Housing Wealth and Overpayment: When Money Moves In*

Darren Aiello† Jason Kotter‡ Gregor Schubert§

March 14, 2023

Abstract

We construct a novel dataset tracking households across property purchases covering 25 years of moves within the U.S. We find that information frictions in residential real estate markets cause movers with larger exogenous housing wealth to overpay for their next house, relative to both time varying local prices as well as time invariant characteristics of the property itself. These housing wealth driven overpayments are associated with larger positive price impacts to the surrounding neighborhood and are larger for local movers relative to non-local movers. The aggregate effect of housing wealth inflows is to increase commuting-zone-level house price growth.

Keywords: Household wealth, migration, real estate.

JEL Classification: G51, D83, R31, D10.

*We thank Bronson Argyle, Stuart Gabriel, Mark Garmaise, Andra Ghent, Chris Hair, Barney Hartman-Glaser, Mahyar Kargar, Taylor Nadauld, and seminar participants at BYU and UCLA for comments and suggestions, as well as Michael Flores for his excellent research assistance. We also thank the UCLA Ziman Center's Rosalinde and Arthur Gilbert Program in Real Estate, Finance, and Urban Economics for their generous support. All errors are our own responsibility.

†Brigham Young University; d.a@byu.edu

‡Brigham Young University; jasonkotter@byu.edu

§University of California, Los Angeles; gregor.schubert@anderson.ucla.edu

1 Introduction

The single biggest asset on the balance sheet of most U.S. households is their primary residence (Bhutta et al., 2020). Despite the importance of this asset in a household's portfolio, there is little evidence in the literature regarding how households make housing consumption and investment decisions across properties. Across a large, nationally representative sample of homeowner moves, we document a surprising fact: an increase in housing wealth causes households to overpay for their next house.

Given both the financial costs (overpayment leads to lower future realized returns on the property) and the opportunity costs (buying a larger house or saving the extra money) of this behavior, why do households spend housing wealth on overpayment? Kurlat and Stroebel (2015) show that individuals with greater neighborhood familiarity have higher equity gains on their real estate transactions, consistent with a trade-off between information and overpayment. Because housing markets are segmented and illiquid, information acquisition requires substantial time and effort. We demonstrate that households substitute away from this costly information acquisition by using their sold home equity gains to overpay on their next purchase.

What effect does this have on the household's new neighbors? First, a purchase by a household with a larger equity gain causes increased prices for properties in the nearby neighborhood. Second, we show that the aggregate effect of housing wealth inflows increases commuting-zone-level house prices. Consequently, increases in housing wealth not only have important effects on individual household investment decisions, but also influence housing market dynamics in the purchase location.

We begin by constructing a novel dataset of household moves based on the Zillow Transactions and Assessments (ZTRAX) data. We focus on moves between owner occupied single-family residential properties between 1996 and 2021 where we observe three specific transactions—the

purchase and sale of a particular property (the "sold property"), as well as the purchase of the subsequent property (the "purchased property"). This allows us to measure, over the individual household's holding period, the increase in the value of their housing asset (their "equity gain"[1]) and relate it to observed behavior in their subsequent housing purchase. For that subsequent purchase, we recover the residual (the "price premium") from a repeat sales regression model that controls for both time invariant observable and unobservable property characteristics as well as time-varying zip code and census tract average price levels, measured with respect to the entire universe of housing transactions in the ZTRAX data. This price premium only contains information about the purchase price that is orthogonal to a reasonably objective measure of the quality and neighborhood trend adjusted value of the asset.

Figure 1 demonstrates the key relationship that we are investigating in this paper: there is a strong positive correlation between the equity gain a household realized on their sold property and the price premium observed on their purchased property. However, assessing the causal effects of a household's equity gain is challenging for two reasons. First, households that realize a large equity gain are likely to be more sophisticated or skilled (or have hired better real estate agents) and therefore less likely to overpay for their new house. The existence of sophisticated housing market participants biases downward any naïve estimate of the effect of equity gain on the price premium. Second, potential co-movement in housing prices can bias upward any estimates of the effect of equity gain at sale on outcomes related to the price paid at purchase. Ultimately, while our measure of the price premium is orthogonal to most local price movements of concern, this will make it difficult to identify the spillover effect of equity gains on neighborhood prices.

[1]We measure equity gain as the difference between the price the property sold for and the original purchase price. This abstracts away from household financing decisions, and represents the total change in housing asset wealth over the period of ownership.

We address these concerns in a three-fold manner. First, we use observable characteristics to control for differences in the types of households by including a control for the median house price of the zip code of the sold property (but at the time that property was originally purchased), a fixed effect for the number of years lived in the house, and interacted fixed effects for the characteristics of the sold property. Second, we include zip-year fixed effects for both the sale and purchase locations, allowing us to control for the timing of the move (at the year level) as well as current zip code level price conditions in both the sale and purchase locations separately. Finally, we instrument for a household's equity gain using the change in the zip code level house price index at the sale location over the household's holding period, exploiting variation in equity gain that is orthogonal to any individual household's (or their agent's) housing market sophistication or bargaining ability. This holding period change in the zip code level house price index strongly predicts a household's equity gain; first-stage F-statistics are well above 100. The exclusion restriction is that, conditional on the fixed effects and controls, the change in housing prices in the zip code of the sold property is uncorrelated with the amount the buyer pays above the current quality and neighborhood trend adjusted value of the property (measured relative to the purchase zip-month and census tract-year) except through the household's equity gain.

Under this set-up, the primary remaining concerns relate to the effect of unobservable property-specific or highly localized economic shocks on transaction prices. For example, if average price growth in the Beverly Hills, CA, zip code (90210) where the household is selling their house is correlated with prices in the purchase neighborhood because of some particular affinity (either it is immediately bordering or there is significant co-movement between "high-end" locations) then our estimates could still be biased. But this purchase neighborhood of concern has to both be small in size and unique relative to its surrounding area—its prices have to move differently from the average property in it's zip-month and census tract-year. Moreover, it is possible that

price premia are driven by increases in house quality due to recent property renovations that are unobservable in our repeat sales model. To alleviate these concerns, we show that a larger price premium negatively predicts future annualized realized returns on the purchased property, which is inconsistent with persistent changes in the valuation of the purchased property. Moreover, we find that this decrease in future annualized returns associated with larger price premia is in fact stronger with shorter holding periods, and largest in cases where the homeowner sells the newly purchased property within a year. Because the value of both recent renovations and micro-neighborhood price trends is unlikely to substantially depreciate within a year, the price premium is unlikely to represent payment for these unobservable characteristics. Instead, we view a larger price premium as consistent with "overpayment".

Within this instrumented framework, we find that for every dollar of exogenous equity gain that a seller receives, they overpay for their next house by 7.9 cents. For the average equity gain in our sample, that represents an overpayment of about 2 percent of the overall purchase price. What leads buyers to use their equity gains to overpay? First, to the extent that equity gains relieve capital constraints, buyers might be able to consider a larger set of houses, potentially leading to a better match. Consequently, the price premium we observe might be driven by buyers using their equity gains to buy a house for which they have a higher private valuation (and thus are more willing to overpay). Second, because acquiring information about neighborhood characteristics in order to become informed about the quality and neighborhood trend adjusted value of a property is costly, buyers with large equity gains might trade off expending that effort with simply overpaying. Consistent with both of these possibilities, Gargano, Giacoletti, and Jarnecic (2020) show that local area price appreciation causes capital constrained homeowners of that area to search across a broader set of prospective properties (potentially leading to a better match), while not changing the amount of attention devoted to any particular listing and spending

less time overall searching (consistent with less information acquisition activity).

We distinguish between these two channels by investigating how the sensitivity of the price premium to equity gains varies based on the characteristics of the local market as well as those of the moving household. First, a high volume of historical transactions over the previous 90 days in a very local geography (1/2 mile radius around the purchase property) represents both a deep pool of properties over which to search as well as an abundance of information about recent comparables. Consistent with these high volume areas having less information asymmetry, but inconsistent with the broader potential search scope leading to better matches for a given effort, we find that the sensitivity of the price premium to equity gain is markedly reduced. Second, for the subset of within-county moves ("local moves"), households that have lived in their sold home for a longer period of time are both more likely to be familiar with the area—thus facing a bigger information advantage relative to their less-tenured neighbors—and to have more strongly developed local housing preferences, perhaps due to relatively older children or more specific community ties. Consistent with more tenured households having relative information advantages, but inconsistent with them having sharper preferences (and thus higher private valuations), we find that their sensitivity of the price premium to equity gain is also reduced. Finally, we find that the sensitivity of the price premium to equity gains is increasing in historical price dispersion, a measure of the quality of information from comparables, providing additional evidence consistent with the information asymmetry channel. Consequently, the evidence suggests that equity gains allow movers to substitute overpayment for costly information acquisition.

One specific version of avoiding costly information gathering is benchmarking—a household might base their offer price on their purchased property on the sale price of their sold property. We find evidence that benchmarking exists in the data. Households that move from high price areas to lower price areas overpay more on average. However, the effect of equity gains on

overpayment persists even after controlling for benchmarking.

Having established that equity gains cause overpayment, we next examine what happens to neighborhood house prices when a household with large equity gains moves in. We show that the equity gain a buyer realized on their recently sold home causes an increase in average housing prices over the next six months in the nearby neighborhood—8.9 cents for every dollar of equity gain. This spillover effect of equity gain varies in the same manner with market and buyer conditions (transaction volume and years in home) as our prior results. This suggests that the spillover effect is also likely a result of the same substitution between equity gain and price discovery.

Surprisingly, while buyers from further distances overpay more on average, their overpayment is markedly less sensitive to their equity gains. The spillover of equity gains onto neighborhood prices is 20% higher for local movers relative to non-local movers. Non-local movers may face limited access to information acquisition technologies (repeated forays to open houses and multiple extended car tours of target neighborhoods) and consequently consistently overpay, irrespective of their equity gain. In contrast, local movers with low equity gains overpay the least—engaging in "shoe leather" information acquisition—whilst local movers with high equity gains overpay the most.

To more fully characterize how housing wealth affects buyer behavior, we also estimate the effect of equity gains on other dimensions of the purchase decision. Buyers with larger equity gains spend more on a house, both because they purchase a property in a more expensive zip code and because they buy a bigger house. We find that for each dollar of equity gain, households spend $0.87 more on their next house. Of that, $0.79 represents a house with a higher quality and neighborhood trend adjusted price, and the remaining $0.08 is spent on overpayment.

Finally, we examine the extent to which these individual spillover effects explain commuting-

zone-level house price growth. For this analysis, we focus on the effect of aggregate out-of-area equity gain inflows on destination commuting zone house price growth. Estimating the causal impact is challenging because households tend to move more between areas with highly correlated house prices (Sinai and Souleles, 2013). As a result, we need to identify exogenous inflows of equity gains. We do this using predicted equity gain inflows based on historical migration routes calculated using IRS data, similar to Schubert (2021). We show that the cumulative effect of equity gain inflows is to drive up local house price growth, with 10% higher equity gains among incoming movers associated with an increase in local house price growth of ~ 0.4 percentage points.

Our paper contributes to a large literature exploring the existence of information asymmetries in real estate markets.[2] Across both commercial and residential real estate, as well as mortgage originations, less informed buyers pay a premium (Chinco and Mayer, 2016; Lambson, McQueen, and Slade, 2004a; Agarwal, Sing, and Wang, 2018; Bhutta, Fuster, and Hizmo, 2019; Cvijanović and Spaenjers, 2021). We expand on these results by showing that households, faced with information asymmetries, use their housing wealth to trade off the costs of becoming informed with overpaying for their next property. This is broadly consistent with Caplin and Dean (2015) who argue that choice "mistakes" are actually a rationale response to costly information acquisition.

We also contribute to the literature exploring migration and house price dynamics (Jia, Molloy, Smith, and Wozniak, 2023). Howard (2020) shows that domestic migration flows predicted by historical links can impact local housing market dynamics, while Badarinza and Ramadorai (2018), Gorback and Keys (2020), Li, Shen, and Zhang (2020), and Davids (2020) show that capital inflows related to foreign ownership drive up local house prices. We differ from this literature by focusing on the *financial characteristics* of movers within the U.S., and show that the house

[2]See, e.g., Garmaise and Moskowitz (2004); Kurlat and Stroebel (2015); Rutherford, Springer, and Yavas (2005); Levitt and Syverson (2008).

price impact of migration varies with the equity gains of migrants, consistent with the model presented in Favilukis and Van Nieuwerburgh (2021).

Finally, our paper's ability to track households across multiple home purchases allows us to contribute relative to a literature exploring the role and behavior of buyers (Gargano, Giacoletti, and Jarnecic, 2020; Reher and Valkanov, 2021; Han and Hong, 2022), sellers (Guren, 2018; Andersen, Badarinza, Liu, Marx, and Ramadorai, 2022; Fu, Jin, and Liu, 2022; Giacoletti and Parsons, 2021), their agents (Aiello, Garmaise, and Nadauld, 2022), and participants' overall performance (Wolff, 2022) in residential real estate.

The remainder of the paper is organized as follows. Section 2 describes our data and the matching process that generated it, as well as descriptions of our main variables of interest. Section 3 introduces a conceptual framework to help contextualize our empirical results. Section 4 explains our identification strategy. Section 5 presents our estimates of the effect of equity gain on housing consumption decisions, as well as the spillover of those decisions to neighborhood prices. Section 6 discusses our analysis of aggregate equity gain flows and their impact on county-level housing prices. Section 7 concludes.

2 Data and Matching

To estimate the effect of equity gains on a subsequent price premium, we first construct a novel dataset of household moves based on the Zillow Transactions and Assessments (ZTRAX) data. The ZTRAX data is based on property deeds and covers the entire U.S. We begin with the ZTRAX universe of 220 million house transactions that occur between 1996 and 2021. We then limit the sample to owner-occupied single family residences to filter out investors as well as vacation properties. This results in a set of 136 million house transactions that we search across

to find potential moves.

Within this subset of transactions, we identify potential moves by matching based on names and addresses listed on the recorded deeds. To qualify as a move in our data, the purchase transaction needs to occur less than 274 days after, but no more than 182 days before, the sale transaction.[3] If there are multiple potential matches within that date range, we privilege the strongest name match available (e.g., matches based on multiple names listed on both deeds), only retaining matches that have unique "strongest" match pairs. This matching process results in a set of 19 million moves.

Because we are interested in the effects of equity gain on a subsequent price premium, we limit our sample to transaction series where we observe the three relevant consecutive prices—the original purchase and subsequent sale of the sold property, as well as the next purchased property—and are able to calculate our residualized measure of a price premium for the purchase transaction. This excludes moves to or from non-disclosure states.[4] We further limit the sample to transactions where all three prices are greater than $30,000 and less than $2 million, occur on a property that was never subject to a foreclosure, and where we are able to calculate all relevant inputs into our regression model. This results in a final sample of around 3.2 million moves. After excluding fixed effect singletons, our final regression sample consists of 3.1 million moves.

Figure 2 shows a map of the counties in our data where households purchase homes. Importantly, our sample covers a broad cross-section of the U.S. Table 1 presents summary statistics of the variables used in our regression sample.

[3]Deeds are recorded in a manner that bunches on particular days of the week. These cut-offs are chosen to represent round week day counts for three-quarters and one-half of a year, respectively.

[4]Non-disclosure states are Idaho, Kansas, Mississippi, Montana, New Mexico, Texas, Utah, Wyoming, and all counties in Missouri except for Jackson County, St. Charles County, St. Louis County, and St. Louis city. We observe prices for a very small set of transactions in non-disclosure states. These prices might be erroneous, or the seller might have voluntarily disclosed the price. For sample consistency, we drop these observations from our sample. However, the results are not sensitive to the decision of whether or not to include these observations.

2.1 Sample Selection Concerns

Our data is a sample of home owner to home owner moves. As a result, we are missing first-time home owners and any owners that are transitioning either to or from the rental market. Furthermore, our sample is conditioned on the decision to move. It is possible that potential equity gains influence the decision to move; thus, we are cautious about extrapolating our results to households that haven't yet made the decision to move. We interpret our results as the effect of exogenous equity dollars on a household's next purchase conditional on moving.

Conditional on moving, our sample of household moves seems broadly representative. ZTRAX has near universal coverage of housing transactions, and we successfully match around 14% of these transactions to moves. Aggregate flows across counties through time in our sample are broadly similar to aggregate flows calculated on the universe of moves with the IRS data.

2.2 Definition of Equity Gain

Our main explanatory variable of interest is a household's equity gain on their sold property, which we define as the change in the value of the house over the length of ownership, i.e.,

$$Equity\ Gain_{i,t} = P^S_{i,t} - P^S_{i,t-tenure} \tag{1}$$

where $P^S_{i,t-tenure}$ is the original purchase price that household i paid for the property being sold, and $P^S_{i,t}$ is the price that the property sells for at time t, the time of the move.

This measure of equity gain represents the change in the house's asset value, and is broadly a measure of the change in the household's housing wealth. Our definition of equity gain does not account for a household's financing decisions; however, conceptually the change in housing wealth is independent of financing choices. While household leverage decisions might change

the timing of when housing wealth is liquidated (e.g., cash-out refinances or home equity lines of credit), the total change in wealth over the period that the household owns the house are captured in our measure of equity gain.

In our sample, the average household lives in a property for 6.6 years before selling and realizes an average equity gain of $86,244.

2.3 Definition of Price Premium

We define the price premium as the price residual from a repeat-sales model. Specifically, we estimate:

$$\begin{aligned} P_{i,t} &= \alpha_i + \alpha_{z,m} + \alpha_{n,y} + \varepsilon_{i,t}^{Price} \\ PricePremium_{i,t} &= \varepsilon_{i,t}^{Price} \end{aligned} \tag{2}$$

where $P_{i,t}$ is the price of property i purchased at time t, α_i is a property-level fixed effect, $\alpha_{z,m}$ is a fixed effect for the zip code by month of the property being purchased, and $\alpha_{n,y}$ is a fixed effect for the census tract by year of the property being purchased. We refer to the residual from this model, $\varepsilon_{i,t}^{Price}$, as the price premium.

This measure of the price premium accounts for any time-invariant differences in characteristics between properties though the property-level fixed effects and captures any changes in property value arising from neighborhood-level trends through zip-month and census tract-year fixed effects. Consequently, we interpret the price premium as the amount paid in excess of a reasonably objective measure of the current quality and neighborhood trend adjusted value of the property. While our data do not allow us to observe changes in property characteristics over time (such as renovations or additions), we provide evidence in Section 5.1 to suggest that changes in property characteristics are unlikely to explain our estimates. The median price premium in our

sample is about $4,300, or about 1.5% of the purchase price. However, price premia exhibit substantial variation, with the interquartile range extending from about -$17,500 to $31,200 (or -6.4% to 9.7% of the purchase price).

3 Conceptual framework

In this section, we develop a partial equilibrium framework based on the sequential search approach in Turnbull and Sirmans (1993) and Lambson, McQueen, and Slade (2004b) to motivate the empirical analysis that follows and provide one possible interpretation of our results. We modify prior versions of these models by showing how the effect of search costs on price premia paid by heterogeneous buyers would vary with different information environments.

Sellers. We focus on the buyer's search problem and therefore abstract away from the bargaining process between buyers and sellers. We assume that sellers differ randomly in the price premium (or discount) that they charge relative to the prevailing neighborhood price for a house of their type. From the perspective of the buyer, this means that each seller sets a price p for a given house type that is distributed with cumulative distribution function F(p). Moreover, these are "transaction" prices rather than listing prices, such that buyers can choose to accept or reject a seller's posted price when they meet and the seller always transacts at that price.

Buyers. Buyers i enter the housing market looking to buy housing of a particular type, which has expected value V to the buyer (which may, for example, reflect expected future resale value, as well as the monetary equivalent of the housing consumption). For simplicity, we assume that they are risk-neutral and have zero discount rates. However, each round of housing search incurs a cost c that can vary across buyers. These costs may be comprised of a variety of components, including travel to a neighborhood, physical inspection, attending open houses, and due dili-

gence on a property. Moreover, as is discussed in later sections, it may also include acquiring information about the character of the local neighborhood and the pricing of comparable recent transactions. These costs may vary across individuals—for example, out-of-state buyers may incur higher costs due to the time and expense of traveling, a lack of local networks from which to source information, and lower awareness of which properties are on the market.

After paying this search cost, buyers randomly draw a seller of a house with their desired characteristics. They can either accept (i.e. "buy") at the seller's price p, or reject it (without being able to return to it later) and continue searching, paying the search cost again at the beginning of the next round. If a match is successful and the buyer accepts, they end up with a surplus of $(V - p)$.

Optimal search. Buyers continue searching until they encounter a seller who offers a price that no longer makes it worthwhile to continue. The optimal stopping rule is characterized by a reservation price p^*, such that the buyer accepts—buys the house—for any offered price $p \leq p^*$. Following Lambson, McQueen, and Slade (2004b), the optimal reservation price is set such that the benefit of immediately accepting and obtaining surplus $(V - p^*)$ needs to be equal to the expected benefit of continuing:

$$V - p^* = (1 - F(p^*))(V - p^*) + \int_0^{p^*} (V - p) f(p) dp - c,$$

Here, the terms on the right-hand-side correspond to the value of not finding an acceptable price in the next round, the expected surplus from finding a seller with a price below the reservation price, and the cost of continuing the search. Rearranging, this simplifies to

$$c = \int_0^{p^*} (p^* - p) f(p) dp, \tag{3}$$

which implicitly characterizes the optimal reservation price p^*: the buyer sets the reservation price and continues search in such a way that the expected marginal benefits of an additional search round are equal to the marginal costs. Note that this means that higher search cost buyers set higher reservation prices, are more likely to encounter a property at a price that they are willing to accept, and will therefore on average pay a higher premium for a given house type—an "overpayment" relative to the quality and neighborhood trend adjusted value of the property.

Information costs. Given our focus on how overpayment varies with buyer characteristics—in particular, buyers' housing equity gain and the information environment—we want to relate search costs to potential factors that change the cost of acquiring additional housing market information. In anticipation of our later discussion, we parameterize marginal search costs as $c = \frac{h^{-\theta}}{\tau}$, where h is the amount of time that a household can devote to the housing search, τ is the level of the buyer's existing knowledge about a neighborhood (e.g. due to "tenure" in the area), and θ is the effectiveness of search, i.e. the ability to convert search time into matches, which may depend, for instance, on the distance to the housing market of interest.

As previous studies have argued (Aguiar, Hurst, and Karabarbounis (2013); Huo and Ríos-Rull (2015); Stroebel and Vavra (2019)), higher wealth is likely associated with less time spent searching for lower prices, which we capture here by assuming that the time available for housing search is a decreasing function of a buyer's housing equity gains e, such that $h = e^{-1}$, and

$$c = \frac{e^{\theta}}{\tau}. \tag{4}$$

Comparative statics. Implicitly differentiating equation 3, we can find the change in the

reservation price of the buyer with regard to overall search costs:

$$\frac{dp^*}{dc} = \frac{1}{F(p^*)} > 0$$

That is, reservation prices, and therefore also expected price premia, are on average higher for high search cost buyers. This suggests, for instance, that out-of-town buyers should *on average* pay more for a given housing type if they experience higher search costs, which is supported by empirical findings in the literature (e.g. Chinco and Mayer (2016); Lambson, McQueen, and Slade (2004b)). We are interested in relating the reservation price—and thus the price premium paid—to equity gains of the buyer. Differentiating the reservation price with regard to the equity gain component of search costs, we find that

$$\frac{dp^*}{de} = \frac{1}{F(p^*)} \frac{\theta e^{\theta-1}}{\tau} > 0 \tag{5}$$

$$\frac{dp^*}{ded\tau} = -\frac{1}{F(p^*)} \frac{\theta e^{\theta-1}}{\tau^2} < 0 \tag{6}$$

$$\frac{dp^*}{ded\theta} = \frac{(1+\theta)}{F(p^*)} \frac{e^{\theta-1}}{\tau} > 0. \tag{7}$$

The first line shows that higher housing equity gains are associated with higher overpayment in this framework because search costs increase with housing wealth. In other words, buyers use housing equity wealth to substitute paying additional search costs with simply paying a higher price premium.

Next, we consider how this overpayment effect varies with the information environment and find that the effect of equity gains on the price premium decreases with a greater baseline level of information about the neighborhood, for instance due to better comps or longer tenure in the area. Intuitively, the overpayment effect of housing equity gains comes from the *incremental*

option to either use additional search time to find a better price or forego such an effort. If a buyer already has good information, the potential impact of such incremental efforts is limited, and so equity gains have a smaller effect on the premium paid.

The third line shows how the effect of equity gains varies with the effectiveness of search efforts: the more efficiently time spent searching translates into better information, the higher the effect of housing equity gains on overpayment. The intuition is that as housing equity gains decrease the time devoted to searching, the resulting disadvantage in terms of the price premium paid will depend on how much information is foregone by not searching. Thus, buyers for whom the efficiency of search time is low, for instance because they live far away from the destination housing market and are not well-connected to local agents, should see a smaller impact of high equity gains on the price premium that they pay.

The following sections provide empirical evidence in line with this framework.

4 Empirical Strategy

We are interested in estimating the effect that the equity gain, which a household realized on their sold property, has on the extent to which they pay more (on a quality and neighborhood trend adjusted basis) on their purchased property. Figure 1 shows that equity gains are positively correlated with a price premium. In this section, we discuss the potential concerns with interpreting this correlation, and introduce a strategy to identify the causal effect of equity gains on the price premium.

4.1 Identification Concerns

We investigate the correlation visible in Figure 1 by estimating regression models of the following form:

$$PricePremium_{i,t} = \beta\, Equity\ Gain_{i,t} + \phi' \Gamma_{i,t} + \varepsilon_{i,t} \tag{8}$$

where $PricePremium_{i,t}$ is the dollar amount the household pays on their next purchased home relative to a quality and neighborhood trend adjusted price (defined in Equation 2), $Equity\ Gain_{i,t}$ is the dollar change in house price the household experienced over the years lived in their sold home (defined in Equation 1), and $\Gamma_{i,t}$ represents control variables at the property or area level and any included fixed effects. All of our property-level regressions include a fixed effect $\alpha^P_{z,y}$ for the zip-year of the purchase, in addition to the other controls and fixed effects described in the next section. This fixed effect controls for housing market conditions at the time and place of the purchase. Throughout the paper, we cluster our standard errors by both year and purchase county separately to account for potential correlation both within and across county-level housing markets.

For Equation 8 to recover the causal effect of equity gain on the price premium, it is necessary that the household's equity gain in their previous home is uncorrelated with any unobserved, price-relevant characteristics of the transaction, conditional on the fixed effects. There are two reasons that this is unlikely to be the case. First, due to differences in experience, bargaining ability, and the quality of real estate agents, households are likely to have varying amounts of housing market sophistication. A more sophisticated household is likely to both receive a higher price when selling a house (and thus realize a larger equity gain) and also pay a lower price when buying a house (and thus be less likely to overpay). To the extent housing market sophistication is common, this will bias our OLS estimate downward. Second, Sinai and Souleles (2013) show

that prices tend to be highly correlated across move locations. This implies that households that experience a high equity gain are likely to purchase a home that has a higher price. Because the price premium is orthogonal to zip-month and census-tract year average prices, this correlation is unlikely to bias estimates of the effect of equity gain on the price premium. However, this correlation will make it difficult to identify spillover effects of equity gains to neighborhood prices.

4.2 Identification Strategy

To address concerns that households with high equity gains are more likely to exhibit housing market sophistication, as well as potential concerns over co-movement in house prices across move locations, we take a three-fold approach. First, we include a set of controls/fixed effects to account for potential differences in the types of households moving. While we do not observe individual demographic characteristics in our data, we do observe a rich set of property/transaction characteristics that we use as proxies for household wealth and experience. Specifically, we control for the zip code house price index of the sold property at the time that the household originally bought the home. We further include a set of fixed effects for the property characteristics of the sold property, constructed as the interaction of the square footage of the home (in percentiles), lot size (in deciles), number of bedrooms, number of bathrooms, and age of the home (in 5-year buckets). Finally, we include a fixed effect for the number of years lived in the sold home. Together, this strategy controls for differences in the household wealth (as proxied by the zip-level house price when the household originally bought the sold home as well as the characteristics of the home itself), and differences across long- and short-tenured movers (such as the likelihood of the transaction being an owner-occupied flip).

Second, we include zip-year fixed effects separately for both the sold and purchased property

locations. These control for the timing of the move, and absorb house price dynamics surrounding both the sold and purchased properties.

Finally, we use a two-stage least squares (2SLS) specification that isolates variation in equity gains that is plausibly exogenous to idiosyncratic household choices. We construct an instrument, $\Delta \text{HPI}^{S}_{z,t}$, for a household's equity gain using the change in the median zip code housing prices over the years that the household lived in the home,

$$\Delta \text{HPI}^{S}_{z,t} = \text{HPI}^{S}_{z,t} - \text{HPI}^{S}_{z,t-tenure} \tag{9}$$

where $\text{HPI}^{S}_{z,t}$ is the Zillow house price index for the zip code of the sold property at the time that the house is sold, and $\text{HPI}^{S}_{z,t-tenure}$ is the house price index at the time that the house was originally purchased. By using local area house price appreciation as an instrument for equity gain, we exclude the variation in a household's equity gain that is due to bargaining ability or other market sophistication of the household when they sold their previous home.[5]

Using the change in the zip code house price index as an instrument, we estimate the first stage regression:

$$Equity\ Gain_{i,t} = \beta\, \Delta \text{HPI}^{S}_{z,t} + \lambda\, \text{HPI}^{S}_{z,t-tenure} + \alpha^{S}_{z,y} + \alpha^{P}_{z,y} + \alpha^{tenure}_{i} + \alpha^{S}_{h} + \varepsilon_{i,t} \tag{10}$$

where $\text{HPI}^{S}_{z,t-tenure}$ is the zip-level house price index of the sold property in the year that the household originally bought the home, $\alpha^{S}_{z,y}$ is a sold property zip code by year fixed effect, $\alpha^{P}_{z,y}$ is a purchased property zip code by year fixed effect, α^{tenure}_{i} is a years lived in sold home fixed

[5]The one possible exception is within-year strategic timing. However, our results are robust to including an additional calendar month fixed effect to absorb seasonal differences in house prices. We can also replicate our results with sold and purchased zip-quarter fixed effects, which effectively eliminates the potential for any meaningful differences in strategic timing. See Appendix Table A.6.

effect, and α_h^S is a fixed effect for the set of interacted house characteristics described above. Unsurprisingly, the change in zip code house prices over the years lived in the house strongly predicts the realized equity gain; the first stage F-statistic is 241.4 in our main specification.

We then use the predicted equity gain from Equation 10 to estimate the following second stage regression.

$$PricePremium_{i,t} = \beta \widehat{Equity\ Gain}_{i,t} + \lambda \text{HPI}^S_{z,t-tenure} + \alpha^S_{z,y} + \alpha^P_{z,y} + \alpha_i^{tenure} + \alpha_h^S + \varepsilon_{i,t} \tag{11}$$

We expect that $\hat{\beta}_{IV} > \hat{\beta}_{OLS}$, since our instrument is predominantly designed to eliminate the effects of housing market sophistication which are likely biasing our naïve estimate downward. To interpret $\hat{\beta}_{IV}$ as the causal effect of equity gains on the price premium, the instrument must satisfy the exclusion restriction that changes in the average house price in the sold property zip code over the holding period do not affect the amount paid in excess of the current quality and neighborhood trend adjusted value of the purchased property except through changes in the household's housing wealth, conditional on our fixed effects and controls.

The fixed effects used in Equation 11 severely limit potential violations of this exclusion restriction. Any remaining concern is limited to a shock that simultaneously affects the average price of houses in the entire sold property zip code and the price of houses only in a small neighborhood within the purchased property zip code. The shock cannot affect the entire purchased property zip code average, or it will be absorbed by the purchased property zip-year fixed effect. Additionally, the shock has to differentially affect properties within the purchased census tract-year, since the price premium is orthogonal to average prices in this dimension. One possibility, consistent with this type of shock, could be the opening of a new, highly regarded school that has boundaries that include all of the sold zip code, but only a very small corner of the neighbor-

ing zip code, wherein the household purchased their new property. It is unlikely that this type of narrow story explains the average effect of our results.[6] Moreover, if paying higher prices reflects persistent changes in very local amenities, we would expect the household to receive a higher price when selling the house in the future. However, we show below that a higher price premium predicts *lower* future returns from holding the property.

One additional concern is that the realized equity gain on a sold property influences the decision of where to move. For example, a household that experiences a substantial appreciation in local house prices might endogenously decide to sell and move to a lower priced area, either to significantly upgrade their house or to extract the equity gains. It is not obvious how this selection affects the price premium, since it is defined relative to average prices in the purchase zip-month and census tract-year. Moreover, we include both sold and purchased property zip-year fixed effects to account for any differences in house price dynamics across move locations. To more fully account for potentially endogenous location decisions, in the Appendix Table A.6 we replicate our results using a different specification that includes a purchase county × sale county × year fixed effect. This specification is conditional on move location choices—we compare two buyers moving from the same county to the same county in the same year, eliminating any differences due to move location decisions. We find similar results using this alternative specification. In Section 6, we further account for endogenous move location decisions at the aggregate commuting-zone-level using pre-established migration routes as an instrument for the move decision.

[6]Our results are also robust to both excluding within-county moves and also to excluding moves of less than 3 miles (see Appendix Tables A.6 and A.7), which is where these types of stories are most prevalent.

5 The Effect of Equity Gains on Housing Transactions

While there is a substantial literature examining frictions that impact residential real estate transactions (Kurlat and Stroebel, 2015; Gargano, Giacoletti, and Jarnecic, 2020), lack of data has prevented researchers from exploring how households make housing investment decisions across properties. The novel transaction-level panel dataset that we construct in Section 2 provides the first large, nationally representative source of data that tracks households across housing transactions. Using this data, combined with the two-stage least squares strategy described in Section 4, we examine how changes in household wealth affect housing decisions when homeowners change their primary residence.

5.1 Equity Gains Cause Overpayment

Figure 1 provides preliminary evidence of the relation between equity gain and price premia. We begin to investigate the positive correlation visible in the figure by estimating the OLS model specified in Equation 8. The regression result, reported in Table 2 column (1), confirms the correlation visible in the figure. There is a positive and significant relation between equity gain and price premia—$1 of equity gain on the sold property leads to $0.01 of overpayment on the purchased property. However, as discussed in Section 4, this estimate is likely biased downward due to a housing market sophistication effect that simultaneously results in a household receiving a larger equity gain on the sold property and paying a lower price on the purchased property.

To identify the causal relation between equity gains and price premia, we use the change in the zip-level house price index over the period that the household owned the home as an instrument for the household's equity gain as specified in Equations 10 and 11. Because this instrument is designed to address housing market sophistication concerns, we expect the IV estimate to be

larger than the OLS estimate. The results confirm that this is the case. In Table 2 column (2) we show that the instrumented effect of equity gain on price premia increases to \$0.08 per dollar of equity gain. For the average household in our sample, which has an equity gain of \$86,000, this implies an overpayment amount of about \$6,800.

The large increase in $\hat{\beta}_{IV}$ relative to $\hat{\beta}_{OLS}$ suggests that housing market sophistication is prevalent in the data. To explore how plausible this is, we calculate the annualized net return a household earned on their sold property over their holding period. We then estimate the future returns that a household will earn on their purchased property as a function of their realized returns on their current sold property.[7] We include the same set of fixed effects as in Equation 11, which absorbs average prices at both the sold and purchased property location. The results are reported in Appendix Table A.1. We find a positive and highly significant relation between the realized return on a sold property and the future realized return on the purchased property, consistent with persistent housing market sophistication. Because of the dollar on dollar specification, we can interpret the coefficient as an 8% marginal propensity to overpay out of a dollar of exogenous equity gains.[8]

While the marginal propensity to overpay out of housing wealth is an important economic concept, it is also interesting to understand the extent to which equity gain-induced overpayment scales up with the price of the purchased property. To investigate this, we define *Price Premium Percent* as the dollar price premium on the purchased property (i.e., Equation 2) divided by the total price paid for the property. Using the same two-stage least squares framework, we estimate the effect of equity gain on *Price Premium Percent* and report the results in Table 2 column (3). We find that equity gains cause a statistically significant increase in the price premium as a percentage

[7]Both are winsorized at a 1% and 99% level.

[8]This share of housing wealth gains spent on a higher price premium is economically quite significant. It is comparable in magnitude to estimates of marginal propensities to consume out of housing wealth that range from 6–11% (Campbell and Cocco, 2007; Carroll, Otsuka, and Slacalek, 2011) .

of the purchase price. A household with a $100,000 equity gain on their sold house overpays by 2.1% of the purchase price on their subsequent house purchase.

One potential concern with the interpretation of these results is that our measure of the quality and neighborhood trend adjusted price might not accurately reflect all information. In particular, perhaps households with large realized equity gains are especially likely to purchase houses that have value-enhancing characteristics that are not absorbed in the price premium model specified in Equation 2 such as substantially renovated properties. We cannot test this directly because our data does not allow us to observe property improvements.[9] To get a sense of how plausible this concern is, we estimate an OLS regression of the future realized returns on the purchased property as a function of the price premium. To the extent that our measure of the price premium represents the excess of the price over the correctly adjusted value of the housing asset, we expect future returns to be lower when the price premium is higher. In contrast, if the price premium represents paying the correct price for, e.g., recently completed renovations, we would not expect to see any effect on future returns. In column (4), we find that the price premium is associated with significantly lower future returns—a household that pays more relative to the quality and neighborhood trend adjusted value by the average amount in our sample ($11,538) receives a 1.4% lower return when selling the property in the future. This suggests that a substantial portion of what we define as the price premium actually is an amount in excess of the correctly adjusted value of the house.

Could the measured positive price premia associated with higher equity gains reflect recent renovations that depreciate over time such that they are both not fully captured by our property fixed effects and also reduce returns in the long run in line with the finding above? To test for this possibility, in Appendix Table A.2 we explore how the negative returns associated with greater

[9]Importantly, problematic property-level improvements would have to be performed in excess of the average level of improvements in the census tract.

price premia vary with the holding period. Recent renovations driving the measured price premia should reduce the returns more in the long-run (as they depreciate over time). We find that the decrease in the annualized return associated with overpayment is in fact larger for households that sell again quickly, suggesting that overpayment is not being driven by paying for recent renovations to the house that depreciate over long holding periods.

The conceptual framework described in Section 3 shows that equity gains can lead households to rationally substitute overpayment for search effort. However, an alternative explanation for these results could be that households benchmark the purchase price to previous transactions. A household that offers a price based on benchmarking is especially likely to overpay when moving from a high price to a lower price area. We examine this in column (5) of Table 2 by including an indicator variable for moving to a lower price zip.[10] We also interact this indicator with the equity gain. We find evidence that benchmarking exists—households that move to lower priced areas overpay more on average. However, the relation between equity gains and the price premium is not primarily explained by benchmarking. After controlling for moving to cheaper areas, the total effect of equity gains on the price premium falls to about \$0.06 per dollar of gain, or roughly 80% of the original magnitude.

5.2 Why do Equity Gains Cause Overpayment?

It is puzzling that households spend some of their realized wealth on overpayment, particularly given the financial costs documented above. Additionally, households sacrifice the opportunity costs of these overpayment dollars. Absent frictions, it seems likely that households would be better off using these overpayment dollars to either buy a bigger/better house or to consume or save the money. The conceptual framework in Section 3 suggests a possible explanation: buyers

[10]Indicator is based on the relative house price indices of sold vs. purchased zip.

trade off the costs of information gathering and additional search with overpaying for a house. In this section, we provide evidence consistent with that framework..

On the one hand, buyers in residential real estate markets face substantial information asymmetry (Kurlat and Stroebel, 2015). Overcoming these information frictions is costly—buyers have to expend substantial time and effort to become informed about the quality and neighborhood trend adjusted value of a particular house. Buyers with large equity gains might choose to remain ignorant and overpay rather than exerting the effort necessary to discover the value of the house. On the other hand, to the extent that equity gains relieve capital constraints, buyers can consider a larger set of houses. This increases the probability that the buyer finds a match for which they have a higher private valuation, and consequently is more willing to overpay. Gargano, Giacoletti, and Jarnecic (2020) provide evidence that is consistent with both of these channels. They show that local area price appreciation causes capital-constrained potential movers to search across a broader set (both in terms of geography and in terms of house characteristics) of potential properties. This broader search could lead to a better, higher private valuation match. However, Gargano, Giacoletti, and Jarnecic (2020) also show that this increase in local price appreciation does not change the amount of time potential movers devote to any particular listing and furthermore shortens the ultimate duration of the search. Searching across a broader set of properties in a shorter period of time could be consistent with putting less effort into acquiring information about specific properties.

We distinguish between these two channels by investigating how the sensitivity of the price premium to equity gains varies based on the characteristics of the local market as well as those of the moving household. We first examine the interaction of equity gains with historical transaction volume, defined as the number of transactions that occurred within a half-mile radius of the purchased property over the prior 90 days. A deep pool of recent transactions suggest that

there is an abundance of information about recent comparable transactions. Because this reduces information asymmetry, the information channel predicts that the sensitivity to equity gains will be lower when transaction volume is high. However, high historical transaction volume also suggests that the market is very liquid with lots of potential properties for sale. Because this increases the probability that the household finds a good match, the private valuation channel predicts that the sensitivity to equity gains will be higher when transaction volume is high.

We estimate the interaction between historical transaction volume and equity gain using our instrumental variable model for the price premium.[11] The results are reported in Table 3 column (1). Consistent with the information channel, and inconsistent with the private valuation channel, we find that the sensitivity of the price premium to equity gains is lower in markets with higher historical transaction volume. Moving from the 25^{th} to the 75^{th} percentile of historical transaction volume decreases the effect of equity gain on overpayment by about 7%.

In Table 3 columns (2) and (3), we use neighborhood price dispersion as an additional proxy for the information environment of the local area. We measure price dispersion in two ways: the standard deviation of price premia and the standard deviation of transaction prices for all transactions that occur within 90 days before the purchase and within a 1/2 mile radius of the purchase property.[12] While the transaction volume result reported in column (1) measures the availability of comparable historical sales, columns (2) and (3) measure the precision of those comps. Higher price dispersion in the recent local market increases the difficulty of determining the correct price for a property. Consequently, the information channel predicts that the sensitivity of overpayment to equity gains will be higher for areas with higher price dispersion. Across both measures of price dispersion we find that the sensitivity of overpayment to equity gain is higher in more

[11] These specifications have two potentially endogenous variables: equity gain and equity gain×characteristic. We use our original instrument (change in zip-level house prices) interacted with the characteristic as a second instrument for these specifications.

[12] We divide the standard deviation by 100 to make the coefficient easier to read.

uncertain environments.

In Table 3 columns (4) and (5), we investigate the interaction of equity gain with the number of years that the household lived in their sold property. For these analyses, it is necessary to subset the sample to within-county moves. For this set of local moves, a household that has lived in their sold home for a longer period of time is likely to be more familiar with the local area, and thus face less information asymmetry when moving. Consequently, the information channel predicts that the sensitivity of price premia to equity gains will be lower for these long-tenured local movers. In contrast, households that have lived in the area longer and that are also moving within that same area are likely to have more strongly developed housing preferences, perhaps due to having relatively older children or more specific community ties. Moreover, these long-tenured households have implicitly had a longer period of time to search the local housing market (even if only passively), and so they are more likely to have found houses that are particularly well-matched for their preferences. As a result, the private valuation channel predicts that the sensitivity to equity gains will be higher for local, long-tenured movers.

We estimate the interaction of equity gain both with a continuous measure of the years lived in the sold property in column (4), as well an indicator variable for households that have lived in their home for more years than the sample median in column (5). Across both definitions we find that the sensitivity of the price premium to equity gain is lower for long-tenured local movers. A household that has lived in their house more than the median number of years has a sensitivity of the price premium to equity gain that is nearly 30% smaller than a household that has lived in their home less than the median number of years. Similar to the transaction volume results, this is consistent with the information channel and inconsistent with the private valuation channel.

Appendix Table A.3 shows that the results in Table 3 are robust to using the price premium percentage, rather than dollar price premium, as the dependent variable. Combined, this evidence

suggests that information frictions drive the relationship between equity gain and overpayment. Consistent with the conceptual framework in Section 3, households substitute the effort costs of acquiring information about a house with simply overpaying.

5.3 Equity Gains Cause Neighborhood Price Spillovers

In this section, we investigate the impact that a household with large equity gains has on neighborhood house prices when they move in. Because residential housing markets are relatively illiquid and because prices are often based, in part, on previous transactions used as comparables, it is possible that households with large equity gains that overpay for their house cause prices to go up in the surrounding neighborhood.

To investigate this possibility, we first define the neighborhood price following a property purchase as

$$\text{NeighborhoodPrice}_{i,N,t,T} = \frac{\sum_{s=0}^{T} \sum_{j \neq i \in N} P_{j,t+s}}{\sum_{s=0}^{T} \sum_{j \neq i \in N} 1} \tag{12}$$

where household i purchases a house in neighborhood N at time t. The neighborhood price is the average price of all homes j located in neighborhood N that sell within T days of the purchase. We examine spillovers for two time windows T (180 and 360 days) and four different definitions of neighborhood N (houses within 0.5, 1, 1.5, and 2 miles of the purchased house).

We estimate the spillover effect that a household's equity gain from their sold house has on prices in the neighborhood surrounding their purchased house using the same two-stage least squares framework that we use to study overpayment. The results are reported in Table 4. Focusing on column (1), we find that a household moving in with a larger equity gain causes a larger increase in neighborhood prices. An extra \$1 of equity gain causes the average price of homes sold in the half mile radius around the purchased property to go up by \$0.09 in the 180

days after the purchase. The magnitude of this price spillover is very similar to the magnitude of the effect of equity gain on the price premium, which is consistent with the overpayment being incorporated into neighborhood house prices.

Looking across the columns of Table 4, we find that the impact of equity gain on neighborhood prices diminishes both with time and over distance. By one year after the purchase, the effect of equity gain on house prices within a half-mile is still significant, but 20% smaller. In contrast, for houses within 2 miles, the effect has fallen by 60% and is no longer statistically significant. This pattern is strongly consistent with price spillovers operating through a comparables channel, either directly (through appraisals) or indirectly (through affecting sellers' reservation prices).[13]

Our identification strategy is designed to estimate the causal effect of equity gain on price premia and spillovers. Thus, while we believe that the most likely channel through which equity gains cause price spillovers is overpayment, we cannot test that directly. However, in Table 5 we provide additional evidence that suggests that the spillover effect operates through the overpayment channel by showing that the spillover effect of equity gain on neighborhood prices varies with the information environment in strikingly similar ways as the effect of equity gain on the price premium. Specifically, we find that the effect of equity gain on price spillovers is smaller in areas with high historical transaction volume and for households moving locally that have lived in the area for longer periods of time, and larger for areas with increased historical price dispersion.

While not dispositive, the collective evidence in Table 5, combined with the fact that the

[13] In line with the declining spillover effects when going beyond a 1-mile distance that we find here, Freddie Mac's "Single Family Seller Servicer Guide" notes that an appraiser would most likely use comparables "in the immediate vicinity" of the property if it is in a suburban or urban area. (URL: https://guide.freddiemac.com/app/guide/home). Similarly, the "HUD Instructions for Completing the Uniform Residential Appraisal Report" declared in 1994 that a separate explanation was needed, "if comparable is more than 1 mile from subject" (URL: https://www.hud.gov/sites/documents/DOC_36119.TXT). Additionally, this measure is similar to those utilized in the foreclosure externality literature. See, e.g., Fisher, Lambie-Hanson, and Willen (2015) and Gupta (2019).

magnitude of the effect of equity gain on price spillovers is very similar to the magnitude of the effect of equity gain on price premia, strongly suggests that equity gain-induced overpayment pushes up prices in the neighborhood surrounding the purchased property.

5.4 Local vs. Non-local Movers

A remaining interesting question is the extent to which the sensitivity of price premia and spillovers to equity gain varies by the distance of the move. A popularly held opinion in the press is that wealthy, out-of-town home buyers are driving prices up for everyone.[14] However, the conceptual framework in Section 3 suggests that the effect of equity gains on price premia should actually be lower for non-local movers. In this section, we explore the extent to which that is true for individual transactions. However, because the effects of spillovers might build on each other, it is possible that the aggregate effects of housing equity flowing into an area differ from the transaction-level effects. We explore the aggregate effects of housing equity inflows on commuting-zone-level prices in a separate analysis described in Section 6.

To investigate the effect of move distance, we split household moves into four categories based on the sale and purchase locations: within zip; out-of-zip, but within county; out-of-county, but within state; and out-of-state. We include indicators for each of these move types, as well as the interaction of these indicators with equity gain, in our two-stage least squares estimates of the effects of equity gain on price premia and price spillovers.[15] Note that our baseline category is within-zip moves, so the estimated coefficients represent the marginal difference in the sensitivity of price premia and spillovers to equity gains relative to households that move within the same zip code.

[14]For example, consider this recent headline from Bloomberg, "Out-of-Town Home Buyers Will Pay 30% More Than Locals in Hottest U.S. Markets." See https://www.bloomberg.com/news/articles/2022-02-15/top-10-most-competitive-housing-markets-where-out-of-towners-outspend-locals

[15]As with Table 3, we use the indicator×IV as an instrument for the interactions.

The results are reported in Table 6. In columns (1) and (2), we show how the sensitivity of the price premium to equity gain varies across move distances. Unsurprisingly, and consistent with overpayment being a function of information asymmetry, the average price premium is monotonically increasing with distance. However, the sensitivity of the price premium to equity gains follows the exact opposite pattern. Non-local movers use less of their equity gain to overpay—out-of-state movers have a sensitivity of price premia to equity gain that is half the magnitude of within-zip movers.[16] In Column (2), we show that this result is not driven by our choice of move categories—the effect of equity gain on price premia continuously decreases with the distance of the move.

Not only does local equity gain result in a larger overpayment than non-local equity gain, but it also results in higher spillovers. Columns (3) and (4) in Table 6 show that the effect of equity gains on price spillovers also monotonically decreases with distance. Relative to within-zip movers, out-of-state movers have an spillover sensitivity to equity gains that is 15% smaller.

Our results may be driven by large and wealthy counties (such as Los Angeles County, CA) where locals overpaying more than non-locals might seem intuitive. To test this, we run a specification similar to that of column (2) of Table 6 at an individual county level, for the 100 largest counties by observation count in our main regression sample. Rather than indicators for four classes of move distances across political boundaries, we collapse to a single binary indicator for whether the move was a non-local (i.e., out-of-county but within-state or out-of-state) move.[17] Appendix Figure A.1 displays a map of these 100 counties as well as information regarding the

[16] For the average equity gain in our sample, the total effect of equity gain on the price premium varies very little across move distance. Interestingly, if spillovers are driven by overpayment, this would suggest that the average effect of equity gain on spillovers will not vary by distance. This is consistent with the level effects of the distance indicators being small and insignificant in Table 6 Column (3).

[17] The specification is also altered from that of Equation 11 because the focus on a single county obviates some of the fixed effect and clustering dimensions. The zip-year fixed effects for both the sale and purchase locations are replaced by a transaction year fixed effect and the purchase zip clustering dimension is dropped.

sign and significance of the (instrumented) coefficient on the interaction between equity gain and a non-local move. Red counties have negative coefficients (non-locals overpay less than locals), while blue counties have positive coefficients (non-locals overpay more than locals). Significance at the 95% level is displayed by a deeper shade. Appendix Figure A.2 displays the coefficients and 95% confidence intervals for each of the 100 counties underlying the map. There are no counties where non-locals are overpaying more than locals as a function of their equity gain at a 95% confidence level.[18] These results suggest both that our results are not being driven by large metropolitan areas and that the popular press narrative is not true for any individual major county.

As a final robustness check, we replicate columns (2) and (4) of Table 6 for the pre-COVID-19 period (1996-2019) and the COVID-19 period (2020-2021) in Appendix Table A.4. During the pre-pandemic period, displayed in columns (1) and (3), we recover results consistent with our main sample effects, demonstrating our results are not driven by pandemic induced shifts in household behavior.

Together, the evidence in this section suggests that, at least at the individual level, equity gains in the hands of local movers actually drive prices up more than equity gains in the hands of non-locals. The conceptual model in Section 3 provides the intuition behind this result. In a simple model of housing search, a search technology that translates effort more effectively into lower search costs will be associated with a higher reservation price and a higher expected price premium paid by the buyer. For a local mover, information gathering is particularly efficient, and so the reduction in time spent gathering information after experiencing an equity gain results in a large increase in the price premium. As move distance increases, search efficiency decreases. As a result, long-distance buyers overpay more on average (because they are less informed),

[18] Although one county out of the hundred tested, Naples, FL, is significantly positive at the 90% level.

but simultaneously use less of their equity gains to overpay (because the benefits to avoiding information gathering are small). The popular press narrative regarding wealthy out-of-towners outbidding locals requires some qualification: while out-of-towners overpay more on average, this is less a function of their wealth—proxied by housing equity gains—than for locals. However, in some areas non-local movers may on average have experienced higher equity gains than locals, which may more than make up for their smaller sensitivity to each dollar of housing equity. The results in this section show that the popular discussion misses the important distinction between the sensitivity to equity gain, the total housing equity gains of movers, and their overall tendency to overpay due to information disadvantages.

5.5 Characterizing the full effects of equity gain

We have established that increases in housing wealth cause households to overpay for their next home. How do equity gains affect other aspects of the housing purchase decision? Our data, combined with our two-stage least squares approach, allow us to explore the effect of equity gains on various aspects of the purchase price, as well as on the characteristics of the home that is purchased.

In column (1) of Table 7, we present the results from estimating the effect of the sold property equity gain on the purchase price of the subsequently purchased property. We find that for every $1 of equity gain, households spend another $0.87 on their next house. We can decompose this spending into the price premium (see Table 2) and the quality and neighborhood trend adjusted value of the house (defined as the predicted price from the price premium model in Equation 2). Mechanically, these two coefficients sum to the overall effect of equity gain on purchase price. In Column (2) we show that $1 of equity gain causes buyers to purchase a house that has $0.79 higher quality and neighborhood trend adjusted value. This implies that the spending on over-

payment ($.08 per dollar of equity gain) represents about 9% of the total equity gain-induced house spending.

A well-documented fact in the literature is that buyers who pay in cash receive a discounted price (Reher and Valkanov, 2021). If equity gains are large enough, it is possible that they allow a household to purchase a home in cash, which would push against the overpayment results we document. In Appendix Table A.5, we find that equity gains do not predict the probability of purchasing a home in cash.[19] The results further confirm that households that pay in cash pay less relative to the quality and neighborhood trend adjusted value (i.e., they receive a cash discount), but even accounting for this fact households with large equity gains still overpay on net.

Across the remaining columns of Table 7, we show how equity gain affects the purchase neighborhood and the purchased property characteristics. We find that equity gains cause households to move to a more expensive zip code and to purchase a modestly older, but significantly larger house. For example, a household with the average equity gain in our sample purchases a 5% larger home (or 120 square feet larger based on the average house in our data). While overpayment is the most salient and interesting aspect of a household's response to equity gains (see Section 5), these results highlight the many and varied ways that a household's housing consumption and investment decisions are determined across multiple transactions.

6 Aggregate Impact of Housing Equity Gains

An important policy concern related to housing capital flows is what share of overall house price growth in destination locations can be attributed to the inflow of housing capital gains.

[19] This is less surprising given that our average equity gain is $86,000, and even the 95th percentile is only around $300,000; not enough to purchase the average home in our sample.

There are several reasons why the spillover effects estimated above cannot easily be translated into an estimate of the aggregate effect: (1) Our property-level estimates condition on characteristics of the buyers' areas of origin, which at least partially eliminates the effects coming from an area's exposure to particular *geographies* (rather than households within geographies) that have higher equity gains. (2) The spillover effects on other properties may lead to knock-on effects on further properties and may interact with the spillovers from other purchases in the area—as a result, the aggregate effect may differ from the partial equilibrium effects. (3) If high housing equity purchases cluster in the same year and county, then the effect on individual properties and their neighborhood may cumulate as overpayment for the first purchases enters the comparables for the later ones.

6.1 Aggregate effect specification

Therefore, we analyze the overall effect of equity gains moving into a neighborhood by considering the impact on house price growth at the commuting zone (CZ) level.[20] We focus on identifying the effect of buyers moving into a commuting zone from outside, as it is easier to construct plausibly exogenous variation with regard to destination CZ trends in such long-distance moves.[21]

We estimate CZ-year-level regressions of the form:

$$\Delta \ln P_{c,t} = \alpha_c + \alpha_t + \beta \ln eg_{c,t}^{\text{non-local}} + \epsilon_{c,t}, \tag{13}$$

where we include CZ c and year t fixed effects (α_c and α_t) in order to capture general differences

[20] We use commuting zone as a proxy for contiguous housing markets.

[21] The moving decision of households that come from the same CZ in which we are measuring our outcome variable is more likely to be endogenous with regard to local house price trends.

in the growth rate of house prices across housing markets, as well as macroeconomic trends (e.g. interest rate changes) that affect all locations equally. The dependent variable is the CZ's growth rate (log change) in the Federal Housing Finance Authority's repeat-sales house price index.[22] The equity gain variable $eg_{c,t}^{\text{non-local}}$ is the expected total equity gains brought by in-migrants, scaled to be in units of dollars per local household in the destination CZ.

This total inflow of equity gains cannot be directly measured in our property-level data as we do not necessarily have full coverage of all transactions in a CZ. Instead, we combine our data on average equity gains among buyers with data on total movers between counties from the Internal Revenue Service's Statistics of Income (SOI) database.[23] First, we use the ZTRAX data to estimate the average equity gain among observed out-of-CZ buyers in destination c. Second, we scale this average equity gain $\overline{\text{EquityGain}}^{\text{non-local}}$ into an estimate of total equity gains from all non-local movers per local household by calculating

$$eg_{ct}^{\text{non-local}} = \overline{\text{EquityGain}}_{ct}^{\text{non-local}} \times \frac{\text{OutOfCZMovers}_{ct}^{\text{IRS}}}{\text{ResidentHHs}_{c,2006}^{\text{IRS}}}, \tag{14}$$

where $\text{OutOfCZMovers}_{ct}^{\text{IRS}}$ is the IRS estimate of total household inflows into the commuting zone c from origin CZs at least 50 miles away and $\text{ResidentHHs}_{c,2006}^{\text{IRS}}$ is the total number of local households residing in the CZ in 2006.[24]

Movers between CZs may be selected in a way that would introduce bias into a simple OLS estimation of the effect of these equity gains on local house price growth. For example, movers that sell their house in response to high house price growth in their origin city may be more

[22] This is constructed from county-level HPI data from the FHFA, which we convert into annual log changes at the county level, and then aggregate to CZ level average log changes, weighting each county by its population in the year 2000. For CZs with < 10% of its population in counties with FHFA data in that year, missing counties' data is replaced by state-level log changes in the HPI before aggregating.

[23] See Appendix section A.1 for full details on the construction of equity gain flows and the instrument

[24] 2006 is the first year of our panel for the aggregate effects estimation.

likely to move to other cities experiencing high house price growth (for instance, because there is a common cause for these house price changes, such as a revaluation of shared natural amenities). At the same time, high house price growth destinations may attract additional equity gain inflows from speculators *because* their prices are rising. To mitigate these concerns around the endogenous destination choice of movers, we construct an instrument that captures plausibly exogenous equity gain flows between CZs.

To construct a plausibly exogenous flow of *predicted movers* for each origin-destination CZ pair, we use historical migration links as predictors of contemporaneous migration flows (Howard, 2020; Schubert, 2021). That is, the predicted number of movers from k to c is constructed from IRS migration data as a "shift-share" instrument that captures the degree to which the characteristics of CZ k are prompting outflows in period t, and the average share of those outflows that would be expected to go towards CZ c—for instance, because of low migration costs or historic family connections between these locations[25]—but without using any period t variation from CZ c itself that might potentially be correlated with local house price trends in c. Moreover, when aggregating these predicted flows from other CZs, we again omit migration flows from CZs that are less than 50 miles away, in order to eliminate bias from the migration dynamics of CZs that may have overlapping local housing and labor markets with the destination CZ of interest.

Similarly, the predicted average equity gain among people moving into CZ c is constructed without using data on the characteristics of contemporaneous flows from k to c, which might be selected based on destination characteristics. Instead, we assume that movers from high equity gain locations are also generally more likely to have higher equity gains. Thus, destination CZ c is expected to have higher average equity gains among its incoming households if it has higher historic migration exposure to CZs that are experiencing high equity gains among their house

[25]See Schubert (2021) for an analysis of the determinants of migration flows in recent decades, which include similar industry composition and demographics, physical proximity, and cultural similarity.

sellers in period t, so we construct a weighted average $\overline{EquityGain}_{ct}^{\text{origins}}$ of those origin city characteristics with weights based on historical migration flows.

We combine these measures to construct our instrument for equity gains (per local household) flowing into CZ c from other CZs:

$$\widetilde{eg}_{ct}^{\text{non-local}} = \underbrace{\left(\sum_{k\,:\,dist(\text{CZ(k)},\text{CZ(c)})>50\text{mi.}} \widehat{Mig}^{k\to c} \right)}_{\text{Predicted Out-of-area Inmigration}} \times \overline{EquityGain}_{ct}^{\text{origins}} \times \frac{1}{\text{ResidentHHs}_{c,2006}^{\text{IRS}}} \tag{15}$$

This shift-share approach results in an instrument for the total equity gain inflow into CZ c in year t that does not use any contemporaneous information on the actual number or characteristics of movers with equity gains coming into c and thereby eliminates any correlation between the contemporaneous destination choices of movers and local house price dynamics in c. Instead, the instrument relies on variation across CZs in their historical exposure to migration from particular distant CZs, in conjunction with variation over time in the average equity gains in those distant CZs.

6.2 Aggregate effect estimates

We use this instrument in a two-stage least squares setup to obtain an estimate of the causal effect β of out-of-area housing equity gain inflows on local house price growth in equation 13. We also include CZ fixed effects in the estimation to account for the fact that some areas may have higher house price growth trends on average, as well as year fixed effects to control for common macroeconomic trends in house prices.

The results are shown in Table 8. Column (1) shows the first stage, corresponding to a regression of the log of $eg_{ct}^{\text{non-local}}$ on the log of $\widetilde{eg}_{ct}^{\text{non-local}}$. The instrument has a significant and positive

effect on the predicted actual equity gains flowing into the CZ. In column (2), we first use OLS to estimate the specification shown in equation 13. The raw relationship between housing equity gain inflows and local house price growth is positive but relatively small, with a coefficient of about 0.7. In column (3), we use our instrument to estimate the causal effect of out-of-area housing equity gains flowing into CZ c on its house price growth. We find that out-of-area equity flows have a positive and significant causal effect on the destination CZ's house price growth, with an estimated coefficient on local house price growth with regard to equity inflows per local household of about 4. This implies that a 10% increase in the average housing equity gain per local household brought into a CZ by out-of-area movers causes about a 0.4 percentage point increase in local house price growth. The higher IV coefficient suggests that the OLS estimate is downward-biased. That is, migrants that have experienced high housing equity gains seem to choose on average to move to markets that are experiencing lower house price growth in that period, in line with movers responding endogenously to differences in housing market opportunities across cities.

We implement a number of robustness checks to address potential concerns about the identification of the aggregate equity gains effects: On the one hand, Schubert (2021) and others have shown that migration into an area can increase house price growth, even without considering the equity gains of migrants. We control for overall migration pressures into the area by including the log of the predicted out-of-area inmigration into c (as defined in equation 15) as a control variable in the regression, with results shown in column (4). In column (5), we control for whether the destination CZ is a "hot" market with high housing market turnover, by including the log of the total number of sales transactions in our ZTRAX data per local resident household in the regression. In column (6), we explore whether shocks to local labor demand—perhaps interacting with local supply constraints (Saiz, 2010)—might be jointly driving local house price growth and

equity gains inflows. We control for this possibility by directly including shift-share shocks based on local industry exposure to national wage trends in the regression, as suggested by Chodorow-Reich and Wieland (2020), and also interacting them with the share of local land unavailable for construction as proxy for housing supply constraints (Lutz and Sand, 2022).[26] Our IV estimates of the causal effect of equity gain inflows on local house price growth are significant and little changed in magnitude across these different specifications.

Moreover, the literature on the house price effects of migration generally suggests that the size of the effect of a change in local housing demand—here due to higher equity gains among migrants—should depend on local housing supply constraints (Saiz, 2010). In column (7), we show that this is also the case here: interacting the equity gain inflows (and the instruments) with quintiles of CZ-level land constraints, we find that the house price effect monotonically increases with land constraints and is not significant for the two least-constrained quintiles of commuting zones.

Note that, while the log-change-to-log specification here is not directly comparable to the property-level results in the previous section, the aggregate effect estimates are generally larger than our estimates of property-level impacts of housing equity gains on house prices. There are two possible reasons for this. First, our overpayment results deliberately control for zip code level price trends in order to avoid conflating a property's higher price with increasing prices in an area in general. This means that if multiple high equity gain movers were to move into the same neighborhood at the same time then the overpayment estimates we recover are net of the *common effect* of high equity gain movers on the zip code house prices. In contrast, the

[26]The reason why we use labor demand shocks rather than local wages themselves is that local wages are endogenous with regard to local housing market dynamics, so actual local wage growth would be a "bad control" in this setting as it may capture—and thus eliminate—the variation of interest in the dependent variable. See Schubert (2021) for the construction of the included labor demand "Bartik" shocks based on multiplying a vector of local CZ industry composition in 1990 by industry-specific leave-one-out national wage growth.

aggregate effect analysis shown here includes the common effects caused by multiple moves into the same area, which is the relevant effect for quantifying the importance of these flows for housing market dynamics. Second, the spillover effects found above could cumulate over time within a year. If one high equity gain purchase increases a neighborhood's prices, it may then feed into the comparables for later purchases, which, in turn, may be affected by overpayment from other high equity gain movers. The resulting aggregate effect of multiple movers into a housing market can be higher than individual property effects because of this multiplicative dynamic.

In general, the large positive effects from inflows of out-of-area housing equity gains on local housing markets provide causal evidence of a mechanism oftentimes postulated in anecdotal reports of why house prices lift off in smaller markets experiencing inflows. We find that the housing equity gains of the migrants indeed play an important role by amplifying the impact on the destination housing markets of the migrants with higher gains from their previous sale.

7 Conclusion

House price appreciation over the last several decades has generated substantial wealth for homeowners. Using a novel dataset that allows researchers to track homeowners across moves, we document a surprising fact: households use this housing wealth to overpay for their next house. This overpayment has real consequences on household wealth; overpayment leads to significantly lower future returns when selling the property. Given this, why do households use their housing wealth to overpay?

We show that the relation between equity gains and overpayment is driven by information frictions in the residential real estate market. Because housing markets are segmented and illiquid, acquiring information about the quality and neighborhood trend adjusted value of a house

takes a significant amount of time and effort. Households use equity gains to substitute for the effort of costly information acquisition; they choose to overpay and remain ignorant about the price rather than pay the effort costs necessary to become informed. This individual behavior creates externalities—households with large equity gains that move into an area drive up their new neighbor's home prices by approximately the same amount that they overpaid.

Unexpectedly, the individual effect of equity gains on price premia and on price spillovers is higher for local than for non-local movers, and this appears to be true for the vast majority of counties in our data. In this sense, the common story that rich out-of-towners drive up prices is incomplete—local movers with lots of housing wealth drive up neighborhood prices by more than similarly wealthy non-local movers, albeit relative to a baseline of higher overpayments by out-of-towners regardless of their wealth. At the aggregate-level, though, the effects of housing wealth flowing into an area do put a significant amount of price pressure on local house prices.

Together, our results show that heterogeneity in housing wealth plays an important role both in explaining household-level housing decisions and in explaining aggregate-level housing dynamics.

References

Agarwal, Sumit, Tien Foo Sing, and Long Wang, 2018, Information asymmetries and learning in commercial real estate markets, *Available at SSRN 3022705* .

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis, 2013, Time use during the great recession, *American Economic Review* 103, 1664–96.

Aiello, Darren, Mark J Garmaise, and Taylor Nadauld, 2022, What problem do intermediaries solve?, *Available at SSRN 4105923* .

Andersen, Steffen, Cristian Badarinza, Lu Liu, Julie Marx, and Tarun Ramadorai, 2022, Reference dependence in the housing market, *American Economic Review* .

Badarinza, Cristian, and Tarun Ramadorai, 2018, Home away from home? foreign demand and london house prices, *Journal of Financial Economics* 130, 532–555.

Bhutta, Neil, Jesse Bricker, Andrew C Chang, Lisa J Dettling, Sarena Goodman, Joanne W Hsu, Kevin B Moore, Sarah Reber, Alice Henriques Volz, and Richard Windle, 2020, Changes in us family finances from 2016 to 2019: Evidence from the survey of consumer finances, *Federal Reserve Bulletin* 106.

Bhutta, Neil, Andreas Fuster, and Aurel Hizmo, 2019, Paying too much? borrower sophistication and overpayment in the us mortgage market, *Borrower Sophistication and Overpayment in the US Mortgage Market (July 19, 2019)* .

Campbell, John Y, and Joao F Cocco, 2007, How do house prices affect consumption? evidence from micro data, *Journal of monetary Economics* 54, 591–621.

Caplin, Andrew, and Mark Dean, 2015, Revealed preference, rational inattention, and costly information acquisition, *American Economic Review* 105, 2183–2203.

Carroll, Christopher D, Misuzu Otsuka, and Jiri Slacalek, 2011, How large are housing and financial wealth effects? a new approach, *Journal of Money, Credit and Banking* 43, 55–79.

Chinco, Alex, and Christopher Mayer, 2016, Misinformed speculators and mispricing in the housing market, *The Review of Financial Studies* 29, 486–522.

Chodorow-Reich, Gabriel, and Johannes Wieland, 2020, Secular labor reallocation and business cycles, *Journal of Political Economy* 128, 2245–2287.

Cvijanović, Dragana, and Christophe Spaenjers, 2021, "we'll always have paris": Out-of-country buyers in the housing market, *Management Science* 67, 4120–4138.

Davids, Allan, 2020, The cape of good homes: exchange rate depreciations, foreign demand and house prices, *Foreign Demand and House Prices (May 30, 2020)* .

Favilukis, Jack, and Stijn Van Nieuwerburgh, 2021, Out-of-town home buyers and city welfare, *The Journal of Finance* 76, 2577–2638.

Fisher, Lynn M, Lauren Lambie-Hanson, and Paul Willen, 2015, The role of proximity in foreclosure externalities: Evidence from condominiums, *American Economic Journal: Economic Policy* 7, 119–140.

Fu, Runshan, Ginger Zhe Jin, and Meng Liu, 2022, Human-algorithm interactions: Evidence from zillow. com, Technical report, National Bureau of Economic Research.

Gargano, Antonio, Marco Giacoletti, and Elvis Jarnecic, 2020, Local experiences, attention and spillovers in the housing market.", *Available at SSRN* .

Garmaise, Mark J, and Tobias J Moskowitz, 2004, Confronting information asymmetries: Evidence from real estate markets, *The Review of Financial Studies* 17, 405–437.

Giacoletti, Marco, and Christopher A Parsons, 2021, Propagating experiences by competing: Micro-level evidence from real estate, *Available at SSRN 3875752* .

Gorback, Caitlin S, and Benjamin J Keys, 2020, Global capital and local assets: House prices, quantities, and elasticities, Technical report, National Bureau of Economic Research.

Gupta, Arpit, 2019, Foreclosure contagion and the neighborhood spillover effects of mortgage defaults, *The Journal of Finance* 74, 2249–2301.

Guren, Adam M, 2018, House price momentum and strategic complementarity, *Journal of Political Economy* 126, 1172–1218.

Han, Lu, and Seung-Hyun Hong, 2022, Cash is king? understanding financing risk in housing markets, *Available at SSRN 3436797* .

Howard, Greg, 2020, The migration accelerator: Labor mobility, housing, and demand, *American Economic Journal: Macroeconomics* 12, 147–79.

Huo, Zhen, and José-Víctor Ríos-Rull, 2015, Tightening financial frictions on households, recessions, and price reallocations, *Review of Economic Dynamics* 18, 118–139.

Jia, Ning, Raven Molloy, Christopher Smith, and Abigail Wozniak, 2023, The economics of internal migration: Advances and policy questions, *Journal of Economic Literature* 61, 144–180.

Kurlat, Pablo, and Johannes Stroebel, 2015, Testing for information asymmetries in real estate markets, *The Review of Financial Studies* 28, 2429–2461.

Lambson, Val E, Grant R McQueen, and Barrett A Slade, 2004a, Do out-of-state buyers pay more for real estate? an examination of anchoring-induced bias and search costs, *Real Estate Economics* 32, 85–126.

Lambson, Val E, Grant R McQueen, and Barrett A Slade, 2004b, Do out-of-state buyers pay more for real estate? an examination of anchoring-induced bias and search costs, *Real Estate Economics* 32, 85–126.

Levitt, Steven D, and Chad Syverson, 2008, Market distortions when agents are better informed: The value of information in real estate transactions, *The Review of Economics and Statistics* 90, 599–611.

Li, Zhimin, Leslie Sheng Shen, and Calvin Zhang, 2020, Capital flows, asset prices, and the real economy: A " china shock" in the us real estate market, *International Finance Discussion Paper* .

Lutz, Chandler, and Ben Sand, 2022, Highly disaggregated land unavailability, *Available at SSRN 3478900* .

Reher, Michael, and Rossen I Valkanov, 2021, The mortgage-cash premium puzzle, *Available at SSRN 3751917* .

Rutherford, Ronald C, Thomas M Springer, and Abdullah Yavas, 2005, Conflicts between principals and agents: evidence from residential brokerage, *Journal of financial Economics* 76, 627–665.

Saiz, Albert, 2010, The geographic determinants of housing supply, *The Quarterly Journal of Economics* 125, 1253–1296.

Schubert, Gregor, 2021, House price contagion and u.s. city migration networks, Working paper, University of California, Los Angeles.

Sinai, Todd, and Nicholas Souleles, 2013, Can owning a home hedge the risk of moving?, *American Economic Journal: Economic Policy* 5, 282–312.

Stroebel, Johannes, and Joseph Vavra, 2019, House prices, local demand, and retail prices, *Journal of Political Economy* 127, 1391–1436.

Turnbull, Geoffrey K, and Casey F Sirmans, 1993, Information, search, and house prices, *Regional Science and Urban Economics* 23, 545–557.

Wolff, Edward N, 2022, Heterogenous rates of return on homes and other real estate: Do the rich do better? do blacks do worse?, *NBER Working Paper 30543* .

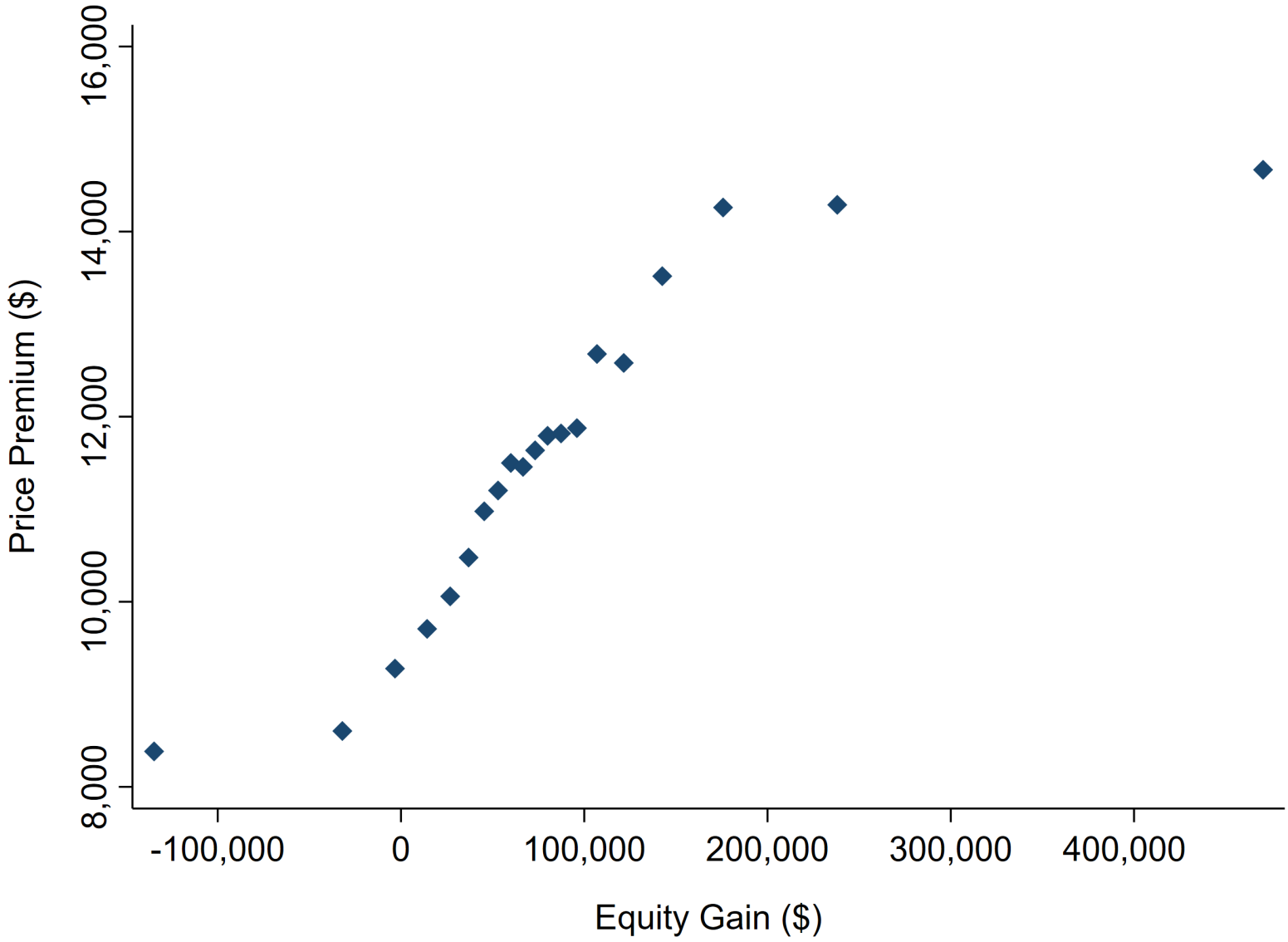

Figure 1. Overpayment and Equity Gains This figure presents the average $PricePremium_{i,t}$ (as calculated in Equation 2) in 20 bins of $Equity\,Gain_{i,t}$ (as calculated in Equation 1), after absorbing an interacted fixed effect for the purchase property zip and transaction year following the specification found in Equation 8.

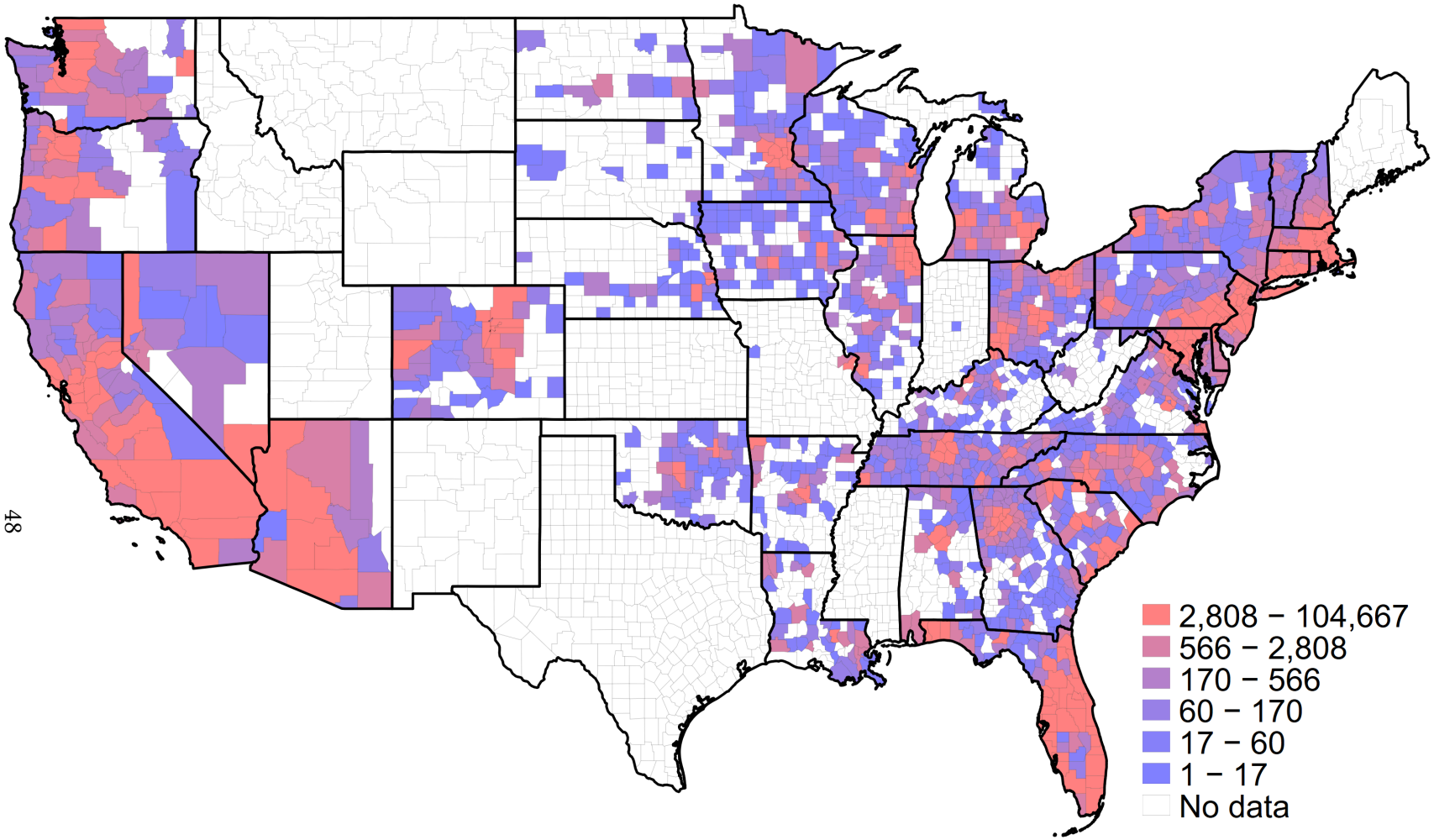

Figure 2. County-Level Observation Count This figure presents a county-level representation of the location of the purchased property for the 3.1 million household moves in our main regression sample.

Table 1
Summary Statistics

This table reports summary statistics related to the full sample of matched household moves. Panel A reports the characteristics associated with the sold property and the sale transaction. Panel B reports the characteristics of the move itself. Panel C reports the characteristics associated with the purchased property. Panel D reports results related to the neighborhood (at various levels) surrounding the purchased property.

Variable	Obs.	Mean	Std. Dev.	Q5	Q25	Q50	Q75	Q95
Panel A: Sold Property								
Equity Gain	3,103,108	86,244	142,872	-35,000	16,000	50,000	116,000	333,100
Years Lived in Home	3,103,108	6.58	5.27	0.90	2.84	5.13	9.10	16.93
Realized Return	3,040,584	9.6%	25.1%	-1.9%	1.7%	4.6%	9.4%	28.1%
Zip ΔHPI over Holding Period	3,103,108	65,304	100,434	-44,000	10,000	43,550	98,600	245,583
Zip Median HP at Purchase	3,103,108	215,038	133,045	83,089	130,000	176,500	257,939	473,000
Square Footage	2,854,888	2,024.81	1,063.63	931	1,331	1,776	2,424	3,862
House Age	2,955,949	29.51	24.86	3	10	22	44	81
Number of Bedrooms	2,337,181	3.21	0.89	2	3	3	4	5
Number of Bathrooms	2,114,884	2.51	0.93	1	2	2	3	4
Panel B: The Move								
Transaction Date	3,103,108	05/13/2011	6 Yrs 11 Mos	02/10/2000	02/22/2005	09/17/2012	08/09/2017	09/04/2020
Move Distance (Miles)	3,054,146	330.60	616.06	0.53	3.33	12.20	373.11	1,965.03
In-Zip Move	3,103,108	18.3%						
Out-of-Zip, In-County Move	3,103,108	30.5%						
Out-of-County, In-State Move	3,103,108	19.4%						
Out-of-State Move	3,103,108	31.7%						
Panel C: Purchased Property								
Transaction Price	3,103,108	372,314	268,691	94,000	197,304	301,000	460,000	885,000
Price Premium	3,103,108	11,538	76,607	-77,461	-17,528	4,296	31,210	128,469
Price Premium Percent	3,103,108	-1.9%	37.3%	-32.3%	-6.4%	1.5%	9.7%	29.9%
Future Realized Return	1,480,493	11.8%	41.5%	-6.5%	0.4%	3.5%	8.9%	42.9%
Cash Only	3,103,108	18.0%						
Square Footage	2,866,987	2,318.19	1,154.00	1,030	1,542	2,089	2,798	4,319
House Age	2,891,616	25.13	24.07	0	7	18	38	75
Number of Bedrooms	2,331,311	3.41	0.94	2	3	3	4	5
Number of Bathrooms	2,098,338	2.76	0.99	1	2	3	3	4
Panel D: Purchase Neighborhood								
Zip Median House Price	3,101,240	279,165	169,713	103,500	165,000	236,000	340,000	612,500
Average Sales Price, 1/2-Mile, 0-180 Days	3,013,245	331,264	205,802	112,768	191,768	277,691	409,421	740,554
Historical Average Sales Price, 1/2-Mile, 0-180 Days	3,016,753	321,483	200,519	109,438	185,682	269,311	397,183	720,000
Historical Transaction Volume, 1/2-Mile, 0-90 Days	3,076,397	18.23	23.00	1	6	12	23	52

Table 2

Equity Gains Cause Buyers to Overpay for Their Next House

Column (1) estimates Equation 8 reporting the correlation between $PricePremium_{i,t}$ (as calculated in Equation 2) and $Equity\,Gain_{i,t}$ (as calculated in Equation 1). Column (2) estimates Equation 11 and reports the causal effect of $Equity\,Gain_{i,t}$, instrumented with $\Delta HPI^{S}_{z,t}$ (as calculated in Equation 9), on $PricePremium_{i,t}$. Column (3) replicates column (2) but instead reports the causal effect of $Equity\,Gain_{i,t}$ on the fraction of the purchased property's transaction price that is represented by $PricePremium_{i,t}$. Column (4) reports the OLS estimate recovered from regressing the future realized return on the house being purchased on to $PricePremium_{i,t}$. Results relate to the full sample of matched household moves. All dollar amounts are in $100,000 units. Controls and fixed effects are included as indicated. The Adjusted Reported R^2 is reported for all OLS specifications and the Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. t-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Price Premium, House Being Purchased		Price Premium Percent, House Being Purchased	Annualized Future Realized Return, House Being Purchased	Price Premium, House Being Purchased
	OLS	2SLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)
Equity Gain, House Being Sold	0.0106**	0.0785***	0.0205***		0.0741***
	(2.36)	(6.21)	(4.31)		(5.30)
Price Premium, House Being Purchased				-0.125***	
				(-17.67)	
Purchase Zip Lower Price Than Sale Zip, Indicator					0.0246***
					(8.41)
Equity Gain, × Purchase Zip Lower Price					-0.0109***
					(-4.67)
Zip Median HP of Sold Home at Purchase		X	X	X	X
Five Interacted Property Sold Characteristic FE		X	X	X	X
Years Lived in Sold Home FE		X	X	X	X
Sale Zip × Transaction Year FE		X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X	X
Instrumental Variable		Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period		Zip ΔHPI over Sold Home Holding Period and Interaction
Sample	Full	Full	Full	Full	Out-of-Zip Moves
Observations	3,103,108	3,103,108	3,103,108	1,412,419	2,514,792
Adj. R^2	0.018			0.244	
Weak ID KP F Stat		241.4	241.4		118.7

Table 3
Overpayment is Decreasing in Information

This table reports results related to estimations of specifications similar to that of Equation 11 and column (2) of Table 2. Column (1) includes an additional interaction between $Equity\ Gain_{i,t}$ and a measure of the local purchase neighborhood's historical transaction volume, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and historical transaction volume, as well as a level of transaction volume term. Column (2) includes an additional interaction between $Equity\ Gain_{i,t}$ and a measure of the standard deviation of historical price premia in the local purchase neighborhood, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and that standard deviation, as well as a level of the standard deviation term. Column (3) includes an additional interaction between $Equity\ Gain_{i,t}$ and a measure of the standard deviation of historical transaction prices in the local purchase neighborhood, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and that standard deviation, as well as a level of the standard deviation term. Column (4) includes an additional interaction between $Equity\ Gain_{i,t}$ and a continuous measure of the moving household's years lived in the sold home, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and the years in sold home. Column (5) includes an additional interaction between $Equity\ Gain_{i,t}$ and an indicator for whether the moving household's years lived in the sold home is above the median observed in the sample, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and the years in sold home above median indicator. The level terms in columns (4) and (5) are absorbed by the fixed effects. Columns (1) through (3) report results relating to the full sample of matched household moves, further restricted to observations with valid measurements of historical transaction volume, and standard deviations of price premia and transaction prices, respectively. Columns (4) and (5) report results relating to the sub-set of matched household moves where both the sold and purchased properties are in the same county ("local moves"). All dollar amounts are in $100,000 units. Standard deviations measures are divided by 100. Controls and fixed effects are included as indicated. The Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. Reported t-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Price Premium, House Being Purchased				
	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)
Equity Gain,	0.0871***	0.0578***	0.0532***	0.114***	0.0967***
House Being Sold	(6.50)	(3.66)	(3.58)	(3.46)	(2.95)
Historical Transaction Volume in Purchase Area,	-0.000358**				
1/2-Mile, 0-90 Days	(-2.40)				
Historical Transaction Volume	-0.000340***				
× Equity Gain	(-4.14)				
Historical Price Premium St. Dev. in Purchase Area,		0.0000933***			
1/2-Mile, 0-90 Days		(7.81)			
Historical Price Premium St. Dev.		0.0000167***			
× Equity Gain		(4.73)			
Historical Price St. Dev. in Purchase Area,			0.00000947***		
1/2-Mile, 0-90 Days			(5.94)		
Historical Price St. Dev.			0.0000302***		
× Equity Gain			(7.47)		
Years Lived in Sold Home				-0.00392***	
× Equity Gain				(-6.03)	
Above Median Years Lived in Sold Home (Ind.)					-0.0266***
× Equity Gain					(-3.15)
Zip Median HP of Sold Home at Purchase	X	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X	X
Sale Zip × Transaction Year FE	X	X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X	X
Instrumental Variables	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction
Sample	Full	Full	Full	Local Moves	Local Moves
Observations	3,074,934	2,692,174	2,800,884	1,444,434	1,444,434
Weak ID KP F Stat	116.5	101.2	106.3	67.52	68.15

Table 4

Equity Gains Cause Neighborhood Price Spillovers

This table reports results related to estimations of specification similar to that of Equation 11. The left-hand side variable is, however, instead a measure of the average sales price in the immediate neighborhood surrounding the purchased property in the time period and geography relevant to the respective column, $\text{NeighborhoodPrice}_{i,N,t,T}$ (as calculated in Equation 12). Columns (1), (3), (5), and (7) measure the average price across transactions occurring between 0-180 days after the purchase transaction, while columns (2), (4), (6), and (8) measure across 0-360 days after the purchase transaction. Columns (1) and (2) average transactions across a 1/2-mile radius circle surrounding the purchase property while columns (3) and (4), (5) and (6), and (7) and (8) average across 1-mile, 1 1/2-mile, and 2-mile radii, respectively. In all cases the average price is measured excluding the purchase transaction. Results relate to the full sample of matched household moves, subject to there being at least one subsequent transaction in the time period and geography relevant to the respective column. All dollar amounts are in $100,000 units. Controls and fixed effects are included as indicated, with "at matched distance" referring to the geography relevant to the respective column. The Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. Reported t-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Average Sales Price within 1/2-Mile of Home Purchase		Average Sales Price within 1-Mile of Home Purchase		Average Sales Price within 1 1/2-Mile of Home Purchase		Average Sales Price within 2-Miles of Home Purchase	
	0-180 Days	0-360 Days	0-180 Days	0-360 Days	0-180 Days	0-360 Days	0-180 Days	0-360 Days
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Equity Gain, House Being Sold	0.0882***	0.0682**	0.0822***	0.0527**	0.0681***	0.0409*	0.0625***	0.0371
	(3.25)	(2.29)	(3.49)	(2.13)	(3.03)	(1.78)	(2.86)	(1.67)
Four Quarterly Historical Average Prices, At Matched Distance	X	X	X	X	X	X	X	X
Zip Median HP of Sold Home at Purchase	X	X	X	X	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X	X	X	X	X
Sale Zip × Transaction Year FE	X	X	X	X	X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X	X	X	X	X
Instrumental Variable	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period
Observations	2,735,407	2,739,485	2,969,898	2,971,202	3,028,482	3,028,977	3,050,266	3,050,497
Weak ID KP F Stat	231.8	231.9	248.9	249.5	256.2	256.6	258.7	258.6

Table 5
Spillovers Decreasing in Information

This table combines the specification of Table 4 column 1 with the interaction specifications of Table 3. Columns (1) through (3) report results relating to the full sample of matched household moves, further restricted to observations with valid measurements of historical transaction volume, and standard deviations of price premia and transaction prices, respectively. Columns (4) and (5) report results relating to the sub-set of matched household moves where both the sold and purchased properties are in the same county ("local moves"). All columns are subject to there being at least one subsequent transaction in the time period and geography relevant to the respective column. All dollar amounts are in $100,000 units. Standard deviations measures are divided by 100. Controls and fixed effects are included as indicated, with "at matched distance" referring to the geography relevant to the respective column. The Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. Reported *t*-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Average Sales Price, within 1/2-Mile and 0-180 Days of Home Purchase				
	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)
Equity Gain,	0.0969***	0.0770**	0.0695**	0.170***	0.164***
House Being Sold	(3.55)	(2.52)	(2.35)	(3.46)	(3.32)
Historical Transaction Volume in Purchase Area,	-0.00117***				
1/2-Mile, 0-90 Days	(-5.09)				
Historical Transaction Volume	-0.000158**				
× Equity Gain	(-2.73)				
Historical Price Premium St. Dev. in Purchase Area,		0.0000375***			
1/2-Mile, 0-90 Days		(6.28)			
Historical Price Premium St. Dev.		0.00000433			
× Equity Gain		(1.39)			
Historical Price St. Dev. in Purchase Area,			-0.0000577***		
1/2-Mile, 0-90 Days			(-15.01)		
Historical Price St. Dev.			0.00000435***		
× Equity Gain			(2.96)		
Years Lived in Sold Home				-0.00344***	
× Equity Gain				(-3.54)	
Above Median Years Lived in Sold Home (Ind.)					-0.0380**
× Equity Gain					(-2.66)
Four Quarterly Historical Average Prices, At Matched Distance	X	X	X	X	X
Zip Median HP of Sold Home at Purchase	X	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X	X
Sale Zip × Transaction Year FE	X	X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X	X
Instrumental Variables	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction
Sample	Full	Full	Full	Local Moves	Local Moves
Observations	2,735,407	2,594,042	2,670,788	1,285,159	1,285,159
Weak ID KP *F* Stat	115.4	106.8	109.8	61.59	62.55

Table 6
Local Movers vs. Non-Local Movers

This table reports results related to estimations of specification similar to that of Equation 11. Columns (1) and (2) are similar to column (2) of Table 2 and columns (3) and (4) are similar to column (1) of Table 4. Columns (1) and (3) include additional interactions between $Equity\ Gain_{i,t}$ and indicators for the distance, in terms of political boundaries, between the sold and the purchased properties, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and the relevant indicator, as well as the respective indicators for the move distance. The left out indicator level is for In-Zip moves. Columns (2) and (4) include an additional interaction between $Equity\ Gain_{i,t}$ and the distance, in miles, between the sold and the purchased properties, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and the move distance, as well as a level of move distance term. Results relate to the full sample of matched household moves, with columns (2) and (4) further restricted to observations with valid measurements of distance between the sold and purchased properties, and columns (3) and (4) subject to there being at least one subsequent transaction in the time period and geography relevant to the respective column. All dollar amounts are in $100,000 units. Controls and fixed effects are included as indicated, with "at matched distance" referring to the geography relevant to the respective column. The Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. Reported t-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Price Premium, House Being Purchased		Average Sales Price, within 1/2-Mile and 0-180 Days of Home Purchase	
	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)
Equity Gain, House Being Sold	0.109***	0.0741***	0.0972***	0.0576**
	(8.03)	(5.28)	(3.49)	(2.38)
Out-of-Zip, In-Count, Move Indicator	0.0218***		0.00408	
	(3.43)		(1.33)	
Out-of-County, In-State Move Indicator	0.0370***		0.00437	
	(3.20)		(1.09)	
Out-of-State, Move Indicator	0.0456***		0.00582	
	(2.82)		(0.87)	
Out-of-Zip, In-County Move Indicator × Equity Gain	-0.0206***		-0.00600**	
	(-6.40)		(-2.71)	
Out-of-County, In-State Move Indicator × Equity Gain	-0.0329***		-0.0108***	
	(-6.92)		(-3.62)	
Out-of-State Move Indicator × Equity Gain	-0.0507***		-0.0144***	
	(-8.58)		(-3.78)	
Ratio of Sold to Purchase Property Zip Code House Prices > 1		0.0244***		0.00494**
		(8.38)		(2.24)
Zip Code House Price Ratio > 1 × Equity Gain		-0.0109***		-0.00322
		(-4.66)		(-1.39)
Four Quarterly Historical Average Prices, At Matched Distance			X	X
Zip Median HP of Sold Home at Purchase	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X
Sale Zip × Transaction Year FE	X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X
Instrumental Variables	Zip ΔHPI over Sold Home Holding Period and Interactions	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interactions	Zip ΔHPI over Sold Home Holding Period and Interaction
Sample	Full	Out-of-Zip Moves	Full	Out-of-Zip Moves
Observations	3,103,108	2,517,256	2,735,407	2,212,386
Weak ID KP F Stat	59.12	119.6	56.89	121.7

Table 7

Equity Gains Cause Movers to Spend More On Their Next House

This table reports results related to estimations of specification similar to that of Equation 11. The left-hand side variables are, however, instead various measures related to the overall characteristics of the transaction and purchased property. Columns (1) through (7) utilize, as their left-hand side variables, the transaction price, the quality and neighborhood trend adjusted value of the property (transaction price - $PricePremium_{i,t}$), the median house price of the zip code where the purchased property is located, the log square footage of the purchased property, the log years since the purchased property was built, the number of bedrooms in the purchased property, and the number of bathrooms in the purchased property, respectively. Results relate to the full sample of matched household moves, with columns (3) through (7) further restricted to observations with valid measurements for the respective outcome variables. All dollar amounts are in $100,000 units. Controls and fixed effects are included as indicated. The Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. Reported t-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Transaction Price House Being Purchased	Adjusted Value House Being Purchased	Zip Median HP, House Being Purchased	Log House Size (Sq. Ft.), House Being Purchased	Log House Age (Years), House Being Purchased	Number of Bedrooms, House Being Purchased	Number of Bathrooms, House Being Purchased
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Equity Gain, House Being Sold	0.871***	0.793***	1.234***	0.0591***	0.0329*	0.0808***	0.118***
	(22.31)	(20.68)	(21.22)	(7.87)	(1.75)	(6.44)	(6.62)
Zip Median HP of Sold Home at Purchase	X	X	X	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X	X	X	X
Sale Zip × Transaction Year FE	X	X	X	X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X	X	X	X
Instrumental Variable	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period
Observations	3,103,108	3,103,108	3,100,649	2,857,897	2,600,440	2,304,821	2,068,888
Weak ID KP F Stat	241.4	241.4	239.9	223.3	219.2	253.3	210.9

Table 8

Commuting Zone Level Equity Inflows Increase House Prices

This table reports results related to estimations of specification similar to that of Equation 13. The left-hand side variable is a measure of the log change in the FHFA House Price Index (aggregated to the CZ-level from population-weighted county changes). The key right-hand variable of interest is the estimated total positive housing equity gain of in-migrants per local household $eg_{i,t}^{\text{non-local}}$ (as calculated in Equation 14). Columns estimated with 2SLS use the instrumental variable $\widetilde{eg}_{i,t}^{\text{non-local}}$ as constructed in equation 15, using data only for movers coming from outside a 50 mile radius around the commuting zone. Results relate to the panel of CZ-year cells for 2006-2018 for which the required data is available. The Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. Controls and fixed effects are included as indicated. Reported t-statistics in parentheses are heteroskedasticity-robust and clustered at both the county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Predicted Log Out-of-CZ Equity Inflow, Per Capita	100 × Δ Log House Price Index, CZ-Year					
	First Stage	OLS	IV	IV	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Out-of-CZ Equity Inflow Instrument, Per Capita	0.666***						
	(5.73)						
Predicted Log Out-of-CZ Equity Inflow, Per Capita		0.699***	4.295***	4.695**	4.301***	3.830**	0.632
		(3.73)	(3.40)	(2.63)	(3.34)	(3.02)	(0.38)
Predicted Log Out-of-CZ Equity Inflow, × 2nd Quint. Land Constraints							0.611
							(1.15)
Predicted Log Out-of-CZ Equity Inflow, × 3rd Quint. Land Constraints							2.835**
							(2.28)
Predicted Log Out-of-CZ Equity Inflow, × 4th Quint. Land Constraints							3.534**
							(2.41)
Predicted Log Out-of-CZ Equity Inflow, × 5th Quint. Land Constraints							4.673**
							(2.24)
Commuting Zone FE	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X
Log Predicted Out-of-CZ In-Mig., HHs				X			
Log CZ Home Sales, per HH					X		
Industry Wage Shock						X	
Industry Wage Shock×Land Constraint						X	
Instrumental Variable			Log Exp. Out-of-CZ Equity Inflow, Per Capita	Log Exp. Out-of-CZ Equity Inflow, Per Capita	Log Exp. Out-of-CZ Equity Inflow, Per Capita	Log Exp. Out-of-CZ Equity Inflow, Per Capita	Log Exp. Out-of-CZ Equity Inflow, Per Capita
Observations	4,320	4,327	4,320	4,320	4,320	3,947	4,320
Adj. R^2	0.813	0.485					
Weak ID KP F Stat			32.85	22.87	32.94	29.23	4.888

Appendix

A.1 Aggregate effect estimation details

This section details the methodology for constructing the measures of cross-CZ housing capital flows and instruments.

A.1.1 Estimating housing equity gain inflows

The variable of interest for determining the impact of housing capital flows is the total inflow of equity gains into a CZ as a result of inbound migration. However, this quantity cannot be directly measured in our property-level data as we do not necessarily have full coverage of all transactions in a CZ. Instead, we combine our data on average equity gains among buyers with data on total movers between counties from the Internal Revenue Service's Statistics of Income (SOI) database. Specifically, we use the ZTRAX data to estimate the average equity gain among observed out-of-CZ buyers in destination c as:

$$\overline{\text{EquityGain}}^{\text{non-local}}_{c,t} = \left(\frac{\sum_i^{N^{\text{ZTRAX}}_{c,t}} \text{EquityGain}_{i,t,c} \times \mathbb{1}[\text{Out-of-CZ Origin}]_{i,t,c}}{\sum_i^{N^{\text{ZTRAX}}_{c,t}} \mathbb{1}[\text{Out-of-CZ Origin}]_{i,t,c}} \right),$$

where we define $\mathbb{1}[\text{Out-of-CZ Origin}]_{i,t,c}$ to indicate movers that are coming from a CZ for which each county centroid is at least 50 miles away from each county in the destination CZ. We impose this minimum distance between geographies to account for both the possibility of CZ boundaries mismeasuring housing markets as well as any identification issues arising from origin geographies being subject to the same idiosyncratic local housing market shocks as the destination CZ.

Second, we scale this average equity gain $\overline{\text{EquityGain}}^{\text{non-local}}$ into an estimate of total equity

gains from all non-local movers per local household by calculating[27]

$$eg_{ct}^{\text{non-local}} = \overline{\text{EquityGain}}_{ct}^{\text{non-local}} \times \frac{\text{OutOfCZMovers}_{ct}^{\text{IRS}}}{\text{ResidentHHs}_{c,2006}^{\text{IRS}}}, \tag{A.1.1}$$

where $\text{OutOfCZMovers}_{ct}^{\text{IRS}}$ and $\text{ResidentHHs}_{c,2006}^{\text{IRS}}$ are the IRS estimates of total household inflows into the commuting zone c from origin CZs at least 50 miles away and total local households residing in the CZ in 2006.[28]

A.1.2 Housing equity gain inflow instrument

We construct an instrument that captures plausibly exogenous equity gain flows between CZs by estimating the components corresponding to the following *exogenous* predicted equity gain inflows from out-of-area origins for each CZ:

$$\widetilde{eg}_{ct}^{\text{non-local}} = \frac{\text{[Predicted Avg. Equity Gain for Movers]} \times \text{[Predicted Out-of-Area Inmigration]}}{\text{ResidentHHs}_{c,2006}^{\text{IRS}}}.$$

First, to construct a plausibly exogenous flow of *predicted movers* for each origin-destination CZ pair, we use historical migration links as predictors of contemporaneous migration flows (Howard, 2020; Schubert, 2021). That is, the predicted number of movers from k to c is constructed from IRS migration data as a "shift-share" instrument of the form

$$\widehat{Mig}_t^{k\to c} = \left(\frac{MigShare_{`90-`99}^{k\to c}}{1 - MigShare_{`90-`99}^{k\to c}} \right) \times OutMig_t^{k\to \neg c},$$

[27] Note that we do not scale the total inflow by the share of homeowners among movers, as precise data on homeownership status for origin-destination pair movers is not available on an annual basis (e.g. the ACS data is not granular enough to compute reliable averages for such a small subgroup). However, note that this likely does not affect the estimation much as the log specification with CZ fixed effects would control for any constant differences in homeowner shares among movers into different CZs.

[28] 2006 is the first year of our panel for the aggregate effects estimation.

where $MigShare_{90-99}^{k \to c}$ is the average share of outflows from CZ k that go to CZ c during 1990-1999, and $OutMig_t^{k \to \neg c}$ is the total outflow of migrants in year t to all locations *other* than CZ c that are at least 50 miles away.[29] The intuition for using this measure of expected migration flows is that it captures the degree to which the characteristics of CZ k are prompting outflows in period t, and the average share of those outflows that would be expected to go towards CZ c—for instance, because of low migration costs or historic family connections between these locations—but without using any period t variation from CZ c itself that might potentially be correlated with local house price trends in c. Moreover, when aggregating these predicted flows from other CZs, we will again omit any flows from CZs that are less than 50 miles away, in order to eliminate bias from the migration dynamics of CZs that may have overlapping local housing and labor markets with the destination CZ of interest.

Similarly, the predicted average equity gain among people moving into CZ c is constructed without using data on the characteristics of contemporaneous flows from k to c, which might be selected based on destination characteristics. Instead, we assume that movers from high equity gain locations are also generally more likely to have higher equity gains. Thus, destination CZ c is expected to have higher average equity gains among its incoming households if it has higher historic migration exposure to CZs that are experiencing high equity gains among their house sellers in period t:

$$\overline{EquityGain}_{ct}^{\text{origins}} = \sum_{k\,:\,dist(\text{CZ(k)},\text{CZ(c)})>50\text{mi.}} \left(\frac{\widehat{Mig}_t^{k \to c}}{\sum_{k\,:\,dist(\text{CZ(k)},\text{CZ(c)})>50\text{mi.}} \widehat{Mig}_t^{k \to c}} \right) \times \overline{EquityGain}_{kt}$$

Here, $\overline{EquityGain}_{kt}$ is the average gross equity gain among *all* sellers in CZ k in period t in the ZTRAX data, without netting out equity losses (under the assumption that one mover's lack of

[29] The reason for dividing by $(1 - MigShare_{90-99}^{k \to c})$ is to re-scale the leave-one-out outflows to the expected magnitude if CZ c flows of historic proportions had been included.

funds would not negate another mover's overpayment). Moreover, we again omit the housing equity gains in any origin CZs that are within 50 miles of CZ c to avoid bias from common shocks to nearby local housing markets. Combining these measures as shown above, our instrument for equity gains (per local household) flowing into CZ c from other CZs is then

$$\widetilde{eg}_{ct}^{\text{non-local}} = \underbrace{\left(\sum_{k\,:\,dist(\text{CZ(k)},\text{CZ(c)})>50\text{mi.}} \widehat{Mig}^{k\to c} \right)}_{\text{Predicted Out-of-area Inmigration}} \times \overline{EquityGain}_{ct}^{\text{origins}} \times \frac{1}{\text{ResidentHHs}_{c,2006}^{\text{IRS}}} \qquad \text{(A.1.2)}$$

This shift-share approach results in an instrument for the total equity gain inflow into CZ c in year t that does not use any contemporaneous information on the actual number or characteristics of movers with equity gains coming into c and thereby eliminates any correlation between the contemporaneous destination choices of movers and local house price dynamics in c. Instead, the instrument relies on variation across CZs in their historical exposure to migration from particular distant CZs, in conjunction with variation over time in the average equity gains in those distant CZs.

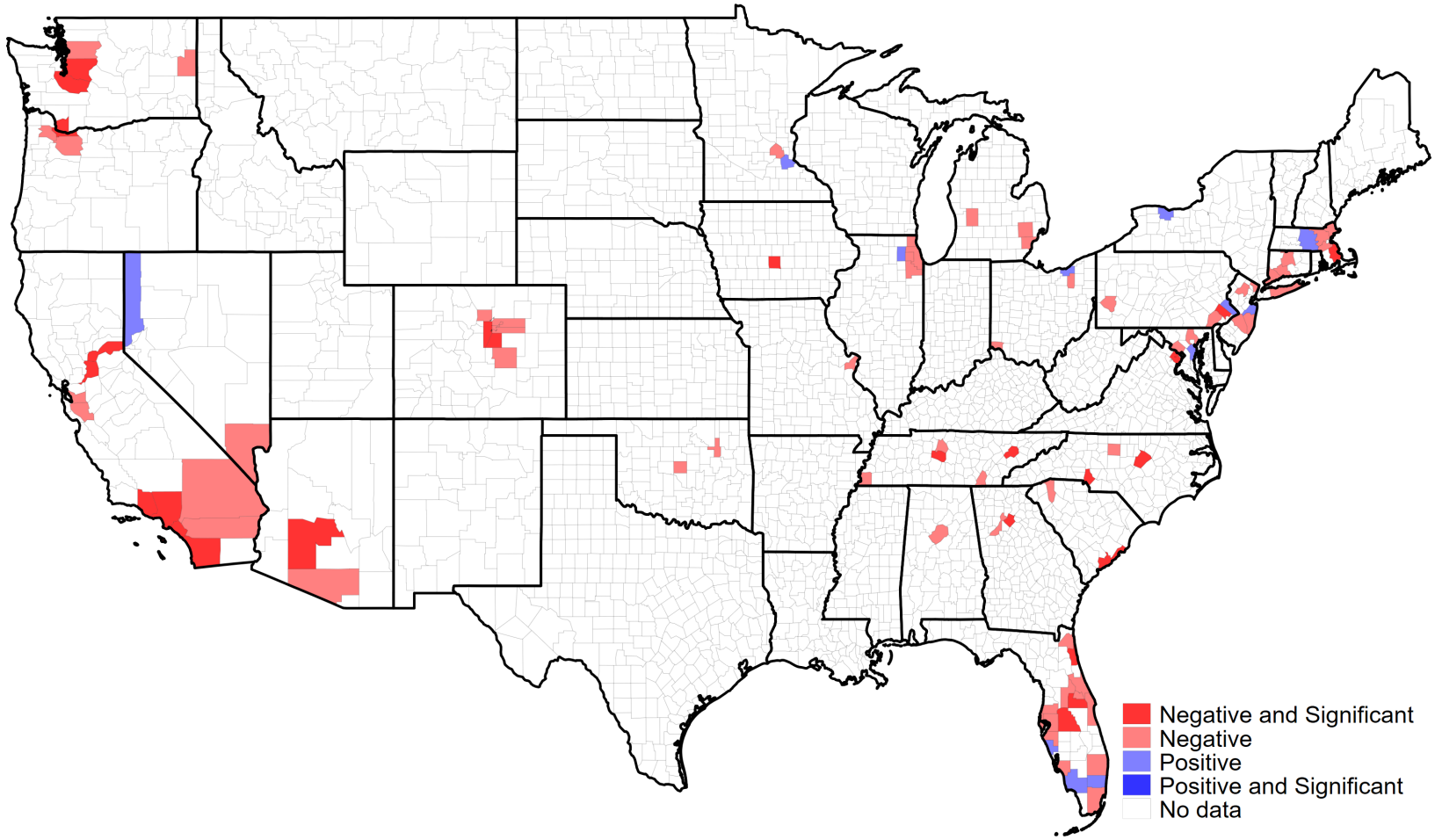

Figure A.1. Effect of Equity Gain on Price Premium for Non-Locals Relative to Locals for the 100 Largest Counties This figure presents estimates conceptually similar to that of column (2) of Table 6. Collapsing the four-fold distinction in moves across political boundary distances to simply an indicator for a non-local move (strictly out-of-county), the map plots the sign and significance (at a 95% level) of the coefficient on the interaction between that non-local move indicator and equity gains, instrumented with an interaction between the indicator and our standard instrument. Fixed effects and clustering differ from the reference specification—the sale zip by year and purchase zip by year fixed effects are replaced with a single transaction year fixed effect and the purchase zip clustering dimension is removed.

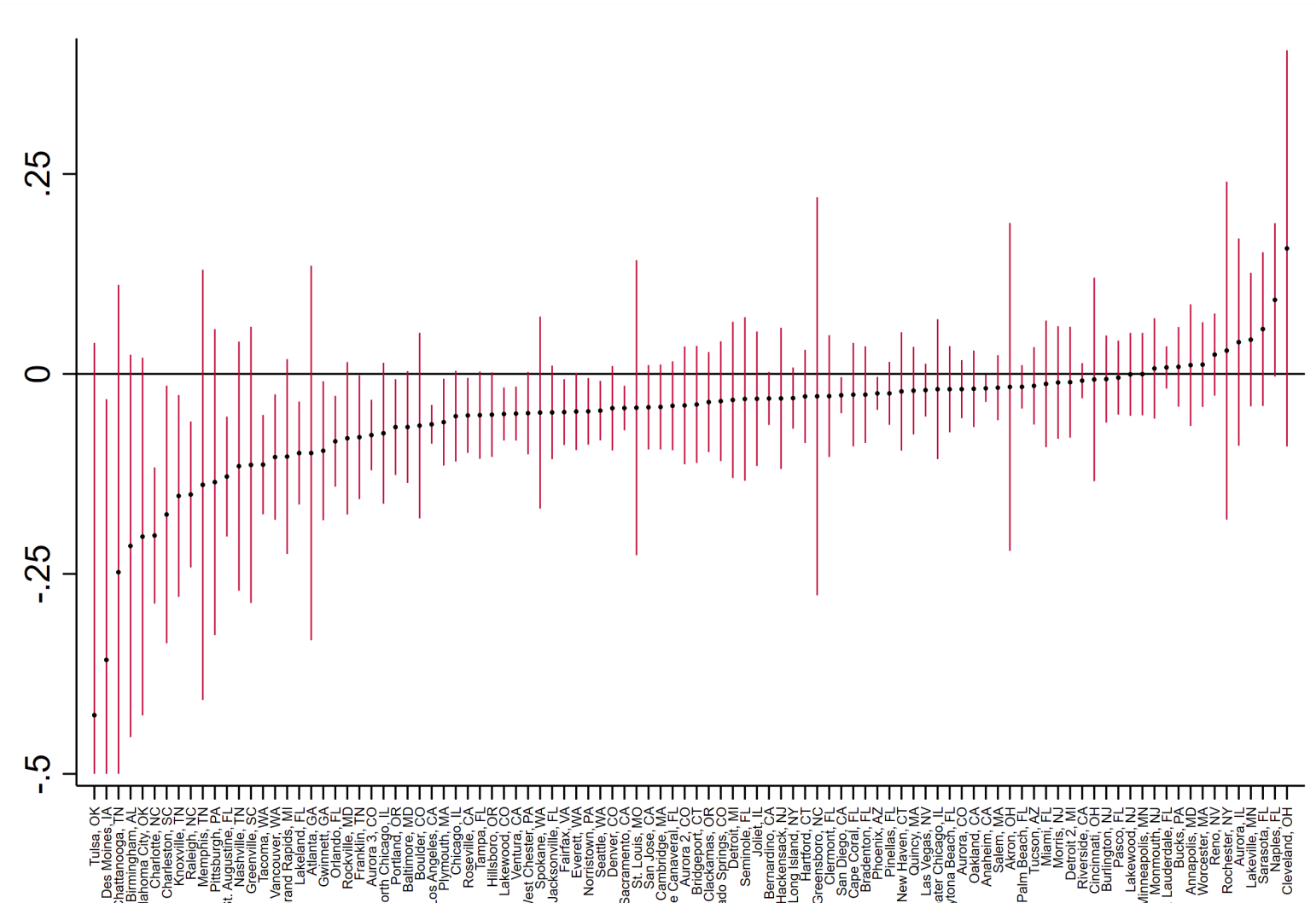

Figure A.2. Effect of Equity Gain on Price Premium for Non-Locals Relative to Locals for the 100 Largest Counties This figure presents the coefficient and significance data underlying the map in Appendix Figure A.1. The left-most three lower confidence intervals are truncated at -0.5 for aesthetic purposes. The names along the x-axis describe county-level bins, where names have been chosen to subjectively refer to the most salient aspect of the county's contents.

Table A.1

Housing Market Sophistication

This table regresses the realized return a household experiences on a home being purchased on the realized return experienced on the home being sold as evidence in support of the existence of housing market sophistication that could be contaminating naïve estimates of the relationship between equity gain and the price premium. Reported t-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Annualized, Future Realized Return, House Being Purchased
	OLS
	(1)
Annualized Realized Return, House Being Sold	0.332***
	(15.22)
Zip Median HP of Sold Home at Purchase	X
Five Interacted Property Sold Characteristic FE	X
Years Lived in Sold Home FE	X
Sale Zip × Transaction Year FE	X
Purchase Zip × Transaction Year FE	X
Observations	1,371,198
Adj. R^2	0.227

Table A.2

Overpayment and Subsequent Performance - Renovations

This table reports results similar to column (4) of Table 2 in differing sub-samples. Column (1) through (4) report results estimating just on the sample where the mover owned the purchased house for less than a year, two years, five years, and ten years, respectively. Column (5) is run on the full sample (where a subsequent sale of the purchased property is observed in our data) and is identical to column (4) of Table 2. All dollar amounts are in $100,000 units. Controls and fixed effects are included as indicated. The Adjusted Reported R^2 is reported for all OLS specifications and the Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. *t*-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Annualized Future Realized Return, House Being Purchased				
	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)
Price Premium,	-0.618***	-0.423***	-0.218***	-0.152***	-0.125***
House Being Purchased	(-9.29)	(-29.72)	(-26.82)	(-20.34)	(-17.67)
Zip Median HP of Sold Home at Purchase	X	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X	X
Sale Zip × Transaction Year FE	X	X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X	X
Sample	Years Held<1	Years Held<2	Years Held<5	Years Held<10	Full
Observations	5,981	136,876	672,321	1,112,709	1,412,419
Adj. R^2	-10.136	0.291	0.282	0.254	0.244

Table A.3
Overpayment as Percent of Price is Decreasing in Information

This table reports results related to estimations of specifications similar to that of Equation 11 and column (3) of Table 2. Column (1) includes an additional interaction between $Equity\ Gain_{i,t}$ and a measure of the local purchase neighborhood's historical transaction volume, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and historical transaction volume, as well as a level of transaction volume term. Column (2) includes an additional interaction between $Equity\ Gain_{i,t}$ and a measure of the standard deviation of historical price premia in the local purchase neighborhood, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and that standard deviation, as well as a level of the standard deviation term. Column (3) includes an additional interaction between $Equity\ Gain_{i,t}$ and a measure of the standard deviation of historical transaction prices in the local purchase neighborhood, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and that standard deviation, as well as a level of the standard deviation term. Column (4) includes an additional interaction between $Equity\ Gain_{i,t}$ and a continuous measure of the moving household's years lived in the sold home, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and the years in sold home. Column (5) includes an additional interaction between $Equity\ Gain_{i,t}$ and an indicator for whether the moving household's years lived in the sold home is above the median observed in the sample, instrumented with an interaction between $\Delta\text{HPI}^{S}_{z,t}$ and the years in sold home above median indicator. The level terms in columns (4) and (5) are absorbed by the fixed effects. Columns (1) through (3) report results relating to the full sample of matched household moves, further restricted to observations with valid measurements of historical transaction volume, and standard deviations of price premia and transaction prices, respectively. Columns (4) and (5) report results relating to the sub-set of matched household moves where both the sold and purchased properties are in the same county ("local moves"). All dollar amounts are in $100,000 units. Standard deviations measures are divided by 100. Controls and fixed effects are included as indicated. The Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. Reported t-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Price Premium Percent, House Being Purchased	Price Premium Percent, House Being Purchased		Price Premium Percent, House Being Purchased	
	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)
Equity Gain, House Being Sold	0.0225***	0.0157***	0.0174***	0.0347***	0.0341***
	(4.49)	(3.07)	(3.39)	(3.27)	(3.40)
Historical Transaction Volume in Purchase Area, 1/2-Mile, 0-90 Days	0.000410***				
	(5.26)				
Historical Transaction Volume × Equity Gain	-0.000146***				
	(-5.53)				
Historical Price Premium St. Dev. in Purchase Area, 1/2-Mile, 0-90 Days		-0.0000108*			
		(-1.99)			
Historical Price Premium St. Dev. × Equity Gain		0.00000487***			
		(3.96)			
Historical Price St. Dev. in Purchase Area, 1/2-Mile, 0-90 Days			-0.00000609***		
			(-4.40)		
Historical Price St. Dev. × Equity Gain			0.00000162***		
			(3.59)		
Years Lived in Sold Home × Equity Gain				-0.000653**	
				(-2.48)	
Above Median Years Lived in Sold Home (Ind.) × Equity Gain					-0.00815***
					(-3.12)
Zip Median HP of Sold Home at Purchase	X	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X	X
Sale Zip × Transaction Year FE	X	X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X	X
Instrumental Variables	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction
Sample	Full	Full	Full	Local Moves	Local Moves
Observations	3,074,934	2,692,174	2,800,884	1,444,434	1,444,434
Weak ID KP F Stat	116.5	101.2	106.3	67.52	68.15

Table A.4
Local Movers vs. Non-Local Movers - Robustness

Column (1) and (2) replicate column (1) of Table 6, but for the pre-COVID-19 period (1996-2019) and the COVID-19 period (2020-2021) respectively. Columns (3) and (4) are structured similarly, but replicating column (3) of Table 6. Our local vs. non-local results are not being driven by shift in household behavior during the COVID-19 pandemic. Reported *t*-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year level for columns one and three and at the purchase county level for columns two and four. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Price Premium, House Being Purchased			Average Sales Price, within 1/2-Mile and 0-180 Days of Home Purchase		
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Equity Gain, House Being Sold	0.104***	0.113***	0.0584**	0.0927***	0.0756***	0.224***
	(8.03)	(7.56)	(2.52)	(3.38)	(3.48)	(8.70)
Log Move Distance (Miles)	0.00716***			0.000874		
	(2.95)			(0.90)		
Log Move Distance (Miles) × Equity Gain	-0.00726***			-0.00181***		
	(-8.41)			(-3.64)		
Out-of-Zip, In-County, Move Indicator		0.0245***	-0.0290***		0.00212	0.0310**
		(3.82)	(-2.65)		(0.86)	(2.04)
Out-of-County, In-State, Move Indicator		0.0405***	-0.0241*		0.00227	0.0159
		(3.42)	(-1.87)		(0.60)	(1.16)
Out-of-State, Move Indicator		0.0540***	-0.0737***		0.00208	0.0305**
		(3.40)	(-5.76)		(0.35)	(2.28)
Out-of-Zip, In-County, Move Indicator × Equity Gain		-0.0200***	-0.00249		-0.00505**	-0.0240**
		(-5.58)	(-0.30)		(-2.48)	(-2.01)
Out-of-County, In-State, Move Indicator × Equity Gain		-0.0323***	-0.0164		-0.0105***	-0.0165
		(-6.71)	(-1.52)		(-3.09)	(-1.57)
Out-of-State, Move Indicator × Equity Gain		-0.0494***	-0.0224**		-0.0147***	-0.0214**
		(-8.12)	(-2.13)		(-3.34)	(-2.15)
Four Quarterly Historical Average Prices, At Matched Distance				X	X	X
Zip Median HP of Sold Home at Purchase	X	X	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X	X	X
Sale Zip × Transaction Year FE	X	X	X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X	X	X
Instrumental Variables	Zip ΔHPI over Sold Home Holding Period and Interactions	Zip ΔHPI over Sold Home Holding Period and Interactions	Zip ΔHPI over Sold Home Holding Period and Interactions	Zip ΔHPI over Sold Home Holding Period and Interactions	Zip ΔHPI over Sold Home Holding Period and Interactions	Zip ΔHPI over Sold Home Holding Period and Interactions
Sample	Full	Before 2020	2020-2021	Full	Before 2020	2020-2021
Observations	3,045,683	2,811,675	267,016	2,709,409	2,485,711	226,297
Weak ID KP *F* Stat	114.4	60.72	147.1	112.7	58.32	121.2

Table A.5
Equity Gains and Cash Discounts

This table reports results exploring the relationship between equity gain and cash purchase on price premia. In column (1) equity gains exhibit no causal relationship on whether or not a household purchases their next house all with cash. Consistent with the literature, we find a negative and significant relationship between a cash purchase and the purchase price of the home in column (4). Similarly, a cash purchase is associated with a decrease in price premia in columns (2) and (3). price premia for cash purchases is significantly less sensitive to equity gain in column (3), whilst total transaction price is more so in column (4). This is, perhaps, consistent with a violation of the exclusion restriction for the Reported t-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Cash Purchase, House Being Purchased	Price Premium, House Being Purchased	Price Premium, House Being Purchased	Log Transaction Price, House Being Purchased
	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)
Equity Gain, House Being Sold	-0.00325	0.0779***	0.0815***	0.186***
	(-0.62)	(6.23)	(6.33)	(15.40)
Cash Purchase Indicator, House Being Purchased		-0.176***	-0.162***	-0.289***
		(-20.91)	(-23.28)	(-24.66)
Cash Purchase Indicator × Equity Gain			-0.0132***	0.0158***
			(-3.35)	(6.29)
Zip Median HP of Sold Home at Purchase	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X
Sale Zip × Transaction Year FE	X	X	X	X
Purchase Zip × Transaction Year FE	X	X	X	X
Instrumental Variable(s)	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period and Interaction	Zip ΔHPI over Sold Home Holding Period and Interaction
Observations	3,103,108	3,103,108	3,103,108	3,103,108
Weak ID KP F Stat	241.4	242.9	122.2	122.2

Table A.6
Alternate Fixed Effects

This table reports results related to estimations of specification similar to that of Equation 11. Columns (1), (3), (5), and (7) are similar to column (2) of Table 2 and columns (2), (4), (6), and (8) are similar to column (1) of Table 4. Results relate to the full sample of matched household moves, with columns (1) and (2) restricted to non-local (out-of-county) moves, and columns (2), (4), (6), and (8) subject to there being at least one subsequent transaction in the time period and geography relevant to the respective column. All dollar amounts are in $100,000 units. Controls and fixed effects are included as indicated, with "at matched distance" referring to the geography relevant to the respective column and fixed effects varying as reported. The Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. Reported *t*-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Price Premium, House Being Purchased	Average Sales Price, within 1/2-Mile and 0-180 Days of Home Purchase	Price Premium, House Being Purchased	Average Sales Price, within 1/2-Mile and 0-180 Days of Home Purchase	Price Premium, House Being Purchased	Average Sales Price, within 1/2-Mile and 0-180 Days of Home Purchase	Price Premium, House Being Purchased	Average Sales Price, within 1/2-Mile and 0-180 Days of Home Purchase
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Equity Gain, House Being Sold	0.0408*** (13.04)	0.0188*** (8.25)	0.0513*** (14.06)	0.0515*** (13.39)	0.0794*** (5.88)	0.0956*** (4.17)	0.0977*** (4.34)	0.0540*** (3.20)
Four Quarterly Historical Average Prices, At Matched Distance		X		X		X		X
Zip Median HP of Sold Home at Purchase	X	X	X	X	X	X	X	X
Five Interacted Property Sold Characteristic FE	X	X	X	X	X	X	X	X
Years Lived in Sold Home FE	X	X	X	X	X	X	X	X
Sale County × Purchase County × Transaction Year FE	X	X	X	X				
Sale Zip × Transaction Year FE					X	X		
Purchase Zip × Transaction Year FE					X	X		
Transaction Calendar Month FE					X	X		
Sale Zip × Transaction Quarter FE							X	X
Purchase Zip × Transaction Quarter FE							X	X
Instrumental Variables	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period
Sample	Non-Local Moves	Non-Local Moves	Full	Full	Full	Full	Full	Full
Observations	1,006,330	887,630	2,536,678	2,259,887	3,103,108	2,735,407	2,753,376	2,413,075
Weak ID KP *F* Stat	3,507	3,037	2,545	2,175	461.4	417.4	389.9	349.5

Table A.7
No Small Moves

This table reports results related to estimations of specification similar to that of Equation 11. Column (1) is similar to column (2) of Table 2 and column (2) is similar to column (1) of Table 4. Results relate to the full sample of matched household moves, with the additional restriction of excluding moves of distance less than 3-miles, and column (2) subject to there being at least one subsequent transaction in the time period and geography relevant to the respective column. All dollar amounts are in $100,000 units. Controls and fixed effects are included as indicated, with "at matched distance" referring to the geography relevant to the respective column. The Kleibergen-Paap rk Wald F statistic is reported for all 2SLS specifications. Reported t-statistics in parentheses are heteroskedasticity-robust and clustered at both the purchase county and transaction year levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Price Premium, House Being Purchased	Average Sales Price, within 1/2-Mile and 0-180 Days of Home Purchase
	2SLS	2SLS
	(1)	(2)
Equity Gain,	0.0685***	0.0576**
House Being Sold	(4.65)	(2.18)
Four Quarterly Historical Average Prices, At Matched Distance		X
Zip Median HP of Sold Home at Purchase	X	X
Five Interacted Property Sold Characteristic FE	X	X
Years Lived in Sold Home FE	X	X
Sale Zip × Transaction Year FE	X	X
Purchase Zip × Transaction Year FE	X	X
Instrumental Variables	Zip ΔHPI over Sold Home Holding Period	Zip ΔHPI over Sold Home Holding Period
Sample	No Small Moves	No Small Moves
Observations	2,372,692	2,051,437
Weak ID KP F Stat	214.0	201.8

Table A.8

Does Overpayment Predict Overpayment?

Table Caption

	Future Price Premium, House Being Purchased	Future Price Premium, House Being Purchased
	OLS	OLS
	(1)	(2)
Price Premium, House Being Purchased	-0.295***	-0.282***
	(-17.08)	(-15.21)
Zip Median HP of Sold Home at Purchase	X	
Five Interacted Property Sold Characteristic FE	X	
Years Lived in Sold Home FE	X	
Years Lived in Purchased Home FE		X
Sale Zip × Transaction Year FE	X	
Purchase Zip × Transaction Year FE	X	X
Purchase Zip × Future Transaction Year FE		X
Observations	1,435,536	1,501,651
Adj. R^2	0.102	0.075