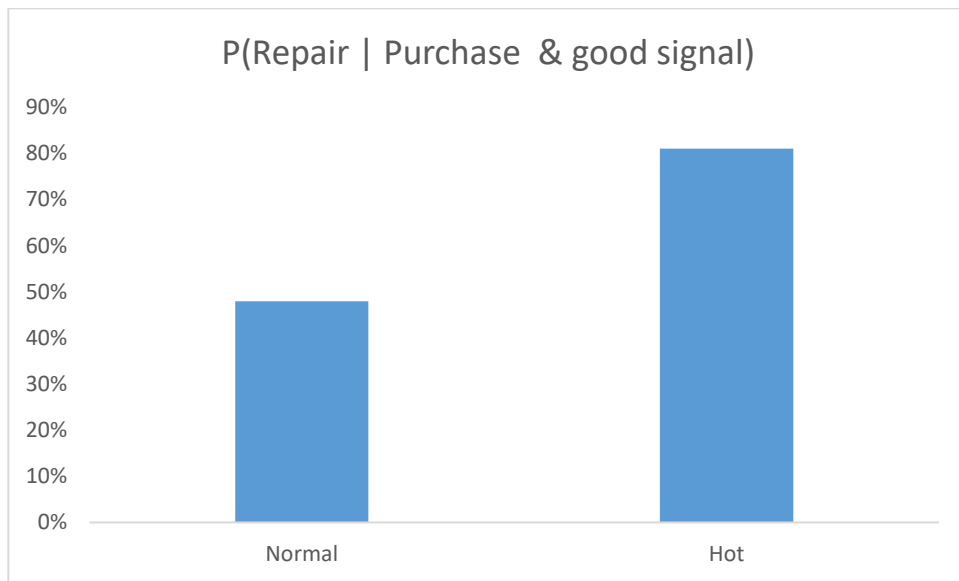
(d) iBuyers and pricing mistakes

Figure 12: iBuyers in “Hot” Markets

This figure shows iBuyers in a “hot” market, where the matching rate is set to generate a time to sale of roughly 30 days. Panel (a) shows iBuyer market share. Note that in the “hot” market, iBuyer market share is close to, but not exactly zero. Panel (b) shows iBuyer pricing mistakes, defined as the fraction of houses that iBuyers purchase that require repairs despite having good signals.



(a) iBuyer market share

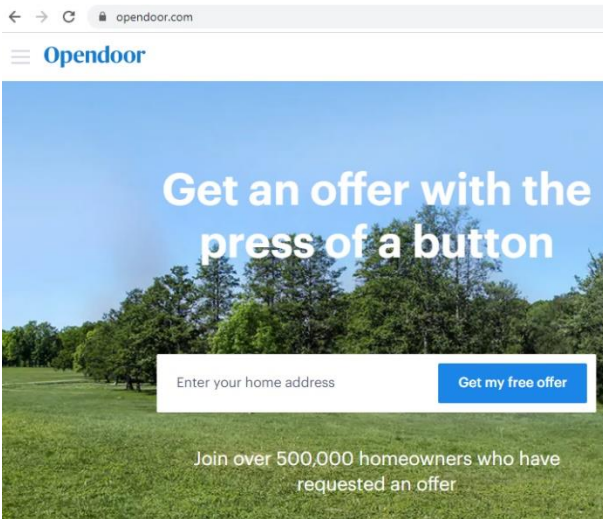


(b) iBuyers and pricing mistakes

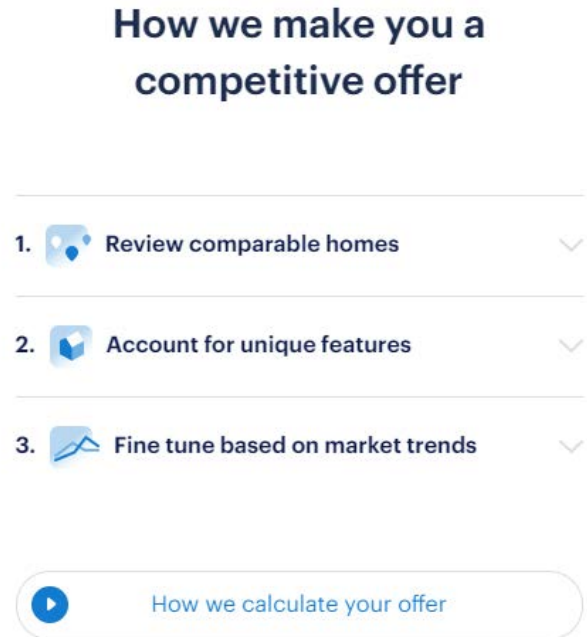
Online Appendix

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

Why Opendoor is better

Selling to Opendoor	VS	Traditional home sale
Competitive cash offer in 24 hours ✓		× Risk of buyer financing fall-through
No listing, prep work, or showings ✓		× Hours of prep work and home showings
Skip the repair work and deduct the costs ✓		× Manage repairs yourself
Choose any close date from 10-60 days ✓		× Uncertain closing timeline

(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,³⁰ Offerpad,³¹ Knock,³² Zillow Offers,³³ and RedfinNow.³⁴ 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

³⁰ <https://www.opendoor.com/>

³¹ <https://www.offerpad.com/>

³² <https://www.knock.com/>

³³ <https://www.zillow.com/offers/>

³⁴ <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: iBuyers and previously listed homes

This table shows match rates for iBuyer purchases and MLS listings and whether iBuyer pricing depends on whether the house was previously listed. Panel A shows summary statistics for matching between MLS and deeds records. Match window is the allowed time between Corelogic sale date and MLS sale date to be considered a match. Panel B shows pricing differences for iBuyer purchases that correspond to listings versus those that do not. : Columns (1), (3), (5), and (7) include only iBuyer buys where there is no seller listing. Columns (2), (4), (6), and (8) include only iBuyer buys where there is a seller listing. Columns (1)/(2), (3)/(4), (5)/(6), and (7)/(8) vary the MLS-deeds sale date match window from exact (day), week, month, and quarter. All columns include house hedonics and zip x quarter fixed effects.

Panel A: iBuyer purchases-to-MLS-listings match rate by match window

Match window	# iBuyer buys with listings	% iBuyer buys with listings	# other buys with listings	% other buys with listings
Day	84	1.3	176,827	14.7
Week	491	7.5	418,414	34.8
Month	505	7.7	501,735	41.7
Quarter	1,464	22.2	654,702	54.4

Panel B: iBuyer purchase price differences by whether there is an MLS listing

	log(Sale price)							
	MLS-Deeds match window							
	Day		Week		Month		Quarter	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
iBuyer Buyer	-0.034*** (0.004)	0.036 (0.035)	-0.035*** (0.004)	-0.017 (0.014)	-0.035*** (0.004)	-0.016 (0.014)	-0.031*** (0.004)	-0.042*** (0.008)
iBuyer sample	Unlisted	Listed	Unlisted	Listed	Unlisted	Listed	Unlisted	Listed
N iBuyer buys	6505	84	6098	491	6084	505	5125	1,464
House hedonics	Y	Y	Y	Y	Y	Y	Y	Y
Zip x Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y
N	1,189,255	1,182,906	1,188,848	1,183,313	1,188,836	1,183,325	1,187,897	1,184,264
R ²	0.753	0.753	0.753	0.753	0.753	0.753	0.753	0.753

Note:

*p<0.1; **p<0.05; ***p<0.01

Appendix A.5: iBuyer purchases and failed listings

This table examines whether iBuyers are likely to buy from houses with failed listings. A failed listing is one with a listing date 90, 182, or 365 days prior to the iBuyer purchase that *does not* result in a sale. Panel A shows summary stats, where we vary the window for finding a corresponding failed listing between 90, 182, and 365 days. Panel B examines pricing differences between iBuyer purchases corresponding to failed listings and non-failed listings. Columns (1), (3), (5), and (7) include only iBuyer buys where there is no failed seller listing within the window. Columns (2), (4), (6), and (8) include only iBuyer buys where there is a failed prior seller listing within the window. Columns (1)/(2), (3)/(4), (5)/(6), and (7)/(8) vary the MLS-deeds sale date match window from 90, 182, and 365 days prior. All columns include house hedonics and zip x quarter fixed effects.

Panel A: iBuyer purchases from failed listings

Window	# iBuyer buys with failed listings	% iBuyer buys with failed listings	# other buys with failed listings	% other buys with failed listings
90 days	12	0.2%	6,627	0.8%
182 days	18	0.3%	10,711	1.3%
365 days	27	0.4%	15,931	2.0%

Panel B: iBuyer pricing and failed listings

	log(Sale price):					
	Prior failed listing window					
	90 days		182 days		365 days	
	(1)	(2)	(3)	(4)	(5)	(6)
iBuyer Buyer	-0.034*** (0.004)	0.092 (0.092)	-0.034*** (0.004)	0.055 (0.078)	-0.034*** (0.004)	0.050 (0.073)
iBuyer sample	No prior	Failed prior	No prior	Failed prior	No prior	Failed prior
N iBuyer buys	5693	12	5687	18	5678	27
House hedonics	Y	Y	Y	Y	Y	Y
Zip x Quarter FE	Y	Y	Y	Y	Y	Y
N	800,442	794,807	800,438	794,811	800,436	794,813
R ²	0.710	0.710	0.710	0.710	0.710	0.710

Note:

*p<0.1; **p<0.05; ***p<0.01

Appendix A.6: 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.

<i>Dependent variable:</i>							
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 (0.003)	0.387 (0.010)	-0.009 (0.002)	4.862 (0.391)	-0.101 (0.005)	-0.099 (0.005)	-0.002 (0.0003)
iBuyer	-0.021 (0.066)	1.444 (0.064)	0.149 (0.011)	29.733 (2.141)	0.240 (0.028)	-0.095 (0.028)	-0.007 (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.7: iBuyer Seasonality

Seasonality plays an important role in residential real estate transactions. For example, Ngai and Tenreyo (2014) document that every year housing markets in the U.K. and U.S. experience systematic above-trend increases in prices and transactions during the spring and summer (“hot season”) and below-trend falls during the autumn and winter (“cold season”). Motivated by this observation we now investigate how iBuyers listings vary across seasons and if they are able to navigate seasonality (off season versus on season) better than other sellers. To study these seasonality patterns formally, we estimate the following sets of regressions:

$$Qtr_{izt}^q = \beta Lister_{izt} + H_i' \mathbf{B} + \mu_{zt} + \epsilon_{izt}$$

Here i indexes a house in zipcode z at quarter t . Qtr^q is an indicator for whether the listing occurs in a quarter q . $Lister_{izt}$ is an indicator for whether the lister is an ordinary seller, a flipper, or an iBuyer. As before, we control for house characteristics in H_i such as square footage, and quarter \times zip fixed effects μ_{zt} .

Panel A of the Table below show the results. Each column corresponds to a quarter of listing. Relative to ordinary sellers, iBuyer listings are concentrated in quarters three and four: they are 8% and 16% less likely to occur in quarters one and two, respectively, and 4% and 20% more likely to occur in quarters three and four, respectively. Flippers, in contrast, have much smaller differences relative to ordinary sellers and do not exhibit the same strong patterns of seasonality.

We next explore whether indeed iBuyers list more aggressively and spend more time on the market during off season by estimating the following specification:

$$\log(List\ Price_{izt}) = Qtr_t + \beta Lister_{izt} \times Qtr_t + H_i' \mathbf{B} + \mu_{zt} + \epsilon_{izt}$$

$$Days\ on\ Market_{izt} = Qtr_t + \beta Lister_{izt} \times Qtr_t + H_i' \mathbf{B} + \mu_{zt} + \epsilon_{izt}$$

This mirrors the earlier regression on listing outcomes with the addition of a *Lister* times *Qtr* interaction. Panel B of the Table below show the results. Column (1) shows that iBuyer listing prices are relatively higher than other market participants in quarters three and four, and not different in quarters one and two, with initial listing prices being roughly 3% higher in the third quarter and 2.6% higher in the fourth quarter, while there are no significant differences in quarters one or two. Column (2) shows that iBuyer houses listed in the third quarter are the ones who spend longest on the market relative to other listings at the same time, spending roughly 47 days longer relative to ordinary sellers. Finally, Column (3) shows that these are also houses iBuyers discount most upon the ultimate sale in terms of how much the eventual sale price has been discounted relative to the aggressive initial listing price.

Listing Seasonality: iBuyers versus Other Sellers

This table examines listing seasonality of iBuyers, home flippers, and other sellers using MLS data between 2013 and 2018 at the combined listing level. Panel (a) shows which quarters iBuyers list houses. Q1-4 are indicator variables for the quarter of listing. *Flipper* and *iBuyer* are indicators for whether the listers are flippers (absentee owners who list within one year of purchase) or iBuyers, with other listers being the base category Panel (b) examines the seasonality of listing characteristics. Columns (1) examines the log initial listing price and include all listings; (2) the days on market and (3) the sale-to-list discount, and include listings that lead to sales only. 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. Additionally, all columns include zip-year fixed effects. Standard errors are shown in parentheses.

Panel A: Listing seasonality

	Q1 (1)	Q2 (2)	Q3 (3)	Q4 (4)
Flipper	0.746 (0.191)	-1.413 (0.194)	-0.687 (0.190)	1.354 (0.181)
iBuyer	-8.379 (1.211)	-16.369 (1.228)	4.015 (1.202)	20.733 (1.149)
Hedonic controls	Y	Y	Y	Y
Zip-year FE	Y	Y	Y	Y
Observations	887,208	887,208	887,208	887,208
R ²	0.037	0.017	0.013	0.074

Panel B: Seasonality of listing characteristics

	Log first list price (1)	Days on market (2)	Sale to list price (3)
iBuyer x Q1	-0.003 (0.021)	19.790 (4.997)	-0.007 (0.004)
iBuyer x Q1	0.007 (0.022)	33.567 (5.032)	-0.010 (0.004)
iBuyer x Q3	0.031 (0.015)	47.228 (3.478)	-0.011 (0.002)
iBuyer x Q4	0.026 (0.014)	25.909 (3.142)	-0.003 (0.002)
Hedonic controls	Y	Y	Y
Zip-Quarter FE	Y	Y	Y
Observations	887,208	621,403	615,825
R ²	0.755	0.429	0.142

Appendix A.8: PnL Decomposition

We augment our analysis of iBuyer PnL with a simple decomposition. The objective is to separate the gross return into a component that is attributable purely to overall house price appreciation and the remainder where iBuyers buy below prevailing (median) market prices and sell above prevailing (median) prices – i.e., the bid/ask spread. In particular, at the three-digit zip code-quarter level, we calculate the median transaction price of all transactions (including iBuyers):³⁵

$$Local\ Price_{zt} = median_{i \in (z,t)}(SalePrice_i)$$

We then define the house price index appreciation in market z between time t and t' as:

$$Index\ Appreciation_{ztt'} = \frac{LocalPrice_{zt'}}{LocalPrice_{zt}} - 1$$

Then, for a house purchased at time t for price $Price_{izt}$ and sold at time t' for price $Price_{izt'}$, we define the *Index Return*, and the *Non Index Return*, as:

$$\begin{aligned} Index\ Return_{iztt'} &= Index\ Appreciation_{ztt'} \\ Non\ Index\ Return_{iztt'} &= GrossReturn_{iztt'} - Index\ Return_{iztt'} \end{aligned}$$

Table 3 Panel B shows the results of this decomposition.

³⁵ In an unreported robustness check, we use Zillow house single family house price indices at the quarter-zip code level rather than median transaction price. This index takes into account compositional changes of the types of houses trading at a given point in time. The results are qualitatively unchanged.

Appendix A.9: Evidence of Adverse Selection

We augment our analysis on iBuyers' preference for liquid and easy to price homes by investigating how iBuyer realized annual gross returns on their transactions relate to our ease-of-pricing and liquidity measures. In the body of the paper, we show that iBuyers concentrate in houses which are easy to price with simple hedonics. Here we show that the realized gross spread for iBuyers on such houses is also the highest, consistent with the notion that iBuyers may face adverse selection in harder to price homes.

To investigate this formally, we regress the realized annualized gross return of sellers on the expected pricing errors and liquidity as follows:

$$\begin{aligned} \text{Gross Return}_{iztt'}^{Ann} = & \beta_0 iBuyer_{izt} + \beta_1 \hat{e}_{izt}^2 + \gamma_1 \widehat{\text{Sells Within 90 Days}}_{izt} + \\ & iBuyer_{izt} \times (\beta_2 \hat{e}_{izt}^2 + \gamma_2 \widehat{\text{Sells Within 90 Days}}_{izt}) + \mu_{zt} + \epsilon_{iztt'} \end{aligned}$$

An observation is a house sale, where i indexes a house in zip code z at quarter t of purchase, and t' is the quarter of sale. $\text{Gross Return}_{iztt'}^{Ann}$ is the annualized return on the given transaction, as defined in Equation (9). \hat{e}_{izt}^2 is the predicted pricing error normalized by the standard deviation of the pricing errors, and $\widehat{\text{Sells Within 90 Days}}_{izt}$ is the predicted probability of a house selling within 90 days of listing. As before, we control for quarter x zip fixed effects μ_{zt} for month of purchase. Our specification therefore compares how the return realized by iBuyers varies with our measures of house's ease of pricing and liquidity as compared to non-iBuyer transactions for similar houses purchased within the same zip code and a point of time.

We include all transactions when estimating the above specification, including non-iBuyer transactions. We then compare how iBuyer returns differ from non-iBuyer systematically with house pricing error and liquidity. It is important to not consider only iBuyer transactions in isolation, because there could be persistent differences in realized returns *on average* across houses with high pricing errors or low liquidity. In particular, the coefficients on pricing error and liquidity absorb these differences, and the interactions of these characteristics with $iBuyer_{izt}$ show how iBuyers and non-iBuyers' returns vary with these characteristics. For example, a negative coefficient on iBuyer interacted with pricing error, β_2 , indicates that when transacting in a hard to price home, iBuyer returns are lower than an individual's return would be when transacting in a similarly hard to price home. The differences here therefore highlight how iBuyer and non-iBuyer transaction strategies relate to gross returns.

The results below show that even among the houses that iBuyers chose to buy, their realized gross returns were greater for easier to price homes. The interaction term in Columns (1) and (3) show that compared to non-iBuyers, iBuyers' realized spread is relatively lower on houses with a high expected squared pricing error. Additionally, iBuyers' realized spread is relatively higher on

houses with a higher probability of selling within 90 days, as shown in Columns (2) and (3). These results are consistent with the idea that iBuyers face more adverse selection in houses, which are more difficult to price. Moreover, their lower returns on homes -- that would otherwise take more time to sell -- can reflect their willingness to sell such homes quickly at a reduced price to avoid costs of carrying empty homes for a longer period of time.

Next, we investigate how time to sell relates to our ease-of-pricing and house liquidity measures. Similar to above, we regress:

$$\begin{aligned} \text{Holding Period}_{iztt'} = & \beta_0 \text{iBuyer}_{izt} + \beta_1 \hat{e}_{izt}^2 + \gamma_1 \widehat{\text{Sells Within 90 Days}}_{izt} + \\ & \text{iBuyer}_{izt} \times (\beta_2 \hat{e}_{izt}^2 + \gamma_2 \widehat{\text{Sells Within 90 Days}}_{izt}) + \mu_{zt} + \epsilon_{izt} \end{aligned}$$

An observation is a house sale, where i indexes a house in zipcode z at quarter t of purchase, and t' is the quarter of sale. *Holding Period* is the time the house remains in inventory, expressed in years. As before, β_1 and γ_1 capture how holding periods differ systematically between easy- and hard-to-price homes and liquid and illiquid homes, respectively. The coefficients of interest are β_2 and γ_2 , which capture how iBuyers outcomes are different from typical sellers along these dimensions. μ_{zt} is a vector of zip-quarter fixed effects.

In the Table below, Column (4), shows that harder-to-price homes remain in iBuyer inventory for relatively longer, and these results are robust to the inclusion of controlling for the liquidity of the house in Column (6). The interaction term on house liquidity is negative, as expected, but not statistically significant either by itself in Column (5) or including the interaction term with pricing error in Column (6). To summarize, when iBuyers purchase houses, which are difficult to price with simple hedonics, they earn lower spreads, and realize higher cost of carrying the house.

iBuyer PnL in Easy-to-Price and Liquid Homes

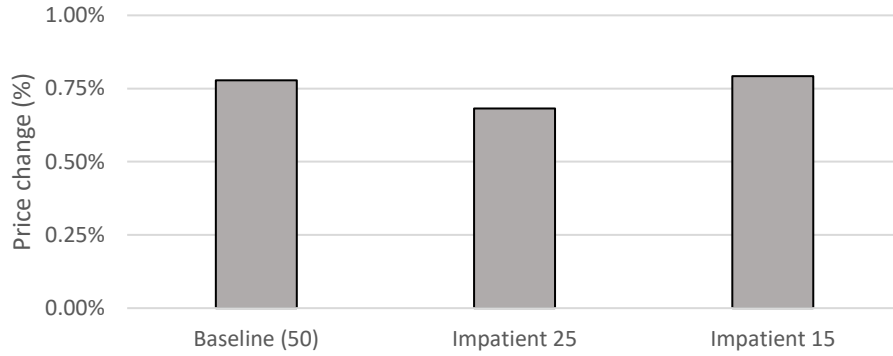
This table shows how iBuyer market penetration and gross returns relate to ease-of-pricing and liquidity measures using Corelogic transaction data from 2013-2018. Panel B shows how the realized gross return and holding period varies by ex-ante pricing error and liquidity among iBuyers and other sellers. The regression includes all transactions where the property is bought and sold within two years. Columns (1)-(3) use realized gross return (annualized, in percentage terms) as the left-hand side variable. Columns (4)-(6) use holding period (in years) as the left-hand side variable. All columns include zip-quarter fixed effects. Standard errors, clustered at the property level, are shown in parentheses.

iBuyer Realized Annualized Gross Return and Ease-of-Pricing and Liquidity Measures

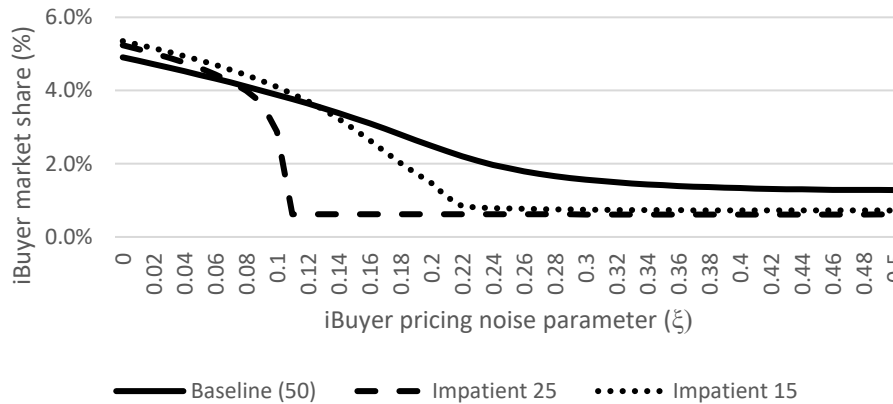
	<i>Dependent variable:</i>					
	<i>Gross Return^{Ann}</i>			<i>Holding period (years)</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
iBuyer	0.281 (0.022)	0.083 (0.066)	0.213 (0.070)	-0.422 (0.017)	-0.382 (0.055)	-0.466 (0.059)
$\hat{\epsilon}^2$	0.799 (0.071)	- (0.121)	0.856 (0.056)	-0.384 (0.056)	- (0.094)	-0.353 (0.094)
<i>SellsWithin90Days</i>	- (0.047)	0.145 (0.047)	0.211 (0.047)	- (0.047)	-0.109 (0.047)	-0.136 (0.048)
iBuyer x $\hat{\epsilon}^2$	-1.105 (0.310)	- (0.504)	-1.573 (0.246)	1.033 (0.246)	- (0.415)	1.211 (0.415)
iBuyer x <i>SellsWithin90Days</i>	- (0.115)	0.314 (0.115)	0.278 (0.115)	- (0.096)	-0.051 (0.096)	-0.047 (0.096)
Zip x Quarter FE	Y	Y	Y	Y	Y	Y
Observations	46,746	19,675	19,675	46,746	19,675	19,675
R ²	0.213	0.263	0.267	0.240	0.301	0.302

Appendix A.10: Model Robustness to Impatient Share

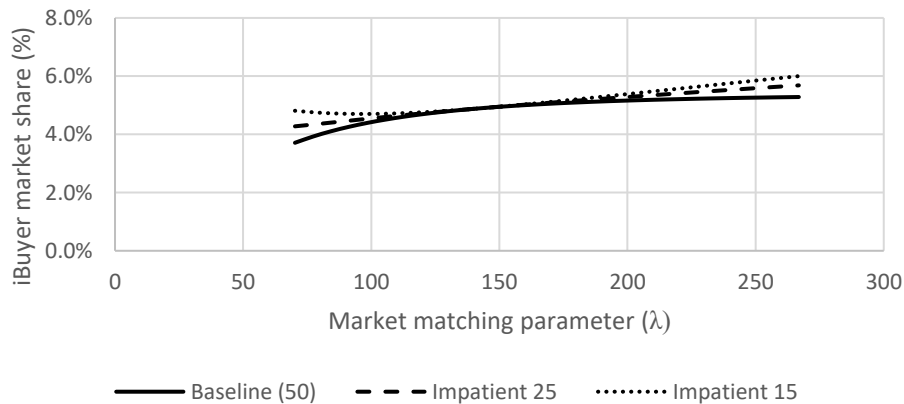
In this section we examine the model's robustness to alternate assumptions about impatient seller share. We compare select counterfactual outcomes for the base model (50% impatient) to two alternative assumptions, (25% and 15%). For the alternate assumptions, we re-estimate the model parameters and provide the resulting counterfactuals. Panel (a) shows equilibrium house price changes after iBuyer entry for the baseline model, and the 25% and 15% impatient models, respectively. Panel (b) shows how iBuyer market share varies with iBuyer pricing error. Panel (c) shows how iBuyer market share varies with market liquidity. In Panels (b) and (c), the solid line is the baseline model, the dashed line is the 25% impatient model, and the dotted line is the 15% impatient model.



(a) Price impact of iBuyer entry



(b) iBuyer market share and signal noise



(c) iBuyer share and market liquidity

Appendix A.11: Model Validation

We undertake two model validation exercises: First, we examine the equilibrium impact on prices and liquidity following iBuyer entry. Second, at the individual level, we explore why consumers transacting with iBuyers may value selling their home quickly.

Equilibrium Evidence

We examine the elasticities of price and liquidity to iBuyer entry using a difference-in-difference style analysis. We utilize the following empirical design to generate elasticities that map to the model-implied elasticities reported above:

$$Liquidity_{zt} = \beta iBuyerShare_z \times Post_t + \gamma_z + \gamma_t + \epsilon_{zy}$$

$$\log(Price)_{izt} = \beta iBuyerShare_z \times Post_t + X'_{izt}\boldsymbol{\Gamma} + \gamma_i + \gamma_t^i + \gamma_z + \epsilon_{izt}$$

$Liquidity_{zt}$ is the fraction of listings in zip z at year t selling within two weeks. $\log(Price)_{izt}$ is the log sale price of property i in zip z at time t . γ_* are fixed effects at the zip, time, property, or ownership tenure (τ)³⁶ level. X'_{izt} is time-varying characteristics of the house (in particular, age fixed effects). $iBuyerShare_z$ is the iBuyer Market share in 2018, $Post_t$ is an indicator for post 2018. Since iBuyer entry may be endogenously correlated with changes in time on the market—indeed, our model suggests that iBuyer is related to liquidity of the market -- 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 2018:

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

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

$$\% \widehat{iBuyer}_z = \frac{1}{N_z} \sum_{i \in z; t \in (2011-2014)} \widehat{iBuyer}_{izt}$$

As before, \widehat{iBuyer}_{izt} 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 2013. We use this measure to instrument for $iBuyer Share_z$. Thus, our instrument for iBuyer market share in 2018 is the

³⁶ That is, for the property-level regression, the number of years between the current sale and previous sale to account for house wear-and-tear or renovations.

predicted iBuyer share based on the physical homes transacted between 2011 and 2014. Figure A.11.1 shows the empirical first stage relationship.

The first stage effect is very strong: A 1% increase in predicted share based on the physical characteristics of houses transacting in 2011-2013 is associated with a 0.964% increase in actual iBuyer market share in 2018 (this regression omits Phoenix, which is used to fit the hedonic model) (shown in Table A.11.2 Column (1)). Columns (2)-(4) show the results for log prices, and Columns (5)-(7) show the results for the fraction of listings sold within two weeks. Columns (2) and (5) are the OLS, (3) and (6) are the reduced form, and (4) and (7) are the IV estimates. Broadly, we find that a 1% increase in iBuyer share is associated with a 1.4% increase in prices (Column (4)) and a 1.6% increase in the fraction of listings sold within two weeks (Column (5)).

Individual-level evidence

We next explore why consumers transacting with iBuyers may value selling their home quickly. One potential reason could be they want to move to a new location, either because they found a new job, or they want to move closer to family. Alternatively, they might live in a house that is too large, and may want to downsize. We now investigate if this is the case in the data.

We follow a panel of individuals' homeownership records through time, and test whether their behavior in terms of sales, moving, house size, and leverage varies following the entry of iBuyers. One approach would be to document changes in these outcomes for households selling to iBuyers relative to other households. A potential concern with that approach is that iBuyer may not facilitate moving to a different location or downsizing. Instead, the same characteristics that drive the household preference for speed are correlated with their preference to move or downsize.

We employ a difference in difference style analysis to address this concern. The event is the entry of iBuyer. We define treatment and control in terms of whether the individual's home is the "type" that an iBuyer would target for purchase. As discussed in Section IV and illustrated again below, iBuyers focus on a predictable subset of homes based on their physical characteristics. This allows us to create control and treatment groups. The treatment group is individuals living in homes that are similar to those typically targeted by iBuyers – the notion is that following iBuyer entry it should be easier for them to sell their homes. Similarly, the control group is individuals living in homes that are not similar to those typically targeted by iBuyers – the notion is that they are unlikely to be directly affected by iBuyer entry. We then evaluate how outcome variables of interest evolve in the two groups following iBuyer entry, using data between 2013 and 2017:³⁷

$$Outcome_{izt} = \beta i\widehat{Buyer}_{izt} \times Post_t + H'_i \mathbf{B} + \mu_{zt} + \epsilon_{izt}$$

³⁷ We end the data in 2017 rather than 2018 so that movers have one year to relocate before the end of our dataset.

Here i indexes an individual in zip code z at quarter t . $Outcome_{izt}$ is an outcome variable of interest. We study three outcomes: whether the individual sells their house, whether the individual moves to a different location, defined as moving to a new MSA relative to the prior house, or downsizes, defined as moving to a house with a lower effective price. $iBuyer_{izt}$ is an indicator for whether the home is likely to be one that an iBuyer transacts in, which we construct as described below. $Post_t$ is an indicator variable that takes a value of 1 after the entry of iBuyer, which we define as 2015. H_i is a vector of house characteristics such as square footage, and μ_{zt} is the zip times quarter (interacted) fixed effect. The identifying variation being used here comparing individuals in iBuyer-targeted homes relative to other individuals in the same zip code and point in time differentially around iBuyer entry.

We construct $iBuyer_{izt}$ as an indicator that effectively sorts individuals into treatment and control groups based on whether they reside in a home that is likely to be targeted by iBuyers. We do this in three steps. First, we estimate the likelihood that a home would be targeted for purchase by an iBuyer using the same method in Section III. As noted there, we do this by estimating specification (3) using 2018 data from Phoenix. This data is then not used in our subsequent analysis. Second, we apply the estimated model to homes in our main regions of analysis, Phoenix, Gwinnet County, Las Vegas, Orlando, and Dallas over the period 2013-2017 to construct a probability that a given home would be targeted by iBuyers for purchase. Finally, we convert (continuous) iBuyer likelihood into a discrete zero-one indicator variable, defining $iBuyer_{izt}$ to be one if house i is predicted to be above median based on the probability that the home would be targeted by iBuyers.

Table A.11.3 presents our main results. We first show that propensity of sale in treatment group relative to control group increases with the entry of iBuyer. The result is not mechanical, because we estimate what homes iBuyer prefer outside of the window of our diff-diff specification. Column (1) shows that the probability of sale of a home that is typically targeted by iBuyers relative to control group increases by roughly 0.5 pp per annum. This is large relative to a mean of 8.3pp. In other words, to the extent our control group is a reasonable comparison group, we can conclude that iBuyers are not simply replacing sales that would have occurred otherwise—they are increasing the rate of sales. We find a strong effect among low LTV individuals -- those with LTVs below the 75th percentile (Column 2). There is no impact among those with high LTVs (Column 3). A possible explanation is that if the need to deleverage compels a household to sell their house -- a more likely scenario with high LTV households -- they will do so whether or not iBuyers enter.

One potential reason why a consumer sells to an iBuyer may be that they want to move to a new location, for example, because they found a new job, or they want to move closer to family. If entry of iBuyer makes it easier to move, it could increase overall mobility. We track individuals in states in which iBuyers enter, and assess whether their propensity of moving out of their market (defined as an MSA) changes. We do this by following individual names and testing whether we

observe a subsequent name match within the same MSA after moving. We find that iBuyers entry increases the mobility of individuals in the treatment group relative to the control group – Column (4) shows that the probability that they leave their market increase by 0.81pp relative to the baseline rate of 21pp. These results are consistent with the idea that entry of iBuyers makes it easier for some individuals to sell their house and relocate.

Finally, unlike individuals who moved to different markets, we now investigate whether iBuyer entry allows some individuals to downsize in the same market. Individuals with houses that are too large or expensive who want to move into a smaller house need to sell their house first. We test whether the presence of iBuyer accelerates this transition. To do this analysis we restrict the sample to individuals who sold their house each year and purchased another house in the same market within the sample period ending in 2018. In doing this we consider the house purchased nearest in time to the sold house. We compare the purchase price of the new home with the original purchase price, adjusted for local market price appreciation since the original purchase. That is, the change in house price is equal to:

$$Price\ Imputed_{it}^{Old} = Purchase\ Price_{it_0}^{Old} \times \frac{PriceIdx_{it}^{New}}{PriceIdx_{it_0}^{Old}}$$

$Price\ Imputed_{it}^{Old}$ represents the estimated value of the sold house i at time t based on local price appreciation in the market between the time of purchase and the time of sale. In particular, $Purchase\ Price_{it_0}^{Old}$ is the original purchase price of property i at original purchase time t_0 . $PriceIdx_{it}^{New}$ is the median sale price in the same county as i at the time of sale t and $PriceIdx_{it_0}^{Old}$ is the median sale price in the same county as i at the original time of purchase t_0 .

The Table below Column (5) shows that, relative to individuals in the control group, those in the treatment group are more likely to move into smaller houses following iBuyer entry. The change in home price is roughly 1.8% lower relative to control group, following iBuyer entry. We also explore whether iBuyers help over-levered individuals to delever. We find essentially no effect: column (6) shows that relative to control group, individuals in treatment group do not tend to increase or decrease their leverage as they transition from their old house to a new one.

Figure A.11.4 shows the timing of these effects. It presents annual differential change in individuals' selling, moving, and downsizing of individuals in treatment group relative to control group, following iBuyer entry in 2015, where Panel (a) corresponds to the probability of selling a home among all owners, (b) among high LTV owners, and (c) among low LTV owners. Panel (d) shows the propensity to remain in the market, (e) shows the change in house prices, and (f) shows the change in LTV. Broadly, these figures show that the timing of these changes is consistent with the timing of iBuyer entry.

Figure A.11.1: 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.

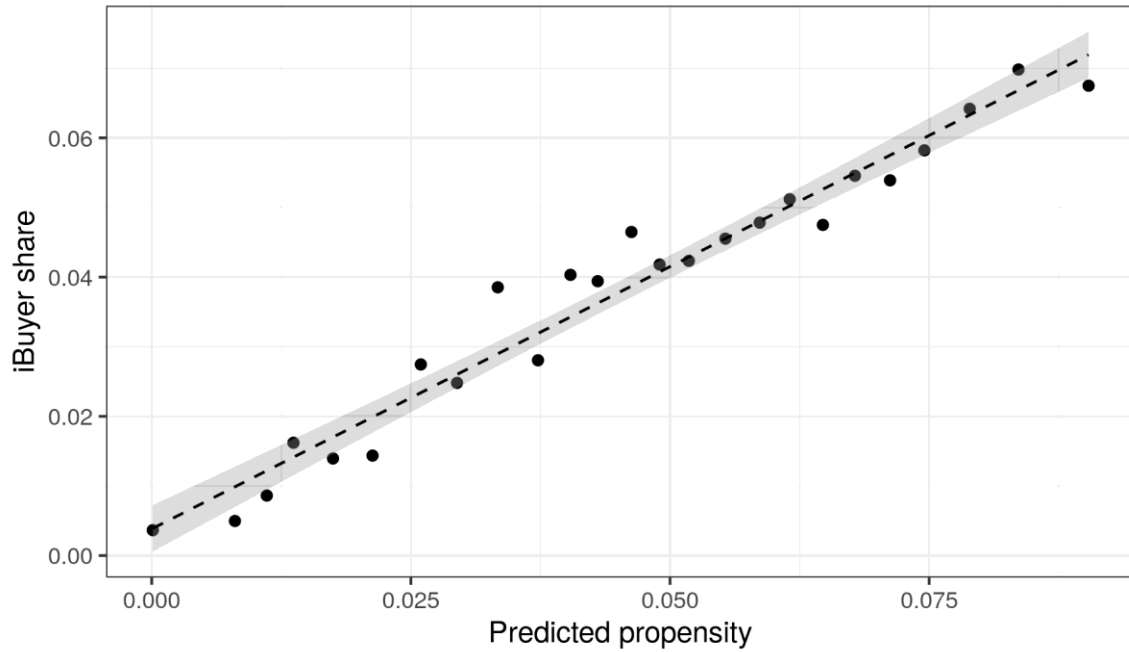


Table A.11.2: Model Validation: iBuyer Entry and Regional Outcome Variables

This table shows the association between iBuyer entry and regional outcome variables using the zip-code level data from deeds records, column (1), house-level transaction data, columns (2)-(4), and zip-year level Redfin data, columns (5)-(7). *iBuyer Share* is the change in iBuyer share in the zip code in 2018. *Log sale amount* is the log of the sale price as recorded in the deeds records. *% Sold in 2 weeks* is the change in houses sold within two weeks of listing and measured at the zip-code level by Redfin. *Propensity* is the instrument used to predict iBuyer market share, calculated by first fitting a hedonic model of iBuyer propensity to purchase a house based on physical characteristics of the house, and then using the model to predict iBuyer market share based on the ex-ante physical characteristics of the housing stock represented by transactions occurring between 2011 and 2013. Column (1) is the first stage regression, regressing the endogenous iBuyer share on the instrument. Columns (2)-(4) examine changes in sale prices for identical homes in areas with high and low iBuyer presence before and after iBuyer entry, excluding homes that iBuyers actually purchase. Columns (5)-(7) examine changes in sale speeds after listing in areas with high and low iBuyer presence. Columns (2) and (5) are the OLS regressions; (3) and (6) are the reduced form regressions using predicted propensity, and (4) and (7) are the IV regressions. Standard errors are shown in parentheses.

Specification	<i>Dependent variable:</i>						
	iBuyer share	log(Sale amount)			% listings sold in two weeks		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	RF	IV	OLS	RF	IV
Propensity	0.964 (0.070)	-	-	-	-	-	-
Post x Share	-	0.451 (0.156)	-	1.405 (0.266)	-0.007 (0.157)	-	1.579 (0.306)
Post x Propensity	-	-	1.824 (0.303)	-	-	1.457 (0.233)	-
Unit of observation	Zip	Property sale	Property sale	Property sale	Zip-Year	Zip-Year	Zip-Year
Zip FE	N	Y	Y	Y	Y	Y	Y
Year FE	N	Y	Y	Y	Y	Y	Y
Property FE	N	Y	Y	Y	N	N	N
House Age & Tenure FE	N	Y	Y	Y	N	N	N
N	348	959,408	959,408	959,408	662	662	662
R ²	0.352	0.946	0.946	0.946	0.680	0.711	0.591

Note:

*p<0.1; **p<0.05; ***p<0.01

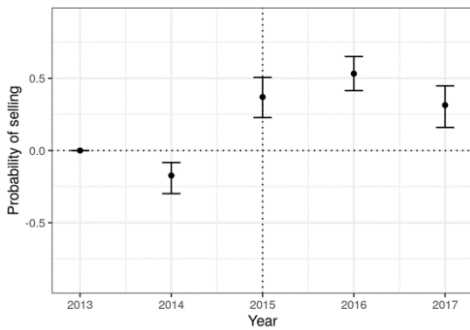
Table A.11.3: Model Validation: iBuyer Entry and Homeowner Mobility

This table shows the impact of iBuyer entry on geographical mobility, home downsizing, and deleveraging of existing homeowners. Data are from Corelogic between 2013 and 2017, two years around iBuyer entry in 2015, at the individual-year level. The outcome variables are as follows. *Sells* is an indicator for whether the individual sells his house in the given year (columns 1-3). *Remains in market* is an indicator for whether the selling individual remains a homeowner in the same market (column 4). *ΔBuy Price (imputed)* is the change (in %) of the current imputed house price to the new house’s price (column 5). *ΔLTV (imputed)* is the change (in %) of the imputed old LTV to the new house’s LTV (column 6). $i\widehat{Buyer}$ is an indicator for whether the individual’s property is in the top 50% of predicted iBuyer shares based on its physical characteristics. *Post* is an indicator for 2015 or later. House controls are those used in Table 2: *Price*, the transaction price in the deeds records, *house age*, the difference between the transaction date and the year of construction, *land square footage*, the tax-assessed property square footage, and *multistory* is an indicator for whether the house has greater than 1 story (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.

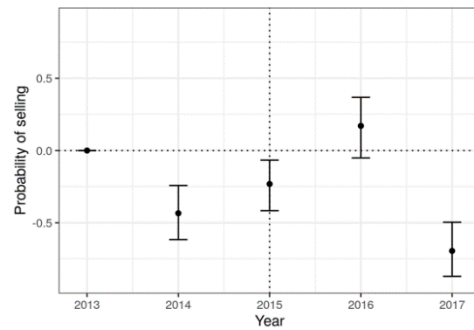
	<i>Dependent variable:</i>					
	Sells			Remains in Market	ΔBuy Price (imputed)	ΔLTV (imputed)
	(1)	(2)	(3)	(4)	(5)	(6)
$i\widehat{Buyer}$	1.677 (0.046)	1.720 (0.055)	1.556 (0.081)	-2.212 (0.243)	-0.543 (0.753)	-0.404 (0.296)
$i\widehat{Buyer} \times Post$	0.490 (0.058)	0.622 (0.069)	-0.032 (0.109)	-0.809 (0.303)	-1.854 (0.919)	-0.572 (0.358)
Sample	All	Low LTV	High LTV	All	All	All
House Controls	Y	Y	Y	Y	Y	Y
Market x Year x Tenure FE	Y	Y	Y	Y	Y	Y
Observations	4,161,751	3,121,313	1,040,438	346,136	73,264	66,563
R ²	0.012	0.010	0.019	0.050	0.028	0.424

Figure A.11.4: Model Validation: iBuyer Entry and Homeowner Mobility

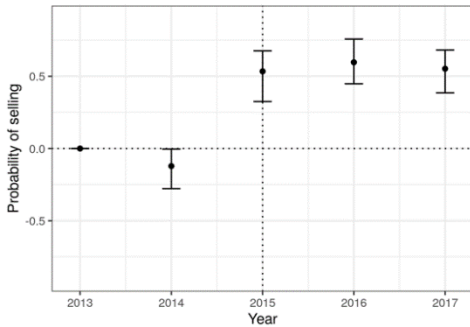
This figure shows the estimated differential change in individuals' selling, moving, and downsizing propensities around the iBuyer entry time of more exposed homeowners to iBuyer entry (treatment) relative to less exposed ones (control). We define treatment and control based on whether the individual's home is the type that an iBuyer would purchase based on the out-of-sample predictive model. The model is based on binned house characteristics, including square footage, age, and price. An iBuyer-type home is then defined as being in the top 50% of predicted iBuyer likelihood. The figures show the estimated coefficient on year times iBuyer-type dummy. The regressions are on the individual-year level. We use Corelogic data from 2013 to 2017, so that sellers in 2017 have one year in the data to find a new house. We plot the estimated differential change along with 95% confidence bounds. Panel (a) considers all individuals; panel (b) considers those with high LTVs (above the 75th percentile relative to other homeowners at origination); panel (c) considers those with low LTVs (below the 75th percentile relative to other homeowners at origination). Panel (d) considers the differential change in the probability of staying within the market, defined as whether the individual purchases another house in the same market; panel (e) considers change in imputed home value from the old house to the new house; panel (f) considers the change in LTV from the old house to the new house.



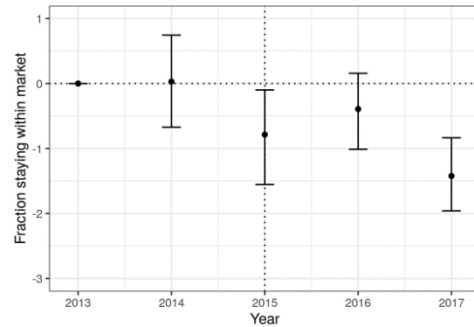
(a) Selling, all iBuyer homes



(b) Selling, high-LTV iBuyer homes



(c) Selling, low-LTV iBuyer homes



(d) Staying in market, all iBuyer homes

Appendix A.12: Speed of Traditional Listing

This figure uses data from Zillow.com and shows the fraction of US metropolitan areas where the median days-to-pending (i.e., the number of days between when a listing appears and when it is marked as a pending sale) is below 30 days. Data are weekly and shown over the entire available date range.

Markets with median days-to-pending < 30

