

2019

# The Effect of New Market-Rate Housing Construction on the Low-Income Housing Market

Evan Mast

*W.E. Upjohn Institute*, [mast@upjohn.org](mailto:mast@upjohn.org)

Upjohn Institute working paper ; 19-307

## **The Effect of New Market-Rate Housing Construction on the Low-Income Housing Market**

**Upjohn Institute Working Paper 19-307**

Evan Mast

*W.E. Upjohn Institute for Employment Research*

email: mast@upjohn.org

July 2019

### **ABSTRACT**

Increasing supply is frequently proposed as a solution to rising housing costs. However, there is little evidence on how new market-rate construction—which is typically expensive—affects the market for lower quality housing in the short run. I begin by using address history data to identify 52,000 residents of new multifamily buildings in large cities, their previous address, the current residents of those addresses, and so on. This sequence quickly adds lower-income neighborhoods, suggesting that strong migratory connections link the low-income market to new construction. Next, I combine the address histories with a simulation model to estimate that building 100 new market-rate units leads 45-70 and 17-39 people to move out of below-median and bottom-quintile income tracts, respectively, with almost all of the effect occurring within five years. This suggests that new construction reduces demand and loosens the housing market in low- and middle-income areas, even in the short run.

**JEL Classification Codes:** R31, R21, R23

**Key Words:** housing supply; housing affordability; filtering

### **Acknowledgments:**

Thanks to Nathan Sotherland and Steve Yesiltepe for excellent research assistance. Thanks to Brian Asquith, George Galster, Atul Gupta, Ray Kluender, Stephen Nei, and Davin Reed for helpful comments. An earlier version of this paper previously circulated under the title “The Effect of New Luxury Housing on Regional Housing Affordability.”

*Upjohn Institute working papers are meant to stimulate discussion and criticism among the policy research community. Content and opinions are the sole responsibility of the author.*

# 1 Introduction

Housing costs have risen rapidly relative to income over the past 60 years in the United States, particularly in large and economically successful cities (Albouy, Ehrlich, and Liu 2016). This trend has important economic implications—Hsieh and Moretti (2019) suggest that rising costs slow aggregate economic growth by limiting the number of workers in high-productivity cities, while Albouy, Ehrlich, and Liu (2016) and Ganong and Shoag (2017) find that the pattern increases real income inequality.

A heated debate on how to reduce housing costs has emerged, and one frequently proposed solution is relaxing land-use regulation and increasing housing supply.<sup>1</sup> While the effect of such policies is obvious in a simple model of homogenous housing units, housing is highly differentiated—new construction is predominately expensive and quite different from units that are affordable to lower-income households. If the housing market is highly segmented, with few households searching or moving across dissimilar housing types, an increase in the supply of expensive new units could have little effect on the market for lower-income housing. The strength of this relationship is crucial to policymakers considering reforms that increase market-rate construction, who must weigh benefits against costs such as objections from neighbors, concerns of gentrification, and reduced political capital for subsidized units or housing vouchers (Been, Ellen, and O’Regan 2019). However, there is little related empirical evidence, especially in the short- or medium-run most relevant to the current debate.<sup>2</sup>

In this paper, I use a large sample of address-level individual migration histories to provide evidence that new market-rate construction substantially loosens the market for middle- and low-income housing by inducing a series of moves that reduces demand for these areas. The effect occurs within a few years of the new units’ completion. I begin my analysis with a simple model of new housing construction in a market with three quality tiers or submarkets, in which new high-quality construction lowers prices for lower quality units through a “migration chain” mechanism.<sup>3</sup> Some households who would have otherwise

---

<sup>1</sup>See, for example, Minneapolis and Oregon’s recent prohibitions of single-family zoning or the California State Senate’s rejected proposals (including State Bills 827 and 50) to loosen local zoning restrictions.

<sup>2</sup>Rosenthal (2014) shows that new units “filter” to become more affordable as they depreciate over the course of decades but does not study their effect on existing units or the broader housing market.

<sup>3</sup>Grigsby (1963) develops an initial theory of housing submarkets, and Rothenberg et al. (1991) further develop the idea into a model of a system of interconnected submarkets.

occupied cheaper units move into new units, reducing demand and lowering prices for the units they leave vacant. The process iterates when a second round of households moves into the units the first round left vacant and so on, eventually reducing prices in low-income areas.<sup>4</sup> However, whether this chain actually reaches such areas in the real world depends on two key factors. A chain has some chance of ending in each round, whether it is due to household formation, a unit being used as a second home, out-of-metro migration, or landlords not reducing rents enough to fully fill vacancies. The longer a chain lasts, the more likely it eventually draws households out of lower quality housing. Second, the stronger the migratory connections between lower quality housing and new housing, the more likely that a chain reaches a lower quality unit in any given round.

Next, I use individual address history data from Infutor Data Solutions to conduct three related empirical exercises. I first broadly consider migratory connections between neighborhoods in 12 major metropolitan areas (CBSAs) and find strong connectivity between census tracts with slightly different characteristics. Individuals originating in, say, the fifth income decile frequently move to the fourth or seventh decile, but rarely the tenth or first. This pattern implies that distinct submarkets exist, but that even quite different areas are connected by a short series of common moves.

In my second exercise, I sharpen focus to the connectivity between new construction and low-income areas and exploit the data's granularity to track moves at the building level. I identify 686 large new market-rate multifamily buildings in central cities and track 52,000 of their current residents to their previous building of residence. I then find the tenants currently living in those buildings and track them to their previous residence, iterating for six rounds and, in order to focus on local connectivity, keeping only within-CBSA moves in each round. About 20 percent of new building residents moved in from tracts with below CBSA-median income, and that proportion rises steadily to 40 percent in round six. Similar patterns emerge for other characteristics, suggesting strong connections between submarkets that are inconsistent with a highly segmented market in which new construction does not affect low-income areas. The results also highlight the geographically diffuse nature

---

<sup>4</sup>Kristoff (1965) was perhaps the first to formulate this mechanism. Similar results occur in richer models, such as Sweeney (1974), Braid (1981), and Nathanson (2019).

of migration chains—only 30 percent of round six originates within the CBSA central city.

The first two exercises show connectivity that strongly suggests that new construction will affect middle- and low-income areas, but they do not provide a quantitative estimate of this effect. To fill this gap, I run a more detailed simulation that allows migration chains to end in each round and households to move whether or not a new unit is built. This simulation allows me to estimate an intuitive metric of a new unit’s effect on other submarkets (defined according to tract characteristics). I define the number of “equivalent units” a new unit produces in a submarket—say, below-median income tracts—as the probability that its migration chain reaches such an area before ending. The intuition behind this metric is simple: inducing a household to leave a submarket is similar to building a new (depreciated) unit in that submarket. The chain reduces demand by one, while building a unit increases supply by one. I focus on this quantity-based outcome because the diffuse nature of migration chains makes estimating a price effect difficult, and it also fits naturally in the policy debate, where “inclusionary zoning” ordinances require developers to build some income-restricted units for each market-rate unit.<sup>5</sup>

While I cannot directly observe either when chains end or where a household would have lived if a new building was never constructed, I use data on vacancy rates, household formation, and within- and across-metro migration to construct a range of reasonable assumptions. In my baseline specification, 100 new market-rate units create 70 equivalent units in below-median income tracts and 39 in bottom-quintile income areas. In my most conservative specification, in which chains end with a much higher probability, I find 45 and 17 equivalent units in below-median and bottom-quintile income areas, respectively. These figures compare favorably to the 5 to 20 income-restricted units that would be required by typical inclusionary zoning ordinances—the connectivity implied by the migration data is strong enough that migration chains frequently reach low-income areas even if they end at a relatively high rate. The effect also appears to cross racial lines. Even for tracts that are in the bottom quintile of percent white and below median income, estimates range from 23 to 49, though these areas are a small percent of the typical CBSA.<sup>6</sup> Effects should be fully

---

<sup>5</sup>Schuetz, Meltzer, and Been (2009) and Thadden and Wang (2017) provide summaries of inclusionary zoning policies in the United States.

<sup>6</sup>Note that these equivalent unit numbers in different housing types should be considered separately

felt within two to five years.

The results from each of these exercises suggest that new market-rate housing construction loosens the market for middle- and low-income housing, even in the short run. This points to an important role for policies that increase construction, as well as less formal interventions such as policymakers pushing development proposals through the often onerous approval process. However, a caveat is that I do not estimate price effects. Because the private market will not provide housing at below marginal cost, market-based strategies may not lower prices in neighborhoods with already very low prices. Alternative policies that either lower the cost of provision or subsidize incomes are likely necessary to improve affordability in such areas. Another limitation is that I study regional effects, and new buildings could have different effects on their neighborhood, where they may change amenities or demographic composition. Empirically, I only observe a housing unit’s location, not characteristics or price, and my simulation requires assumptions on chain decay rate and where individuals would move in the absence of construction. I present some evidence that selection within tracts is minor and explore a variety of alternative assumptions.

This paper contributes to the literature on housing construction and housing prices, sometimes called the filtering literature,<sup>7</sup> and recent empirical work is particularly relevant. Rosenthal (2014) and Weicher, Eggers, and Moumen (2016) find that new units slowly become more affordable over time, particularly after entering the rental stock. Anenberg and Kung (2018) use a neighborhood choice model to estimate extremely small price effects of new housing. Their result may be driven by the assumption that each new unit induces a new migrant to a city, which Nathanson (2019) relaxes in a calibrated spatial equilibrium model, finding a much larger effect. More broadly, Piazzesi, Schneider, and Stroebl (forthcoming) use a model of a segmented housing market to show that a localized shock’s broader effect depends heavily on connections between the shocked area and the rest of the market.

I build on this literature by using novel methodology and granular data to show that new housing has large short-run effects on middle- and low-income submarkets. A small literature

---

rather than summed together, because a new unit in one type starts a migration chain that may nest an equivalent unit in another.

<sup>7</sup>The term *filtering* has been used to refer to the distinct concepts of housing units becoming more affordable over time and households moving through different housing units, as well as other mechanisms.

has also studied migration chains, sometimes called “vacancy chains.” Kristoff (1965) and Lansing, Clifton, and Morgan (1969) construct chains by interviewing households and find substantial decreases in income with each round. More recently, Turner (2008) and Turner and Wessel (2019) use administrative data on Stockholm and Oslo, respectively, to show that the series of moves from new construction is concentrated in high-income areas. I build on these papers by providing a metric to quantify a chain’s effect on lower-income submarkets, by considering that households may have moved in the absence of new construction, and by using a new U.S. data source to study a number of large cities.

Finally, a large literature, reviewed by Gyourko and Molloy (2015), focuses on the adjacent question of how regulation affects housing construction and prices. Studies generally find that regulation increases prices and reduces construction,<sup>8</sup> but typically must contend with small samples where large geographic areas are the unit of observation, as well as endogenous and heterogeneous policies. I contribute to this literature by studying the effect of construction directly, making results relevant to a large set of policies. Additionally, I estimate effects on granular submarkets and use new methodology that is driven by detailed migratory patterns rather than price variation across large geographies.

## 2 Conceptual Framework

### 2.1 Baseline Model and Migration Chains

In this section, I first present a highly stylized model of new housing construction that is quite similar to the graphical examples in Rothenberg et al. (1991). I then use the model to illustrate both the migration chain concept and the equivalent unit metric. Lastly, I discuss how adding realistic complications to the model—household formation, vacation homes, out-of-metro migration, landlord market power—would cause some migration chains to end in each round, potentially before reaching lower income submarkets, highlighting that the magnitude of new housing’s effect on such areas is an empirical question.

Consider a self-contained housing market with a unit mass of households, each of which

---

<sup>8</sup>See, for example, Ihlanfeldt (2007), Glaeser and Gyourko (2003), Glaeser and Ward (2009), and Quigley and Raphael (2005). Saiz (2010) finds similar results for geographic constraints.

has income  $\alpha \sim f$  and lexicographic preferences over housing types and other consumption. Households have a strict preference for housing type  $H$  over  $M$  over  $L$ . As a baseline scenario, suppose that there is a mass of housing units  $\vec{S} = \{S_M, S_L\} = \{.5, \infty\}$  that are rented out by perfectly competitive absentee landlords with no costs who cannot observe  $\alpha$ .

Equilibrium is a vector of prices  $\vec{P} = \{P_M, P_L\}$  and household locations  $\vec{Q} = \{Q_M, Q_L\}$  such that households prefer their housing type to any other housing unit within their budget, landlords have no incentive to change their rent, and  $Q_M = S_M, Q_L \leq S_L$ . Equilibrium prices are then given by  $\vec{P} = \{F^{-1}(0.5), 0\}$ , and the higher tier of housing units is occupied by households with  $\alpha \geq F^{-1}(0.5)$ , while the lower tier is occupied by those with  $\alpha < F^{-1}(0.5)$ . For example, if  $\alpha \sim U[0, 1]$ ,  $\vec{P} = \{0.5, 0\}$ .

Next, consider the equilibrium that would occur in the alternative scenario where there are  $q$  units of top-quality housing type  $H$  in the market, yielding  $\vec{S} = \{S_H, S_M, S_L\} = \{q, .5, \infty\}$ . In order to fully lease the top tier units, landlords will set  $P_H = F^{-1}(1 - q)$ . The households in the top-tier units then have  $\alpha \in [F^{-1}(1 - q), 1]$  and lived in the middle tier in the baseline scenario. Because this group is not occupying type  $M$  units in this scenario, landlords must lower  $P_M$  to  $F^{-1}(0.5 - q)$  in order to attract lower budget households to these units. This induces households with  $\alpha \in [F^{-1}(0.5 - q), F^{-1}(0.5))$  to live in the middle tier instead of the bottom tier. Finally, because supply in the bottom tier is infinite and the price is zero,  $P_L$  does not change.<sup>9</sup> Adding new units thus affects prices in all but the bottom tier and improves housing quality for some households originating in every tier of the market. Braid (1981) and Nathanson (2019) show a similar result in more complex models.<sup>10</sup>

In the world of the model, one can evaluate the effect of high-end housing construction by simply comparing prices and household allocations between the two scenarios. In practice, this is impossible because only one of the two scenarios is observed. Empirical economists typically attempt to approximate the comparison by constructing a natural experiment or parameterizing and estimating a model. The migration chain approach instead follows a

---

<sup>9</sup>The null effect of new construction on bottom tier prices is a consistent theme in the filtering literature and is often cited as a limitation of market-based strategies (e.g., Rothenberg et al. 1991). The intuition is either a minimum cost of housing or defining the bottom tier as homelessness.

<sup>10</sup>These papers further show that new housing of any type will lower prices for all types. This occurs because, for example, new middle-tier construction induces some households to move down from the top tier to take advantage of lower prices, and is sometimes termed *filtering up*. In my model, new housing only lowers prices in qualities below the new construction because of the assumption of lexicographic preferences.

sequence of household moves to identify a set of people who occupy a different housing unit than they would have in the absence of construction. The migration chain from a particular new unit can be formalized as a vector  $C$  of housing types. The first element is the housing type of the new unit, and the second is the type of unit that the household in the new unit would have occupied in the baseline scenario, which is  $M$  in this example. The third element is the baseline-scenario unit of the household occupying the unit vacated by the second round (type  $L$  in the example).

The migration chain contains important information about the effect of new construction on other housing types. For any particular type, the decrease in prices caused by new construction depends directly on how many households are induced to move out of that type. Moreover, in this simple model, inducing  $m$  households to move out of the middle quality tier and building  $m$  new units directly in that tier have exactly the price effect—both lower  $P_M$  to  $F^{-1}(0.5 - m)$ . This provides the intuition for my equivalent unit metric, where I say that a migration chain creates an equivalent unit in a given submarket if it reaches that submarket before ending. However, I note an important limitation of the migration chain approach—it does not account for complications such as amenity or neighborhood composition changes caused by the new building, which could affect prices in a more complicated model.

## 2.2 Why Migration Chains End

There are a number of real-world frictions outside of the model that end chains in each round, potentially before they reach lower-income areas. Starting with the supply side, some chains may end because landlords with market power do not lower prices by enough to completely fill the vacancies created by the chain. This would appear in the model as, for example, middle-tier landlords only lowering prices to  $P_M = F^{-1}(0.5 - q/2)$ . Heterogeneous costs incurred from renters with different housing budgets or discriminatory motives could also lead landlords to not fill all vacancies.

Frictions on the household side of the model can also end chains. First, a housing unit could be a second home or investment property, in which case the owner does not vacate their other unit. In addition, chains could end because a new household forms to fill the new unit. Lastly, if new units induce households to move in from outside the metropolitan

area, the subsequent benefit of the chain will not accrue to the area. While sources of chain decay frequently appear in the policy debate, they have only been accounted for to a limited extent in the theoretical literature.<sup>11</sup> One simple approach is to assume that some percent  $d$  of chains end in each round. In the example where  $q$  new top tier units are built, this would imply that only  $dq$  of the new units are filled, and only  $d^2q$  of the vacancies subsequently created in the middle tier are filled. In this case,  $P_M$  only falls to  $F^{-1}(0.5 - d^2q)$  instead of  $F^{-1}(0.5 - q)$ .

An important complication is that these events only end a chain if a household takes an action that they would not have in the counterfactual world with no construction. For example, the chain only ends with a second home if the owner would not have bought a second home in the absence of new construction. This makes decay harder to assess empirically.

## 3 Data

### 3.1 Infutor Data

My primary data source is individual address histories from Infutor Data Solutions, which was recently introduced to the academic literature by Diamond, McQuade, and Qian (forthcoming). Infutor constructs this information from numerous private and public record sources—such as U.S. Postal Service change of address forms, county assessor records, magazine subscriptions, and phonebooks—and largely sells the data for use in targeted advertisements. Addresses are reported at the unit level and, since they are intended for use in direct mailing, are quite high quality. Each address is accompanied by an estimated date of arrival, and the data contain some limited demographics (age, gender) on each individual.

Because the data contain limited information on housing unit characteristics, I classify units based on their census tract and characteristics from the 2013–2017 American Community Survey (ACS). In addition, because the data track individuals rather than households, I assume that each person occupies a distinct unit. Since children and very young adults are essentially not included in the Infutor data, this assumption largely leads to weighting

---

<sup>11</sup>Nathanson (2019) and Braid (1991) discuss the importance of indivisible housing units for their results, a concept that can be mapped into second homes, and also consider across-MSA migration.

couples more heavily than singles, which may be appropriate given that the former typically occupy larger units.

I examine selection with a number of exercises comparing the Infutor data to established data sources.<sup>12</sup> The Infutor data closely track the census over-25 population at the tract level, with a median of 0.88 observations per census individual. The coverage rate is quite similar across demographic groups, as shown in Figure 1, which plots the ratio of the Infutor and census populations against tract characteristics. The largest differences appear between tracts with different racial composition, with a coverage rate of about 80 percent in the least white tracts versus 95 percent in the whitest tracts.

Because my study primarily uses this data to track household migration, Infutor’s coverage of moves is also important. The data miss a substantial number of moves—the annual individual migration rate in the Infutor data is 5.4 percent, compared to the 9.8 percent reported in the Census Bureau’s 2018 Current Population Survey. This could occur both because of difficulty linking individuals across moves and because the Infutor data has poor coverage of highly mobile young adults. However, because my study uses each move separately, it only requires that the moves that do appear in the Infutor data are randomly selected. To examine this, I next compute the average annual migration rate at the county level in the Infutor data and compare it to census estimates (which are not available at the tract level). Appendix Figure A.1 plots the ratio of the two estimates against county characteristics. Coverage appears to be relatively uncorrelated with county characteristics. There is a slight correlation with county income, where the ratio is about 0.38 in low-income counties and 0.45 in high-income counties. This correlation could lead to an underestimate of the number of individuals who move into a given neighborhood from a lower-income neighborhood.

Another potential concern is that Infutor may not have the correct endpoints for moves. For example, if a household moves from A to B to C, the data could record the move as from A to C, overstating the connectivity between A and C. To examine this, I compute the difference in destination and origin county median household income for moves reported in

---

<sup>12</sup>Because my exercise concerns large metro areas, I restrict the sample in these exercises to counties that are part of a core-based statistical area.

Infutor and in the IRS Statistics of Income data. The distributions are extremely similar, as shown in Appendix Figure A.2. In my simulation exercise, I also run robustness checks in which I remove moves across very different neighborhoods.

## 3.2 New Market-Rate Buildings

I identify new market-rate buildings using the Infutor data. I first collapse the individual-level data by street address and keep buildings with over 16 individuals.<sup>13</sup> I then identify new buildings as those where over 90% of current residents moved in since 2009 and keep only those that were completed prior to 2017 and are within five miles of the central business district.<sup>14</sup> Because the policy debate concentrates on relatively expensive new buildings, I keep only those that are in census tracts that are above the core-based statistical area (CBSA) median in either median household income or income per capita. Finally, I drop buildings—such as student housing, post offices, affordable or subsidized housing, and homeless shelters—that meet the previous criteria but are not market-rate apartments. The algorithm identifies both rental and owner-occupied buildings, as well as some renovated buildings whose previous use was not residential.

In my primary specification, I include 12 of the largest metropolitan areas in the United States: New York City, Chicago, Dallas, Houston, Washington, Philadelphia, Atlanta, Boston, San Francisco/Oakland, Denver, Seattle, and Minneapolis.<sup>15</sup> I define each metropolitan area according to the most recent definition of CBSAs. As shown in Table 1, I identify 52,432 individuals in 686 market-rate multifamily buildings constructed since 2009. The buildings are relatively evenly distributed across cities, with Seattle, New York City, and Chicago having the most (all over 80) and Philadelphia and Boston the least (under 20), and individuals are distributed similarly. About 67 percent of new building residents originate from an address within the same CBSA, and half originate from within the same city.

Figure 2 plots the locations of new market-rate buildings in Chicago, as well as the origin

---

<sup>13</sup>Given the data’s coverage rate of approximately 90 percent, this implies over 20 individuals live in the building.

<sup>14</sup>I hand-checked a number of recently completed buildings to determine that the 10th percentile of arrival dates in a building provides a very good estimate of when a building was completed.

<sup>15</sup>I omit some large metropolitan areas, such as Los Angeles, that do not have a well-defined city center.

addresses of their current tenants.<sup>16</sup> Unsurprisingly, new buildings appear to be concentrated near the central business district, with some slightly farther in wealthy north and northwest neighborhoods. Building residents appear to come from largely wealthy areas, and virtually none come from zip codes with median household income below \$30,000.<sup>17</sup>

Table 2 provides details on the location and size of the sample of new buildings. The median number of Infutor individuals living in a building is 60, and the mean distance to the central business district is 1.95 miles. Unsurprisingly, the buildings are located in very high-income and high-rent tracts—the sample restriction that buildings must be in above-median income tracts is generally not binding. Lastly, the buildings’ tracts have relatively high vacancy rates: 10 percent on average and over 15 percent for many. This may reflect either a high rate of second homes—ACS vacancy rates include only primary residences as occupied—or strong market power for landlords of new buildings.

## 4 Descriptive Exercises

### 4.1 Migratory Connections between Neighborhoods

Before examining new housing directly, I study general migratory connections between different types of neighborhoods. Because the distribution of income and race varies greatly across cities in the sample, as shown in Appendix Table A.1, I focus on migration across tracts within the Chicago CBSA (dropping tracts that are over 20 percent college students) and use moves between 2010 and 2017 to construct graphs similar to transition matrices.

Figure 3 plots distributions of destination tract characteristics conditional on origin tract characteristics. In Panel A, each box shows the median and interquartile range of destination median household income for migrants that originated in a given income decile. Whiskers show the 10th and 90th percentile. Individuals originating in top decile income tracts very rarely move to a below-median income neighborhood, and very few people from lower deciles migrate above the median. While this suggests that submarkets exist, they also appear to

---

<sup>16</sup>I add small amounts of noise to each marker to avoid precisely identifying addresses.

<sup>17</sup>Appendix Figure A.3 includes new building residents who originated anywhere in the Chicago CBSA and tells a largely similar story. Appendix Figure A.4 repeats the exercise in San Francisco, which shows a wider dispersion of both buildings and resident origins, consistent with its generally high incomes.

be permeable—individuals frequently move from the seventh decile to the ninth, the sixth to the fourth, et cetera. The top decile and lower deciles are connected through a series of moves, which is precisely what the migration chain mechanism requires. Panel B shows a similar pattern for median two-bedroom rent.

Panel C plots the same graph for the percent of households in a tract that are white. The least white tracts appear to be more separated from the remainder of the market than were either top or bottom income tracts. People originating outside of these tracts are extremely unlikely to migrate in. However, because individuals do migrate out of these tracts, they are still connected to the broader housing market, and this “outward” connectivity is actually what is required for migration chains to reach these heavily nonwhite neighborhoods. Finally, Panel D depicts median rent burden, which shows much more connectivity across deciles. Appendix Figure [A.5](#) contains the same graphs for the San Francisco area, which are similar.

On the whole, these graphs show that even a housing market that is highly segregated on most measures exhibits substantial migratory connections between tracts with very different characteristics. This implies that there are meaningful connections between housing submarkets that may lead housing construction or other shocks in one submarket to affect others. However, the exercise classifies migrants according to average tract characteristics, which may disguise significant heterogeneity in households and housing units within a tract. It may be that transitions from, say, the sixth and ninth decile income tracts are actually driven by individuals moving from the most expensive unit in the sixth decile to an average unit in the ninth decile, leading the tract-level definition to overstate true connectivity. In my next exercise, I track moves at the building level to mitigate this problem.

## 4.2 Constructing Sequences of Origin Units

I now directly study the connectivity between new construction and low-income areas by constructing sequences of origin units from new buildings. The algorithm to construct this sequence of origin units is simple. I start with the 52,000 individuals currently living in the 686 new market-rate buildings described in Table [2](#). I then use the Infutor data to identify their origin housing units. I then identify the people currently living in the first round’s origin buildings and iterate for six rounds. This exercise is distinct from the migration chain

simulation in the next section because it does not allow chains to end or individuals to move in the absence of construction.

While this exercise is intuitive, there are a few specification choices to note. First, in order to focus on connectivity within metro areas, I only include individuals who moved from within the new building’s CBSA in each round of the chain.<sup>18</sup> This includes about 70 percent of tracked individuals, depending on the round. Second, note that I construct the next round of the chain based on who is currently living in the previous round’s origin units, rather than who was first to occupy that unit after the previous round moved out. This provides a snapshot version of the sequence that may change depending on when measured. However, given that the sample spans only 2009 to 2017, this likely yields similar results to following the first person to move into a given unit after it is vacated.

Finally, I construct each round by taking all people in the previous round’s origin building, not their specific origin unit.<sup>19</sup> When constructing the next round, I then weight the residents of that building so that they sum to one individual. Matching at the building level increases the probability that at least one person originated within the CBSA, allowing me to construct another round. In the event that no one in a building is tracked within the same metro, I proportionally distribute the weight from that building to other similar buildings that are tracked.<sup>20</sup> Sixty-seven percent of new building residents are tracked to a previous within-metro address, and 74 percent of buildings in the next round have at least one person tracked within-metro. This percent gradually falls to 52 percent by the final round as single-family residences and untracked individuals become more common.

### 4.3 Results on Sequences of Origin Units

Figure 4 shows the percent of individuals in each round that originated within the principal city. (Recall that only those who originated within the CBSA are included.) Seventy

---

<sup>18</sup>I also drop individuals who moved from tracts that are over 20 percent undergraduate students.

<sup>19</sup>To avoid major changes over time in building or neighborhood attractiveness, I restrict to people that moved into the building since 2009.

<sup>20</sup>Similar buildings must have the same principal city/suburban status as the untracked building, which helps account for selection on tracking rates between multifamily and single-family buildings. Because it is computationally easier to calculate the chain separately for each category in Figure 5, the characteristics that tracked and untracked buildings are matched on change with each computation. In each case, I require similar buildings to have the same in/out of category status as the untracked building.

percent of round one, the tenants of new buildings, moved from within the principal city. This is intuitive given the new buildings' central city locations. However, this percentage steadily declines to 30 percent in round six, close to the population average. This pattern highlights that the effects of the migration chain mechanism are geographically diffuse. While this implies that diverse regions will benefit, it also makes it less likely that any particular neighborhood will be strongly affected.

Figure 5 shows the percent of each round in five broad overlapping tract categories, defined according to the within-CBSA characteristic deciles shown in Appendix Table A.1. About 20 percent of new building residents originate in tracts with below-median household income, rising to 40 percent by round six. The percent of individuals originating in a bottom-quintile income tract increases from 7 percent to 15 percent from the first to sixth round, and the percent from tracts that are below-median income and in the bottom quintile of percent white starts at six and eventually rises to 14. Finally, almost none of the first round originates in below-median income and rent-burdened tracts, compared to 10 percent of the final round.<sup>21</sup> These migration patterns are not consistent with strong segmentation—even when tracing moves at the building level, it appears that a short series of moves connects new construction and low-income areas.

This approach mitigates one issue with selection within tracts, since each round of the chain is constructed using a building-to-building move, rather than average migration into a type of tract. However, the units in the sequence are still classified according to tract characteristics, which may not match actual unit quality. To diagnose the extent of this problem, I compare the likelihood that the average unit in, say, the fifth income decile is filled by a person moving from a lower-income tract to the same probability for a fifth-decile unit that is included in the sequence. Panel A of Appendix Figure A.7 shows that the units in the first round of the sequence—the origin units of the new building's residents—are somewhat less likely to be filled by a person from a lower income tract than the average unit. In the fifth income decile, the figure is about 30 percent, versus 36 percent in the full

---

<sup>21</sup>While the numbers are substantially larger for below-median income tracts, note that these tracts by definition make up a much larger percent of a CBSA. Appendix Figure A.6 normalizes each line by the percent of the CBSA's population that lives in each group of tracts. Below-median income tracts are still proportionally more represented in the first round, but the gap closes substantially by round six, when all categories fall between 0.6 and 0.8.

sample. However, as shown in Panel B, this gap falls to less than 1.5 percentage points by the third round. This suggests that units in the sequence are somewhat positively selected, particularly in the first round, but are relatively representative of the average unit in their tract.

## 5 Migration Chain Simulation Methodology

### 5.1 Framework

In this section, I simulate migration chains that end with some probability in each round and account for individuals who would have moved even in the absence of new construction. This is a more complicated exercise than the sequence of origin units just described, but the added structure allows me to quantify the effect of new housing on lower-income submarkets.

A migration chain  $C$  is a sequence of housing units.  $C_1$  is a new unit, and  $C_2$  is the unit that the person living in the new unit would have occupied had the new unit never been constructed. I call this the person’s counterfactual location. For example, if the individual living in the new unit left a house on Willow Drive to move to the new unit, but would have moved from Willow Drive to Oak Lane even if the new unit had not been constructed, Oak Lane is the second element in  $C$ . The subsequent elements in  $C$  are then defined recursively from  $C_2$ .

This recursive structure is empirically convenient. To construct  $C_{i+1}$  given  $C_i$ , I first define  $O_i$  as the origin unit of the individual currently living in  $C_i$ . I then introduce the operator  $T$ , which maps an individual’s origin unit to his counterfactual unit:

$$C_{i+1} = T(O_i).$$

In addition, I allow chains to end in each round with some probability  $d(C_i)$  that depends on the current location of the chain.  $O_i$  is directly observable in the data, and I discuss my empirical implementation of  $T$  and  $d$  in the following subsections.

In order to quantify the effect of the chains, I say that a chain creates an “equivalent unit” in submarket  $k$  if it reaches that area before ending. The expected number of equivalent

units in  $k$  created by a new unit in  $h$  is then just the probability that the chain reaches  $k$ . Formally,

$$EU_h(k) = P(C \cap k \neq \emptyset | C_0 \in h).$$

I simulate a migration chain from each new unit and compute this probability empirically for a variety of submarkets. Again, the intuition is that inducing a household to move out of a submarket reduces demand for that submarket by one, which should have a similar effect on prices as building an additional (depreciated) unit in  $k$ . Note that if a chain from a given new unit reaches  $k$  twice, it is still only counted as one equivalent unit. This is because the second occurrence is included in the migration chain that an additional unit in  $k$  would produce.

The following two subsections describe my assumptions on counterfactual locations and the rate of chain decay. Because neither is directly observable in the data and prior literature provides little guidance, I use data on migration, household formation, and vacancy rates to construct a range of reasonable assumptions and repeat the simulation at several points within that range.

## 5.2 Counterfactual Locations and Submarket Definition

For this simulation, I classify each census tract in the sample into a submarket according to its within-CBSA deciles of median household income and percent white households, whether it is in the principal city of its CBSA, and whether it is in the top quintile of median rent burden.<sup>22</sup> This gives a maximum of 400 possible submarkets per metro area, but, on average, only 300 of those contain at least one tract. While I use this granular definition to run the simulation, I use larger groups when presenting results—below-median income or bottom-quintile income tracts, for example.

I first define counterfactual locations for individuals living in the new buildings. The simplest assumption would be that no one would have moved in the absence of construction, yielding  $C_2 = T(O_1) = O_1$ . However, this assumption would likely underestimate the quality

---

<sup>22</sup>The Census Bureau defines rent burden as the ratio of gross rent to household income in renter-occupied housing units.

of  $C_2$ , as some people may have moved from their origin into a nicer unit anyway. To account for this, I instead assume that individuals would have moved up slightly in the absence of construction. Rather than choosing a specific unit, I define  $C_2$  as a submarket.

For within-metro migrants, my baseline assumption is that the counterfactual submarket is one decile higher in the income distribution than the origin submarket and the same on other characteristics. This is approximately the median step up in income in Figure 3, though it overstates upwards mobility for people originating in high-income areas, making the assumption somewhat conservative. I also run the simulation under alternative assumptions that increase the income step size and vary the changes in other characteristics.<sup>23</sup>

Out-of-metro migrants are slightly more complicated. Depending on the specification, I assume that a certain percentage would have moved to the CBSA even in the absence of construction (as discussed in detail in the next section). For those who would have moved to the area anyway, I draw the counterfactual location from the distribution of counterfactual locations of within-metro migrants to the same new building.<sup>24</sup> In contrast, the migrants who were induced to move to the area represent the end of the chain.

After the first round, I modify the algorithm slightly. This is necessary because the  $T$  operator defines  $C_2$  as a submarket, rather than a specific building. To construct the next round of the chain, I replicate the process and assumptions used in the first round of the chain, but use the distribution of origin units of recent (2009–2017) arrivals to that submarket instead of a specific origin building.

### 5.3 Chain Decay Rate

There are several reasons that a chain could end—second homes, household formation, migration from outside the CBSA, or landlord market power. While the prevalence of each force is observable, the exercise actually requires measurement of the marginal increase induced by new construction, which is not observable. I consider two distinct assumptions.

---

<sup>23</sup>Because the submarket definition is granular, there are some categories that no observed tract belong to. If the counterfactual location assumption leads to one of these empty submarkets, I assume that the counterfactual location is equal to the origin.

<sup>24</sup>I make the same assumption for individuals that are not tracked to their previous address in the Infutor data and individuals that originate in heavily college student (>20 percent) tracts.

First, as my baseline estimate, I consider a marginal increase in housing supply. I assume that a small change to the housing market does not affect major decisions like household formation and across-CBSA migration. On the other hand, landlord market power should moderate the effect of even small changes in housing supply. To be conservative, I also assume that second home purchases are always marginal to new construction. To capture these two forces empirically, I set the probability  $d(C_1)$  of a chain ending in the first round equal to the vacancy rate in the block group containing  $C_1$ . In subsequent rounds, I set it equal to the average vacancy rate in  $C_i$ 's submarket.

The intuition for this assumption is best illustrated by considering the new housing units, where this rate captures the units that are unfilled because landlords do not price them low enough or because they are not used as a primary residence (the ACS vacancy rate counts these as vacant), as well as matching frictions causing some units to be vacant at a given time. The intuition is similar in later rounds. When a chain induces households to move out of a submarket, landlords react by lowering prices, but not by enough to reduce the vacancy rate to its preshock level.

Second, I allow new construction to increase household formation and migration to the CBSA. This may provide a better approximation of large changes to housing supply. The Current Population Survey provides an estimate of the percentage of moves that are caused by new household formation, and I use Infutor to estimate the frequency of out-of-CBSA migration. I assume that new construction increases both of these forces by 25 percent, a very large effect that roughly doubles the decay rate.

Using the block group vacancy rate instead of the actual vacancy rate in a new building will be inaccurate if new buildings are quite different from their block group. This is a particular concern given claims of extremely high rates of second homes and investment properties in new luxury buildings.<sup>25</sup> To investigate this, I compare the vacancy rates in the

---

<sup>25</sup>This topic is hotly debated. A number of investigations have found that high numbers of new condos are either owned by shell corporations or do not claim tax exemptions as a primary residence. Logan (2018) finds that only 36 percent of units in 12 new Boston buildings claimed the city's property tax exemption for a primary residence. Solomont and Sun (2019) find that 16 percent of units in Manhattan condo buildings with over 30 units are owned by shell corporations, while over 30 percent did not claim a tax exemption for primary residences. However, these reports are unable to determine if such units actually sit vacant or if they are rented out. Scanlon et al. (2017) estimates that 70 percent of foreign-owned apartments in London were rented to locals. The city of Vancouver reported that 2,538 units in the city were subject to its "empty

tract and block groups containing the new buildings.<sup>26</sup> If new buildings have substantially higher vacancy rates than their surrounding area, this should lead to higher vacancy rates in their containing block groups than tracts. This is not true, as Table 2 shows that the distributions of block group and tract vacancy rates are extremely similar.

## 6 Migration Chain Simulation Results

### 6.1 Baseline Equivalent Unit Estimates

Figure 6 shows baseline estimates of the number of equivalent units in each migration round for five categories of tracts, aggregated from the smaller submarkets to ease presentation. One hundred new market-rate units create 70.2 equivalent units in below-median income tracts. The estimates are also large for areas that are even less similar to high-income areas, with 39.6 created in bottom-quintile income areas and 45.3 in areas that are below-median income and in the top quintile of rent burden. Even for tracts that are below median income and in the bottom quintile of percent white, the figure is 48.8, though these areas are a relatively small percent of the typical CBSA.<sup>27</sup> Most equivalent units are created in early rounds, especially for below-median income areas, and production subsequently slows smoothly. Appendix Figure A.8 repeats the plot but includes only equivalent units within the principal city of the CBSA. The number of equivalent units in each category drops by 20-30 percent, again highlighting that the benefits from new market-rate housing are diffuse throughout a metropolitan area.

Given average rental vacancies and time-on-the-market for home sales, these effects should be felt relatively quickly. Zillow reports that the average house was on the market for one

---

homes tax” in its first year, which is only about 3 percent of the 68,000 units constructed in Vancouver between 2000–2016.

<sup>26</sup>Because Infuture has imperfect coverage, I cannot use it to calculate the vacancy rate in a building.

<sup>27</sup>While there is, to my knowledge, no direct comparison for these results in prior literature, Lansing, Clifton, and Morgan (1969) perform a related exercise. After constructing migration chains through a direct survey, they estimate that for every 100 new housing units, 33 low-income households (defined, in 1969 dollars, as income less than \$1,000 plus \$500 times household size) will move. When I alter my estimation to count total moves from each category rather than equivalent units, I arrive at a similar figure of 43 moves from bottom income quintile and rent burdened areas, which form roughly the same percent of the population as Lansing et al.’s low-income definition.

month before selling in 2018, though this number reached 140 days in 2010.<sup>28</sup> The Federal Reserve Bank of St. Louis reports an average rental vacancy rate of 7 percent in 2018, with a peak near 11 percent in 2010.<sup>29</sup> Taken together, these numbers suggest an upper bound of one to three months for each round of the migration chain.<sup>30</sup> Since nearly all equivalent units in below-median income areas are created by round 15, these benefits should be felt within one to four years. For lower income areas, most equivalent units are created by round 20, which should be reached in two to five years.

One benchmark for these estimates is provided by the inclusionary zoning requirements for new housing developments in many cities. They mandate that developers fund a certain number (typically between 0.05 and 0.2) of income-restricted units per market-rate unit constructed (Schuetz, Meltzer, and Been 2009, Thadden and Wang 2017). While policies vary across cities, these units typically target a maximum income between 50 and 150 percent of area median income, making equivalent units in below-median income tracts a reasonable comparison. My estimates imply that a new market-rate unit has a significantly bigger effect on below-median income housing through market mechanisms than through inclusionary zoning requirements—generally at least three times as large. Demolitions of older, more affordable, housing units on the site of new construction provide another interesting benchmark. Since 100 new market-rate units create about 70 below-median income equivalent units, new construction must contain at least 14 new units for every 10 such units that are demolished in order for the equivalent units to outnumber the demolished units.

## 6.2 Alternative Specifications

In this section, I explore a number of alternative assumptions on chain decay and counterfactual locations in order to construct a range of reasonable equivalent unit estimates. While estimates change somewhat across simulations, they consistently suggest that new

---

<sup>28</sup>Zillow, “What Is the Average Time to Sell a House?” <https://www.zillow.com/sellers-guide/average-time-to-sell-a-house/> (accessed July 2, 2019).

<sup>29</sup>Federal Reserve Bank of St. Louis, “Rental Vacancy Rate for the United States.” <https://fred.stlouisfed.org/series/RRVRUSQ156N> (accessed July 2, 2019).

<sup>30</sup>These numbers represent an upper bound because multiple rounds of the migration chain can happen simultaneously. Suppose, for example, that one household that would have moved from Unit A to Unit B instead moves to the new building, leading another household searching at the same time to locate in Unit B instead of Unit C.

construction has a large effect on lower-income submarkets. The connectivity implied by the migration data is strong enough that migration chains frequently reach low-income areas even if they end at a relatively high rate or if households' counterfactual locations are a large step up from their origins.

First, recall that in the baseline specification I assumed that household formation and migration across CBSAs were unaffected by new construction. This is likely not the case for major expansions in housing supply. Since many proposed and enacted policies—such as the recent elimination of single-family zoning in the city of Minneapolis—could have a large effect on housing supply, I rerun the simulation allowing for new construction to affect these two forces. I draw a baseline average for new household formation from the 2018 Current Population Survey, which estimates that 11.5 percent of moves were to form a new household.<sup>31</sup> For across-metro migration, I refer to Table 1, which shows that 32.8 percent of people in my sample of new units originated from outside the metropolitan area. I then rerun the simulation assuming that these figures represent the average rates in each round of the chain and that new construction has a very large effect, increasing each force by 25 percent. This doubles the mean decay rate of 10 percent in the baseline specification. I use this as an upper bound on the decay rate, but I also consider the more extreme assumption that new construction increases each force by 50 percent.

Results appear in rows 2 and 3 of Table 3. Under the 25 percent assumption, below-median income equivalent units fall from the baseline of 70 to 45, a drop of 37 percent. The number in the lowest income categories falls by more—59 percent in the bottom income quintile—because the effect of a higher decay rate increases exponentially in each round and these types typically appear later in the chain. Under the more extreme 50 percent assumption, below-median income equivalent units fall to 31, and the number in the bottom income quintile falls to 9.3. While this exercise is speculative, it suggests that even a supply shock that sparked very large changes in household formation and migration would still have a meaningful effect on the housing market in below-median income areas. Even in the bottom income quintile, the equivalent unit count remains as large as many inclusionary

---

<sup>31</sup>While the data does not allow me to observe household formation within the chain, Turner (2008) directly observes household formation within chains in Stockholm and arrives at a similar figure of 12.7 percent.

zoning policies.

Next, I explore sensitivity to my counterfactual location assumption. The fourth row of Table 3 shows results under the alternative assumption that individuals would have moved up by two income deciles. The number of equivalent units in below-median income areas decreases by less than 10 percent, from 70 to 64, and the effect is similarly small in other categories. Of course, bottom-quintile income equivalent units fall to zero by assumption. When I instead assume that individuals would have moved up by one income and one percent white decile, the decrease is larger (about 40 percent) for areas that are below-median income and in the bottom quintile of percent white.

Third, I tighten restrictions on the underlying data. Including only new buildings that are above the seventh decile for both median household income and income per capita has minimal effect, as shown in row 6 of the table. In row 7, I eliminate moves to submarkets that are more than four income deciles higher than the origin submarket, reasoning that the housing units underlying these moves may not be reflective of their submarket more broadly. This reduces the number of equivalent units in below-median and bottom-quintile income areas to 55.1 (−21.5 percent) and 20.5 (−48.2 percent), respectively.

Finally, I classify housing units according to their census tract, which may not accurately reflect a unit’s characteristics. As previously shown in Appendix Figure A.7, units are somewhat positively selected in early rounds. To account for this, I compute a transition matrix between submarkets using only the individuals that moved into the units in the first round of the chain. I substitute this matrix into the first round of the simulation and find that equivalent units fall by only 5-10 percent, as shown in row 8 of Table 3.

### 6.3 Heterogeneity across Cities and Mechanisms

Equivalent unit creation is driven by two key factors: connectivity between market-rate units and other submarkets and the decay rate. If neighborhoods are more connected, the migration chain will be more likely to move from wealthy areas to lower income areas. In contrast, the decay rate determines how many rounds the migration chain has to reach the area of interest. To illustrate the relative importance of these forces, I first plot bottom-income quintile equivalent units separately for each city in the sample in Figure 7. The

numbers underlying the figure are shown in the first column of Table 4. There is substantial dispersion around the average of 39.6, from about 18.2 in New York to 60.7 in Denver.

Next, I repeat the simulation with a constant decay rate of 0.9 in all cities. Results appear in the second column of Table 4. The difference between New York and Denver falls from 42.5 to 36.6, suggesting that connectivity is the more important difference between the two cities. However, this depends on the cities being compared—the gap between Atlanta and Minneapolis falls from 25.9 to 8. It appears that both factors can be quantitatively important for explaining differences across cities. Connectivity differences may reflect demographic composition as much as attitudes—the difference between the average tract in the top- and bottom-income quintiles in Minneapolis is \$86,400, versus \$117,100 in New York.

My definition of submarkets uses within-CBSA characteristic deciles, implying that each submarket has more housing units in larger CBSAs. Since within-submarket moves do not change the state variable in the simulation, this could depress equivalent units in larger CBSAs, consistent with New York and Chicago’s low totals. To test this, I remove within-submarket moves from the simulation and rescale the probability of other moves proportionally. Column three shows that this has a small effect that is not larger in bigger CBSAs, likely because the submarket definition is relatively fine. Columns four through six repeat the above exercises with below-median income equivalent units.

## 6.4 Policy Discussion

These results, as well as the descriptive results from Section 4, suggest that new market-rate housing construction can improve housing affordability for middle- and low-income households, even in the short run. The effects are diffuse and appear to benefit diverse areas of a metropolitan area. Policies that increase market-rate construction are thus likely to improve affordability, even outside of the submarkets where new construction occurs. In addition to formal policies, these results also suggest that if policymakers expend the political capital required to get new housing proposals through the often subjective and onerous approval process, there are likely to be benefits throughout the region.

However, there are several shortcomings of market mechanisms, particularly in the lowest cost and most rent-burdened submarkets. Census tracts in the bottom quintile of median

household income and the top quintile of rent burden have an average vacancy rate of 12.8 percent, compared to 8.1 in the rest of my sample. Given that rents are generally already low in such neighborhoods, this suggests that reducing demand through the migration chain mechanism is unlikely to lower costs further, perhaps because rents have reached the minimum cost of housing. In addition to potentially small price effects, there may also be important amenity effects if the migration chain reduces population in these areas, such as reduced retail options, school closures, or increased crime. However, the relationship between income and vacancy rates differs across cities—in New York City, vacancy rates in low-income and rent burdened tracts are 9.7 percent versus 8.8 percent in other tracts, while the figures are 20.8 and 8.4 percent in Chicago. Market mechanisms will likely be more effective at reducing prices in low-income areas where vacancy rates are low.

Inclusionary zoning policies, which directly trade off market-rate construction and subsidized housing, provide an interesting perspective on the two policy approaches. Requiring developers to fund income-restricted units is a costly tax that anecdotally crowds out development (e.g., Dineen 2018), though a small academic literature has found null or small effects (Mukhija et al. 2010; Schuetz, Meltzer, and Been 2011). A back-of-the-envelope calculation suggests that if each required income-restricted unit crowds out more than 1.42 new market-rate units, the lost equivalent units in below-median income areas would outnumber the gain in income-restricted units. However, the income-restricted units offer benefits that market mechanisms do not. They can be rented for arbitrarily low prices, and they do not require a lag after a building’s completion. In addition, policymakers can dictate the location of these units.

## 7 Conclusion

The short-run effect of new market-rate housing on the market for middle- and low-income housing is crucial to the current policy debate, where government intervention and market-based strategies are often pitted against each other. My results suggest that new market-rate construction loosens the housing market in such areas and, moreover, could do so in less than five years. This implies that market-based strategies can play an important

role in improving housing affordability for middle- and low-income households.

However, an important caveat to these results is that I focus on quantity-based metrics rather than prices. This a particular concern for housing that is already extremely low-cost, as market mechanisms cannot induce for-profit landlords to lower prices below marginal cost. Vouchers or policies that reduce the marginal cost of providing housing (such as property tax or utility rate reductions) may be necessary to lower prices in this segment of the market. In addition, while I focus on regional effects, new buildings could have very different effects on their immediate area, where they may change amenities or household composition in ways that affect prices. There is little existing direct evidence on either the price effects or the local effects of new construction, and both could be fruitful areas for future research.

## Bibliography

- Albouy, D., Ehrlich, G., & Liu, Y. (2016). Housing demand, cost-of-living inequality, and the affordability crisis. National Bureau of Economic Research.
- Anenberg, E., & Kung, E. (2018). Can more housing supply solve the affordability crisis? Evidence from a neighborhood choice model. *Regional Science and Urban Economics*.
- Been, V., Ellen, I. G., & O'Regan, K. (2019). Supply skepticism: Housing supply and affordability. *Housing Policy Debate*, 29(1), 25-40.
- Braid, R. M. (1981). The short-run comparative statics of a rental housing market. *Journal of Urban Economics*, 10(3), 286-310.
- Diamond, R., McQuade, T., & Qian, F. (forthcoming). The effects of rent control expansion on tenants, landlords, and inequality: Evidence from San Francisco. *American Economic Review*.
- Dineen, J.K. (2018, August 27). "SF residential projects languish as rising costs force developers to cash out." *San Francisco Chronicle*.
- Ganong, P., & Shoag, D. (2017). Why has regional income convergence in the US declined? *Journal of Urban Economics*, 102, 76-90.
- Glaeser, E. L., & Gyourko, J. (2003). The impact of building restrictions on housing affordability. *FRBNY Economic Policy Review*, June, 21-39.
- Glaeser, E. L., & Ward, B. A. (2009). The causes and consequences of land use regulation: Evidence from Greater Boston. *Journal of Urban Economics*, 65(3), 265-278.
- Gyourko, J., & Molloy, R. (2015). Regulation and housing supply. In *Handbook of Regional and Urban Economics* (Vol. 5, pp. 1289-1337). Elsevier.
- Hsieh, C. T., & Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2), 1-39.
- Ihlanfeldt, K. R. (2007). The effect of land use regulation on housing and land prices. *Journal of Urban Economics*, 61(3), 420-435.
- Kristof, F. S. (1965). Housing policy goals and the turnover of housing. *Journal of the American Institute of Planners*, 31(3), 232-245.
- Lansing, J. B., Clifton, C. W., & Morgan, J. N. (1969). *New Homes and Poor People: A Study of Chains of Moves*. Ann Arbor, MI: Institute for Social Research, University of

Michigan.

- Logan, T. (2018, September 10). "Boston's new luxury towers appear to house few local residents." *Boston Globe*.
- Mukhija, V., Regus, L., Slovin, S., & Das, A. (2010). Can inclusionary zoning be an effective and efficient housing policy? Evidence from Los Angeles and Orange Counties. *Journal of Urban Affairs*, *32*(2), 229-252.
- Nathanson, C. G. (2019). Trickle-down housing economics. *Mimeo*.
- Piazzesi, M., Schneider, M., & Stroebel, J. (forthcoming). Segmented Housing Search. *American Economic Review*
- Quigley, J. M., & Raphael, S. (2005). Regulation and the high cost of housing in California. *American Economic Review*, *95*(2), 323-328.
- Rosenthal, S. S. (2014). Are Private Markets and Filtering a Viable Source of Low-Income Housing? Estimates from a "Repeat Income" Model. *American Economic Review*, *104*(2), 687-706.
- Rothenberg, J., Galster, G. C., Butler, R. V., & Pitkin, J. R. (1991). *The maze of urban housing markets: Theory, evidence, and policy*. University of Chicago Press.
- Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, *125*(3), 1253-1296.
- Scanlon, K., Whitehead, C., Blanc, F., & Moreno-Tabarez, U. (2017). The role of overseas investors in the London new-build residential market. LSE Consulting.
- Schuetz, J., Meltzer, R., & Been, V. (2009). 31 Flavors of inclusionary zoning: Comparing policies from San Francisco, Washington, DC, and suburban Boston. *Journal of the American Planning Association*, *75*(4), 441-456.
- Schuetz, J., Meltzer, R., & Been, V. (2011). Silver bullet or Trojan horse? The effects of inclusionary zoning on local housing markets in the United States. *Urban Studies*, *48*(2), 297-329.
- Solomont, E.B., & Sun, K. (2019, April 1). "NYC's ghost towers." *The Real Deal*.
- Sweeney, J. L. (1974). A commodity hierarchy model of the rental housing market. *Journal of Urban Economics*, *1*(3), 288-323.
- Thaden, E., & Wang, R. (2017). Inclusionary housing in the United States: Prevalence,

impact, and practices. *Mimeo*.

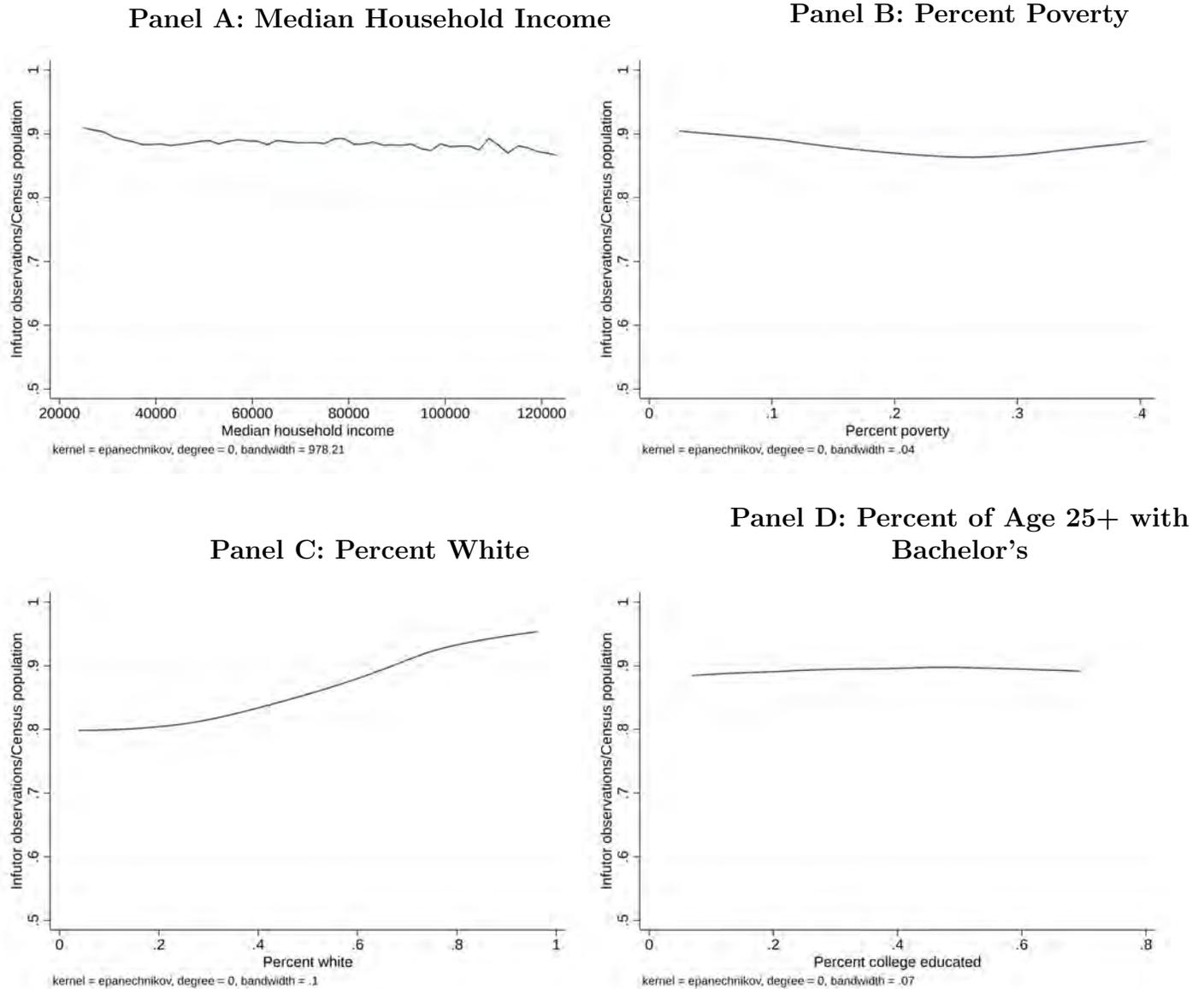
Turner, L. M. (2008). Who gets what and why? Vacancy chains in Stockholm's housing market. *European Journal of Housing Policy*, 8(1), 1-19.

Turner, L. M., & Wessel, T. (2019). Housing market filtering in the Oslo region: pro-market housing policies in a Nordic welfare-state context. *International Journal of Housing Policy*, 1-26.

Weicher, J. C., Eggers, F. J., & Moumen, F. (2016). The long-term dynamics of affordable rental housing. Washington, DC: Hudson Institute.

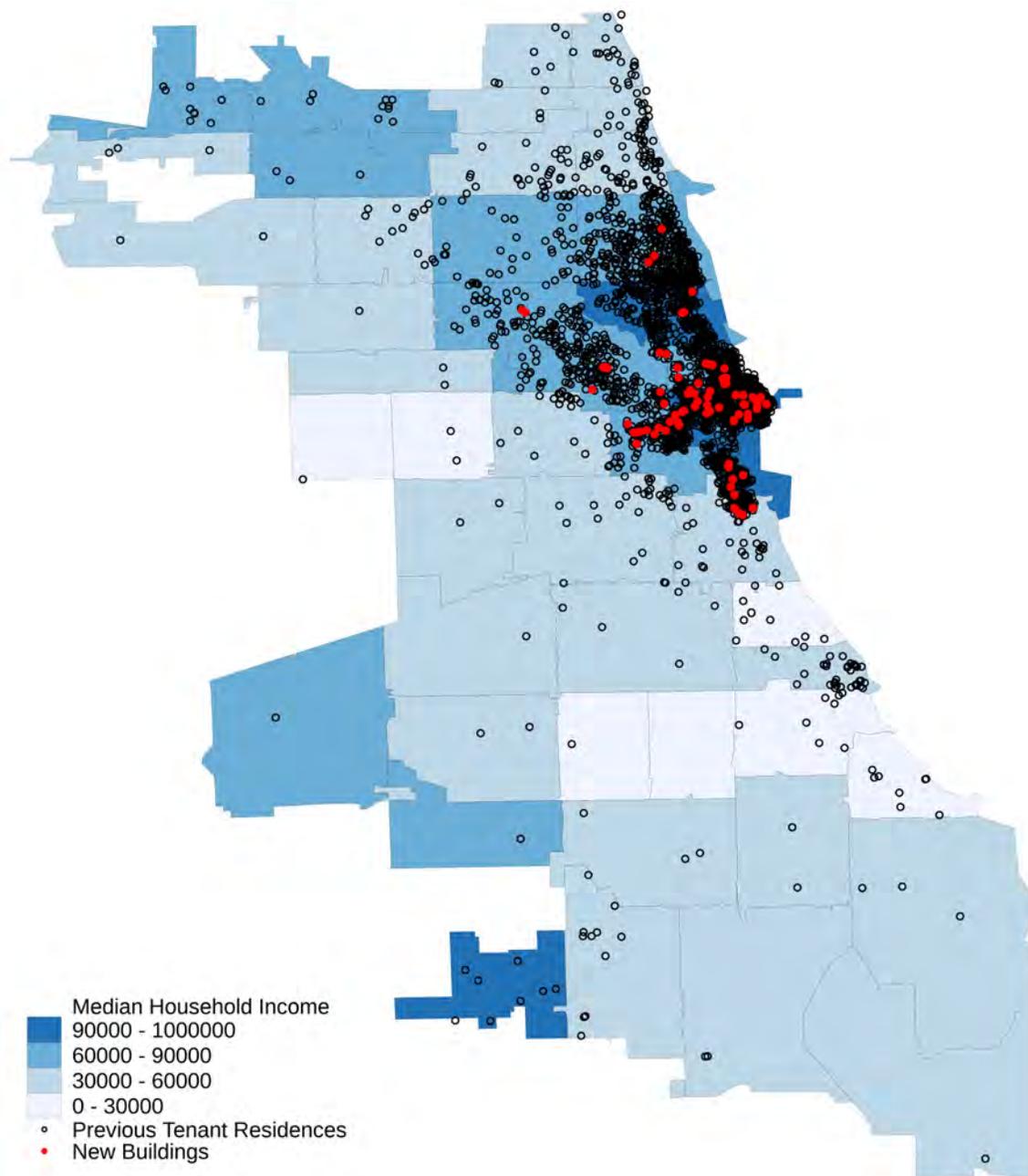
## 8 Figures

Figure 1: Infutor vs. Census Population (census tract level)



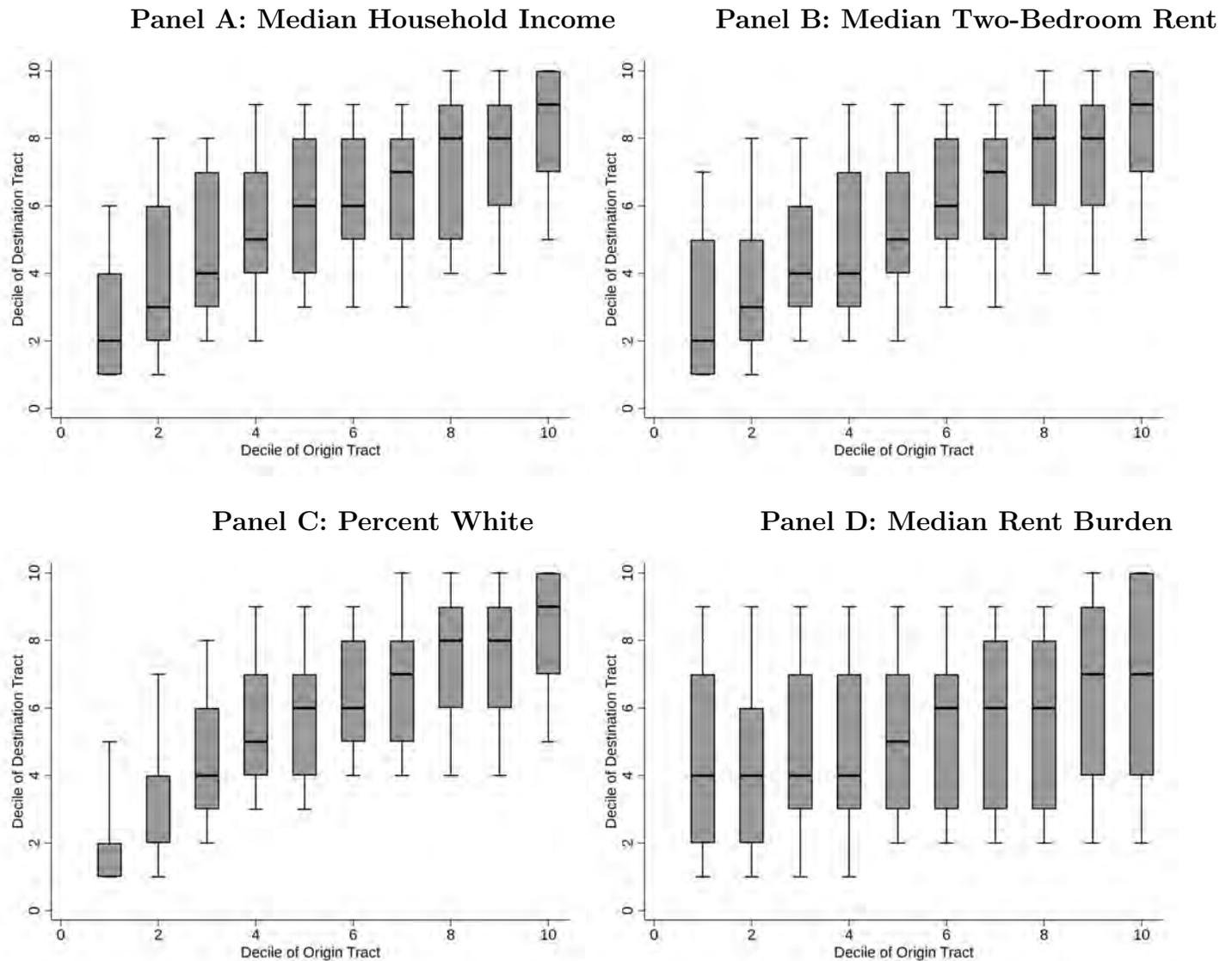
NOTE: Each panel plots a local polynomial regression of Infutor coverage (measured as the ratio of Infutor observations to census over-25 population) in a census tract versus the tract characteristic in the heading. Tract characteristics are drawn from the 2013–2017 ACS.

Figure 2: Prior Residence of Tenants of New Buildings in Chicago



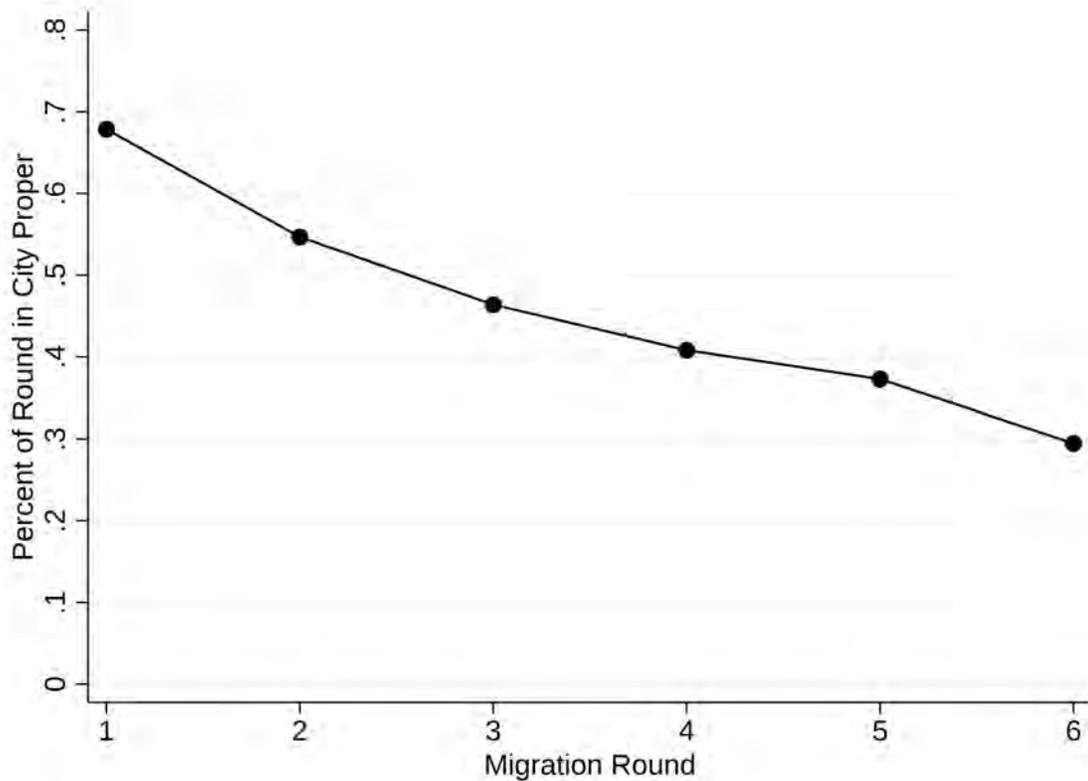
NOTE: Solid red dots represent the location of market-rate apartment buildings completed since 2010. Hollow black dots represent the previous residences of the current tenants in those buildings. The base map polygons are zip codes in Chicago proper, colored according to median household income in the 2013–2017 ACS. Only residents whose prior residence was within the city proper are included. Small amounts of noise are added to each marker to avoid precisely identifying addresses

Figure 3: Migration between Census Tracts in Chicago Metropolitan Area



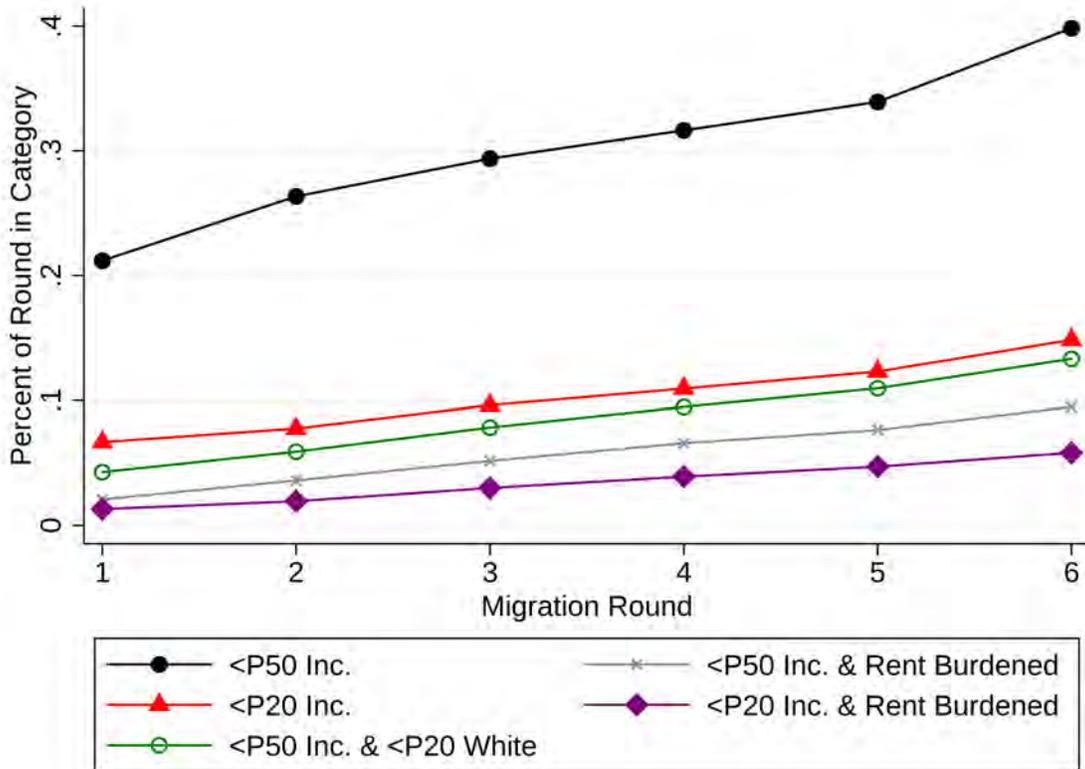
NOTE: This figure shows the distribution of destination neighborhood characteristics conditional on origin neighborhood characteristics for migrants within the Chicago CBSA in 2010–2017. Within each panel, each box plot represents migrants who originated in a tract of a given decile of the characteristic in the heading. The box then shows the median and interquartile range of the same characteristic in the destination tracts of those migrants. The whiskers represent 10th and 90th percentiles. Characteristic deciles are calculated within the CBSA.

Figure 4: Percent of Individuals Originating in CBSA Principal City by Migration Round



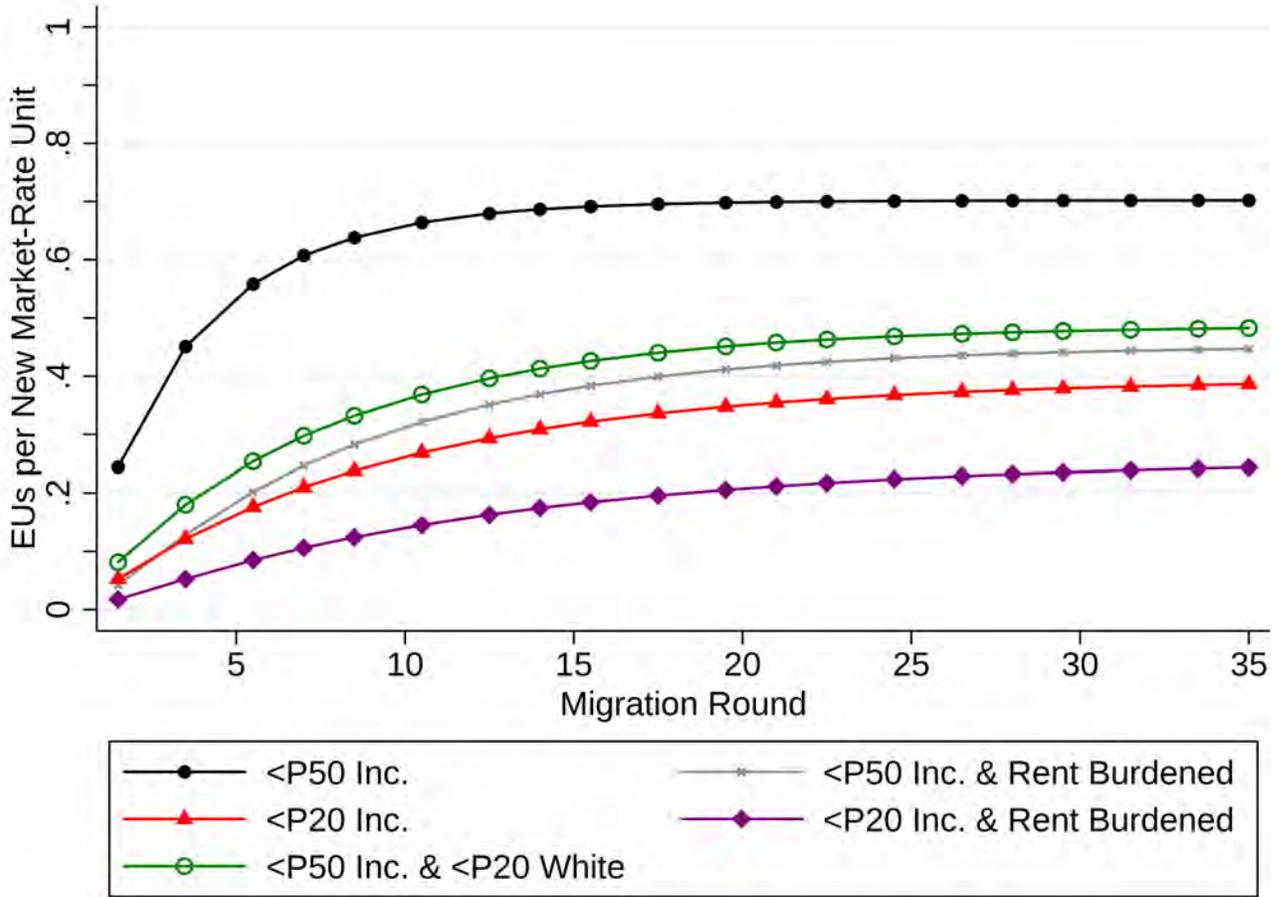
NOTE: This figure plots the percentage of the individuals in each round of the sequence of origin units that originated within the same city as the new building. Note that only migrants from the same metropolitan area as the new building are included in each round. Round 1 is the origin units of the individuals currently occupying the new unit; round 2, the origins of the individuals occupying round 1's origin buildings, and so on. Each subsequent round is constructed by observing the set of individuals currently living in the previous round's origin buildings, not their specific units, and the sequence is reweighted accordingly. The sequences begin with 52,000 individuals living in 686 new market-rate buildings.

Figure 5: Percent of Individuals Originating in Tract Categories by Migration Round



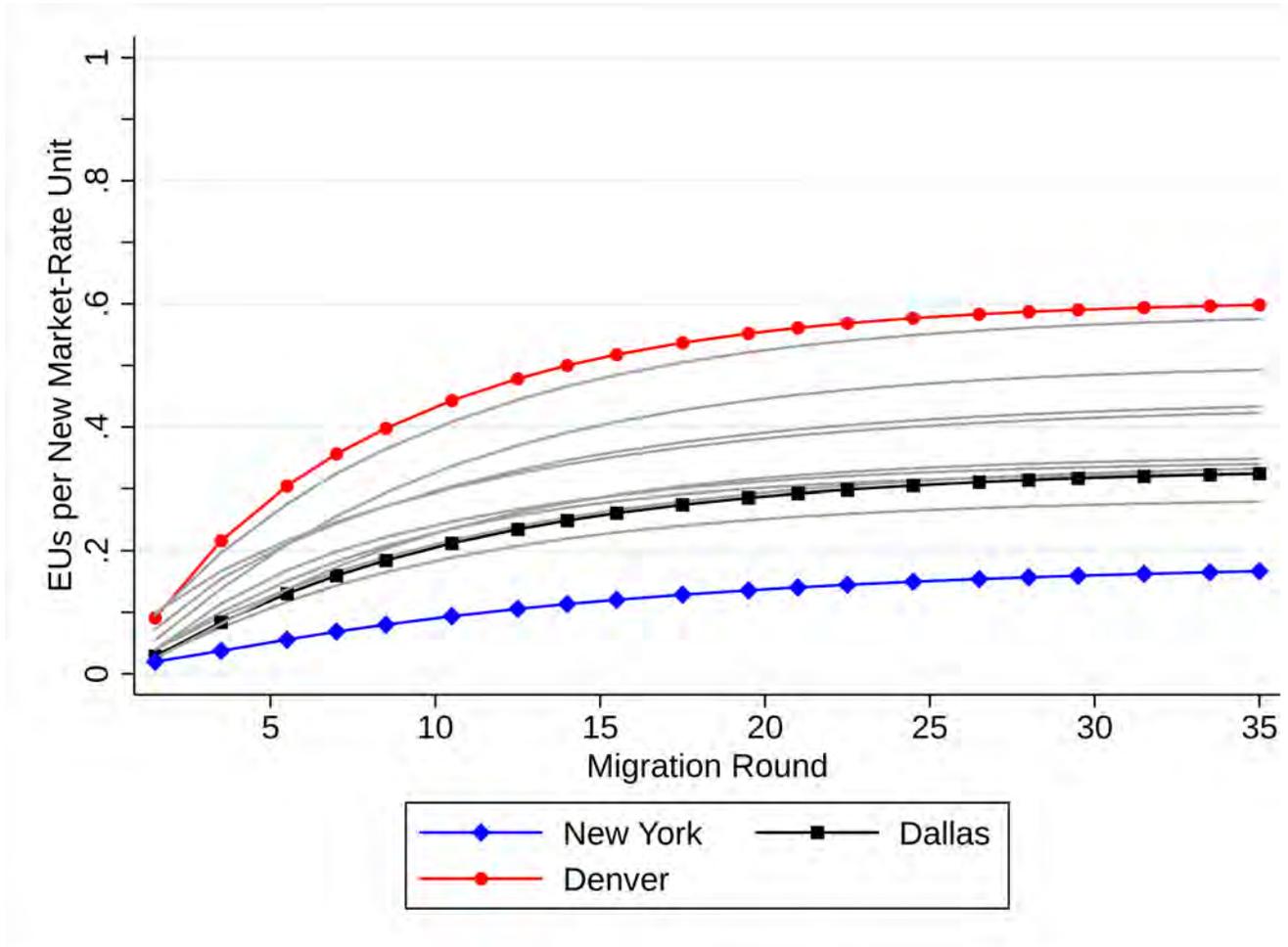
NOTE: This figure plots the percentage of the individuals in each round of the sequence of origin units that originated in a census tract with a given set of characteristics. Note that only migrants from the same metropolitan area as the new building are included in each round. Round 1 is the origin units of the individuals currently occupying the new unit; round 2, the origins of the individuals occupying round 1's origin buildings, and so on. Tract characteristics are taken from the 2013–2017 ACS, and all quantiles are computed within CBSAs. Income is median household income, and rent burdened is defined as in the top quintile of rent burden for the CBSA. Each subsequent round is constructed by observing the set of individuals currently living in the previous round's origin buildings, not their specific units, and the sequence is reweighted accordingly. The sequences begin with 52,000 individuals living in 686 new market-rate buildings.

Figure 6: Equivalent Unit Creation



NOTE: This figure shows the expected number of equivalent units created by a new market-rate unit, cumulative across rounds of the migration chain. An equivalent unit is created in a category when a migration chain reaches such an area for the first time, thus reducing demand for that category by one. Tract characteristics are taken from the 2013-2017 ACS, and all quantiles are computed within CBSAs. Income is median household income, and rent burdened is defined as in the top quintile of rent burden for the CBSA. Appendix Table A.2 shows the numbers underlying the figure.

Figure 7: Heterogeneity across Cities in Equivalent Unit Creation



NOTE: This figure shows heterogeneity across CBSAs in the number of bottom-quintile income equivalent units created by a new market-rate unit, cumulative across rounds of the migration chain. Each unmarked gray line represents a different metropolitan area, with Denver, New York, and Dallas highlighted as examples. Table 4 shows the numbers underlying the figure, as well as results under a number of alternative specifications.

Table 1: Number of New Buildings and Residents across Cities

City	New buildings	Infutor individuals	Percent from same CBSA	Percent from same city
Atlanta	44	3,641	0.687	0.484
Boston	16	1,238	0.700	0.375
Chicago	84	7,068	0.728	0.578
Dallas	76	6,670	0.687	0.487
Denver	49	3,270	0.539	0.411
Houston	69	5,906	0.711	0.584
Minneapolis	37	2,206	0.714	0.483
New York City	89	7,835	0.764	0.682
Philadelphia	12	694	0.642	0.423
Seattle	101	6,334	0.598	0.472
San Francisco	38	1,704	0.632	0.512
Washington	71	5,866	0.658	0.505
Sample	686	52,432	0.672	0.500

NOTE: This table shows the number of new buildings in each city and the number of individuals currently living in those buildings in the Infutor data. The buildings, which are detected using the algorithm described in Section 3, must contain over 16 individuals in the Infutor data, be built since 2009, and be within five miles of their CBSA’s central business district and in a census tract with above median income for the CBSA. Individuals whose immediately previous address is in the same CBSA (city) as the new building are considered from the same CBSA (city).

Table 2: Building Characteristics

Percentile	Infutor individuals	Distance to CBD	Median household income decile	Income per capita decile	Median two-bedroom rent decile	Percent vacant (tract)	Percent vacant (block group)
Min	17	0.04	5	4	1	0.004	0.000
10	24	0.59	5	9	9	0.048	0.029
50	60	1.73	8	10	10	0.109	0.111
75	100	2.67	9	10	10	0.154	0.159
95	183	4.27	10	10	10	0.234	0.268
Max.	468	4.97	10	10	10	0.547	0.547
Mean	76.43	1.95	7.63	9.57	9.42	0.119	0.119
N	686	686	686	686	681	686	686

NOTE: This table shows characteristics of the new buildings. Distance to central business district (CBD) is given in miles. Median household income, income per capita, and median two-bedroom rent are determined using the building's census tract and the 2013–2017 ACS, and deciles of each are computed within CBSAs. Percent vacant is reported at both the tract and block group level and is also drawn from the ACS, which counts second homes as vacant.

Table 3: Equivalent Units under Alternative Specifications

Specification	Number of Equivalent Units in:				
	<P50 Inc.	<P50 Inc. & Rent Burdened	<P20 Inc.	<P20 Inc. & Rent Burdened	<P50 Inc. & <P20 White
Baseline	0.702	0.453	0.396	0.258	0.488
Assume 25% marginal	0.453	0.191	0.167	0.088	0.231
Assume 50% marginal	0.308	0.101	0.093	0.043	0.135
Plus 2 income deciles	0.644	0.422	NA	NA	0.472
Plus 1 income, 1 white decile	0.694	0.408	0.352	0.212	0.281
Only high-income new buildings	0.693	0.449	0.391	0.255	0.481
Trim large transitions	0.551	0.305	0.205	0.127	0.327
Use round 1 transitions	0.655	0.406	0.357	0.230	0.440

NOTE: This table shows the number of equivalent units created under various changes to the baseline specification. Row 1 is the baseline, while Rows 2 and 3 show results under the assumption that 25% and 50%, respectively, of household formation and across-metro migration are marginal to the new construction, which substantially increases the decay rate. Rows 4 and 5 change the counterfactual location assumption from moving up one income decile to moving up two income deciles or one income and one percent white decile, respectively. Row 6 only includes new buildings that are above the seventh decile of both median household income and income per capita. Row 7 removes transitions in which individual's tract income increases by more than four deciles. Row 8 uses a set of transitions computed from the first round of the sequence of origin units, as shown in Figure 5, to compute counterfactual locations in the first round of the simulation, instead of the transitions computed from the full sample.

Table 4: CBSA Heterogeneity in Equivalent Units

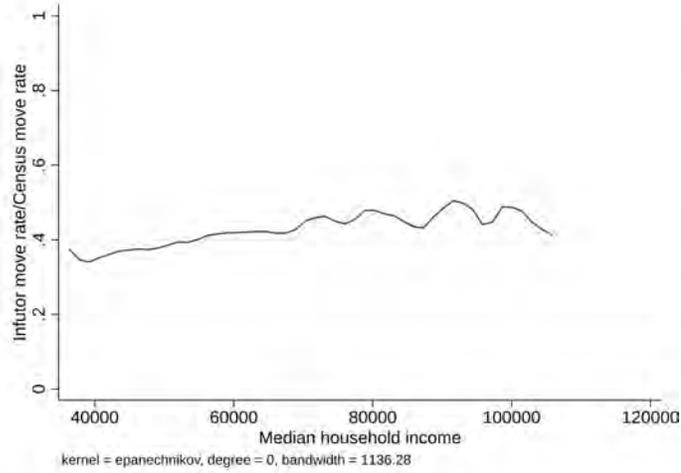
City	(1) <P20 Inc.	(2) <P20 Inc., decay=0.9	(3) <P20 Inc., no within moves	(4) <P50 Inc.	(5) <P50 Inc., decay=0.9	(6) <P50 Inc., no within moves
Atlanta	0.328	0.444	0.356	0.625	0.741	0.646
Boston	0.344	0.389	0.378	0.637	0.745	0.658
Chicago	0.287	0.363	0.310	0.604	0.706	0.625
Dallas	0.334	0.379	0.360	0.686	0.753	0.702
Denver	0.607	0.563	0.623	0.817	0.831	0.823
Houston	0.354	0.423	0.379	0.634	0.721	0.654
Minneapolis	0.587	0.524	0.607	0.844	0.848	0.849
New York	0.182	0.197	0.193	0.600	0.642	0.602
Philadelphia	0.346	0.416	0.373	0.711	0.765	0.728
Seattle	0.503	0.468	0.533	0.802	0.796	0.816
San Francisco	0.436	0.445	0.464	0.717	0.772	0.735
Washington	0.445	0.450	0.471	0.749	0.773	0.765

NOTE: This figure shows heterogeneity across CBSAs in the number of bottom-quintile and below-median income equivalent units created by a new market-rate unit. Columns 1 and 4 use the baseline specification, while columns 2 and 5 set the decay rate in every city to 0.9. Columns 3 and 6 eliminate moves within submarkets and proportionately redistribute the weight to other categories.

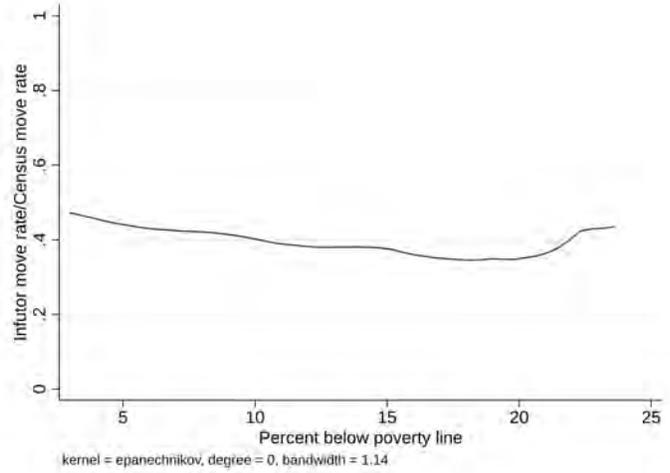
# A Appendix Figures

Figure A.1: Infutor vs. Census Migration Rates (county level)

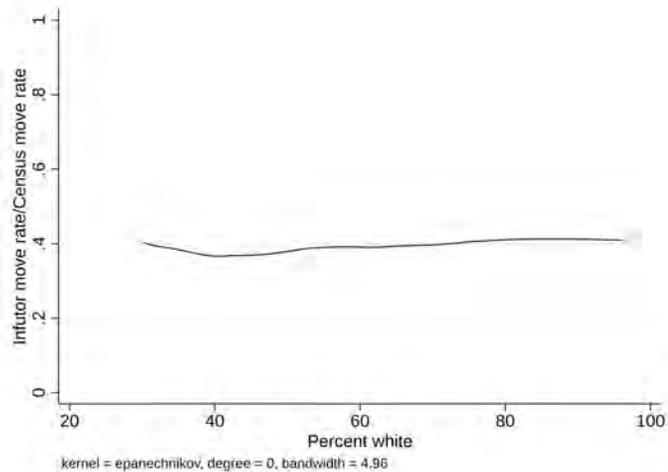
Panel A: Median Household Income



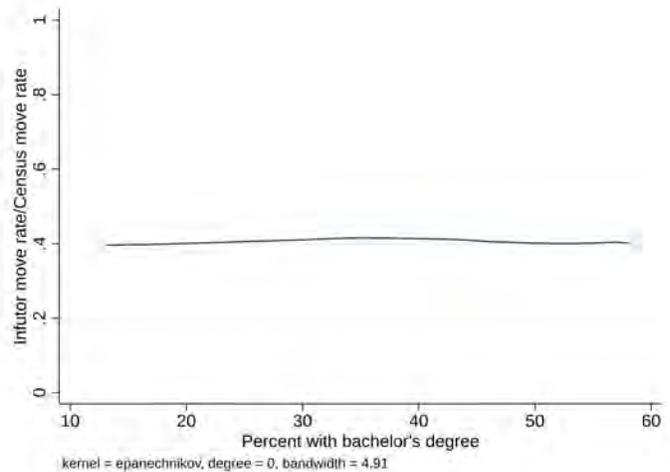
Panel B: Percent Poverty



Panel C: Percent White

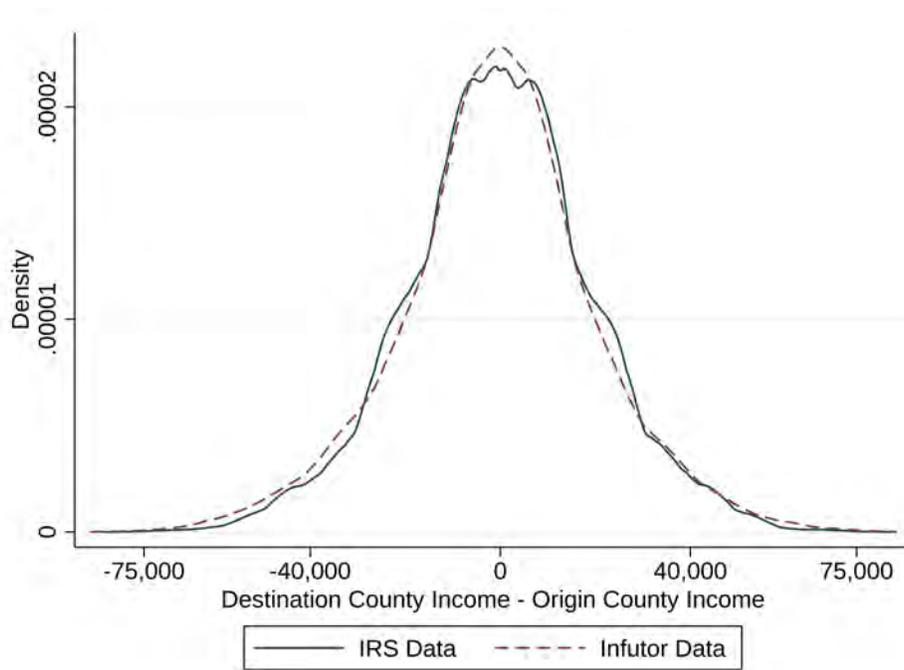


Panel D: Percent of Age 25+ with Some College



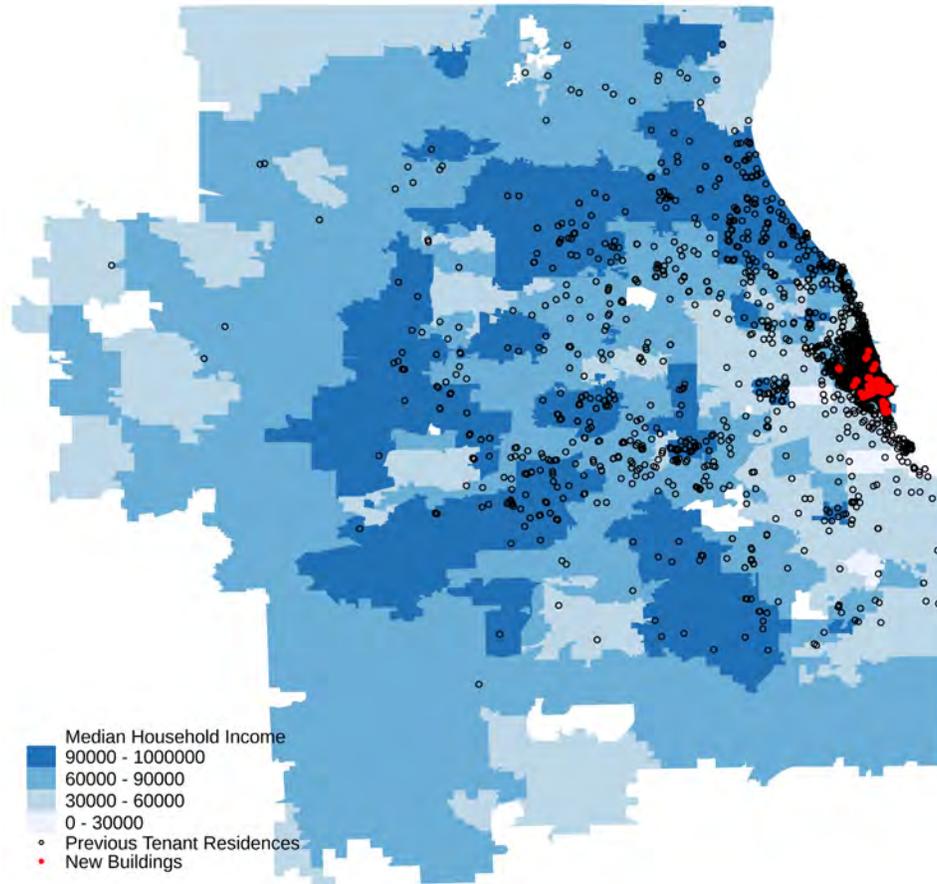
NOTE: Each panel plots a local polynomial regression of the ratio of Infutor to Census annual move rates (measured at the county level) against county characteristics. County characteristics and move rates are drawn from the 2013–2017 ACS.

Figure A.2: Difference between Destination and Origin County Income in Infutor versus IRS Data



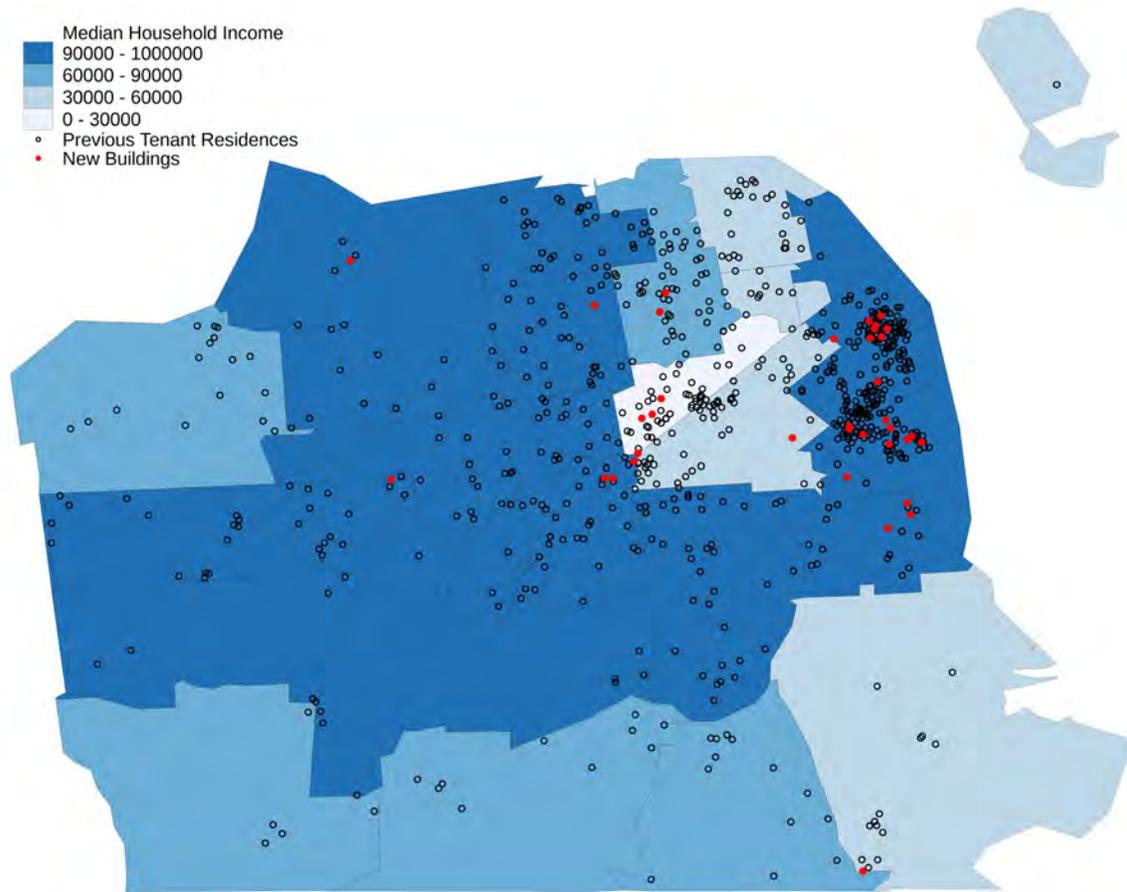
NOTE: This figure plots the distributions of destination county income and origin county income for moves in the Infutor data and 2018 IRS Statistics of Income data. County median household income is taken from the 2013–2017 ACS.

Figure A.3: Chicago Metro Origins



NOTE: Solid red dots represent the location of market-rate apartment buildings completed since 2010. Hollow black dots represent the previous residences of the current tenants in those buildings. The base map polygons are zip codes in the Chicago CBSA, colored according to median household income in the 2013–2017 ACS. Only residents whose prior residence was within the Chicago CBSA are included. Small amounts of noise are added to each marker to avoid precisely identifying addresses.

Figure A.4: San Francisco City Origins

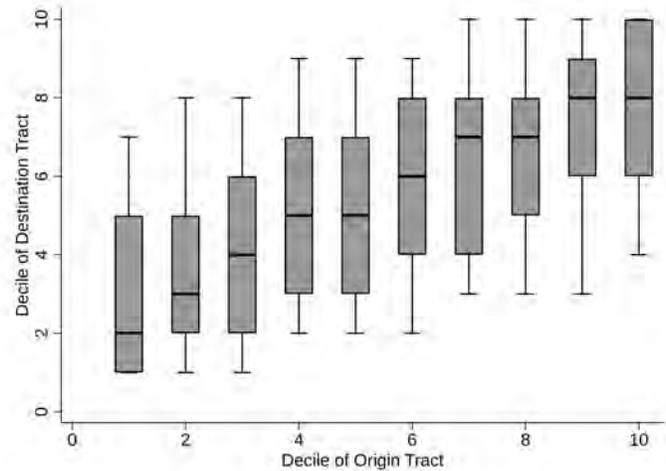
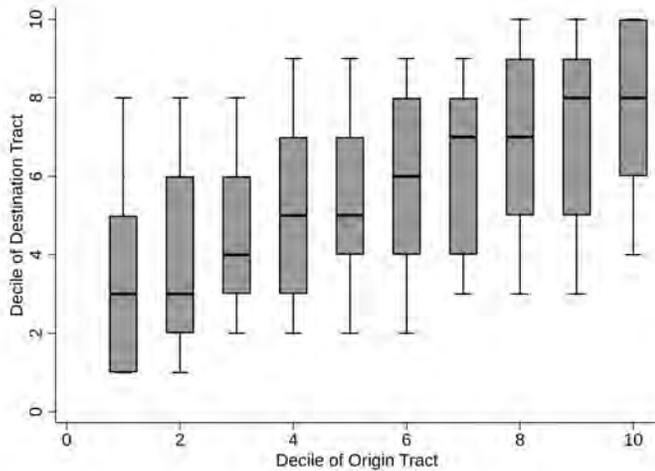


NOTE: Solid red dots represent the location of market-rate apartment buildings completed since 2010. Hollow black dots represent the previous residences of the current tenants in those buildings. The base map polygons are zip codes in San Francisco proper, colored according to median household income in the 2013–2017 ACS. Only residents whose prior residence was within the city proper are included. Small amounts of noise are added to each marker to avoid precisely identifying addresses.

Figure A.5: Migration between Census Tracts in San Francisco Metropolitan Area

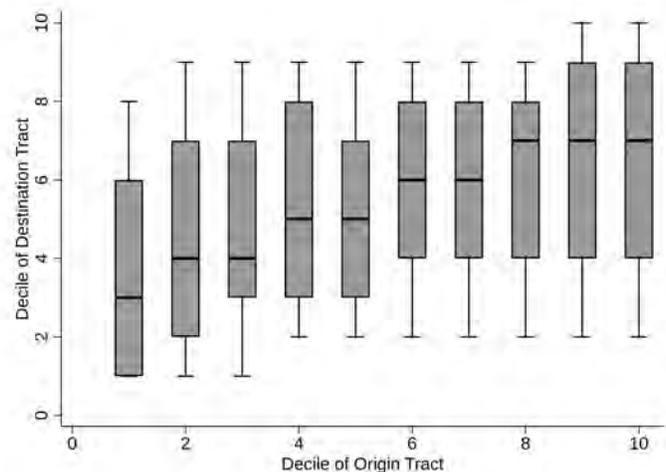
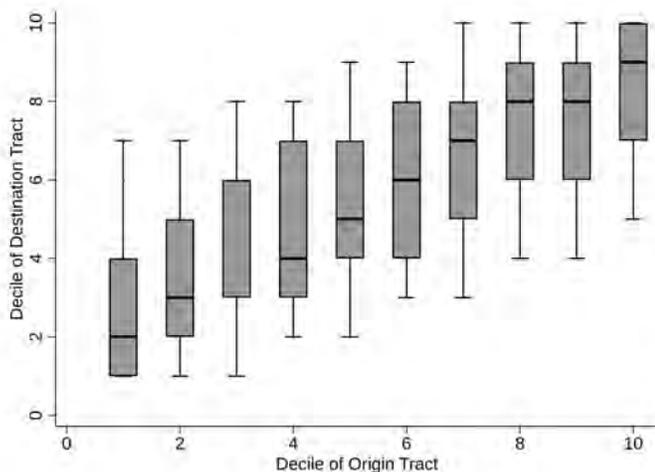
Panel A: Median Household Income

Panel B: Median Two-Bedroom Rent



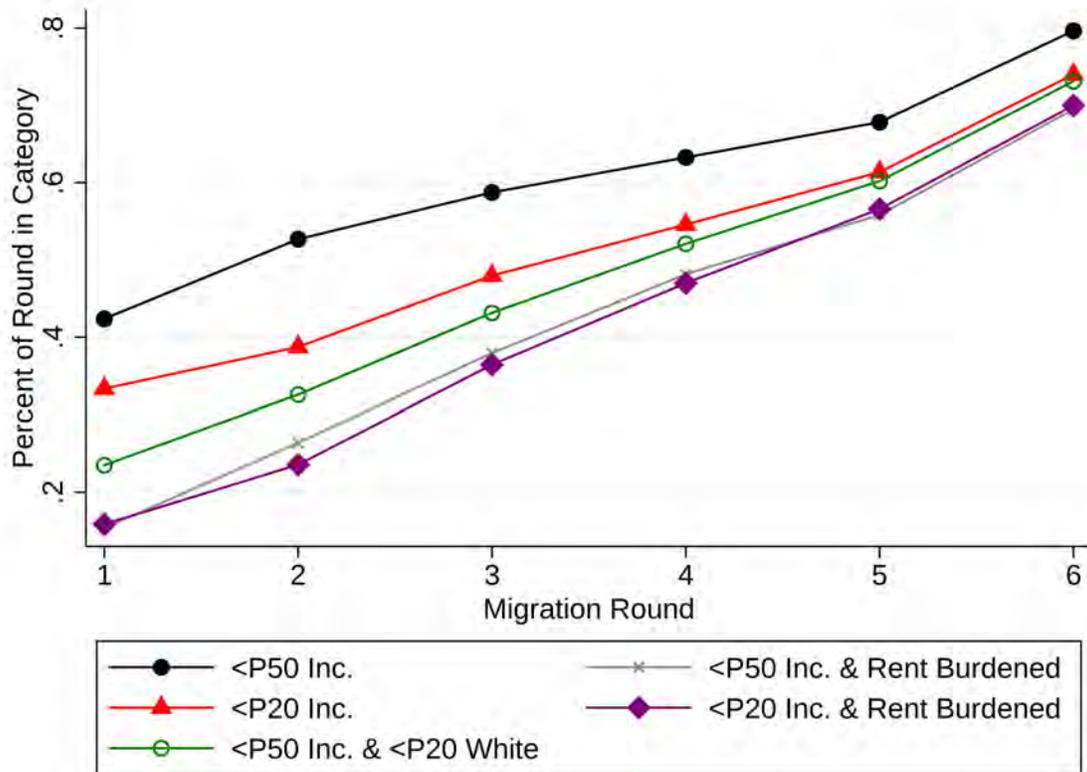
Panel C: Percent White

Panel D: Median Rent Burden



NOTE: This figure shows the distribution of destination neighborhood characteristics conditional on origin neighborhood characteristics for migrants within the San Francisco CBSA in 2010–2017. Within each panel, each box plot represents migrants who originated in a tract of a given decile of the characteristic in the heading. The box then shows the median and interquartile range of the same characteristic in the destination tracts of those migrants. The whiskers represent 10th and 90th percentiles. Characteristic deciles are calculated within the CBSA.

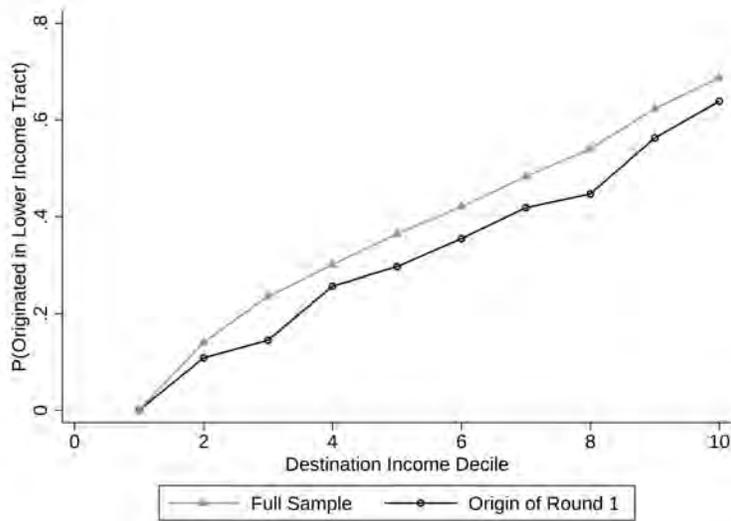
Figure A.6: Normalized Composition of Sequence of Origin Units



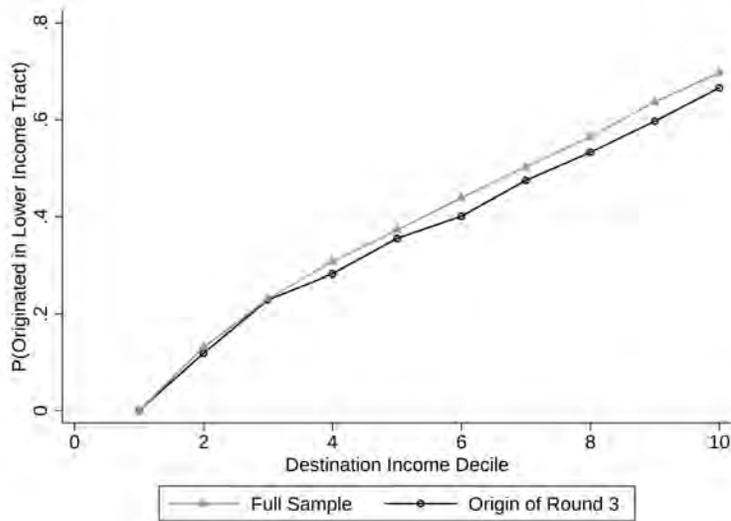
NOTE: This figure repeats Figure 5, normalizing each line by the percent of the CBSA population that lives in a given tract type.

Figure A.7: Percent of Units Filled by Individual from Lower-Income Decile for Full Sample versus Sequence of Origin Units

**Panel A: Round 1**

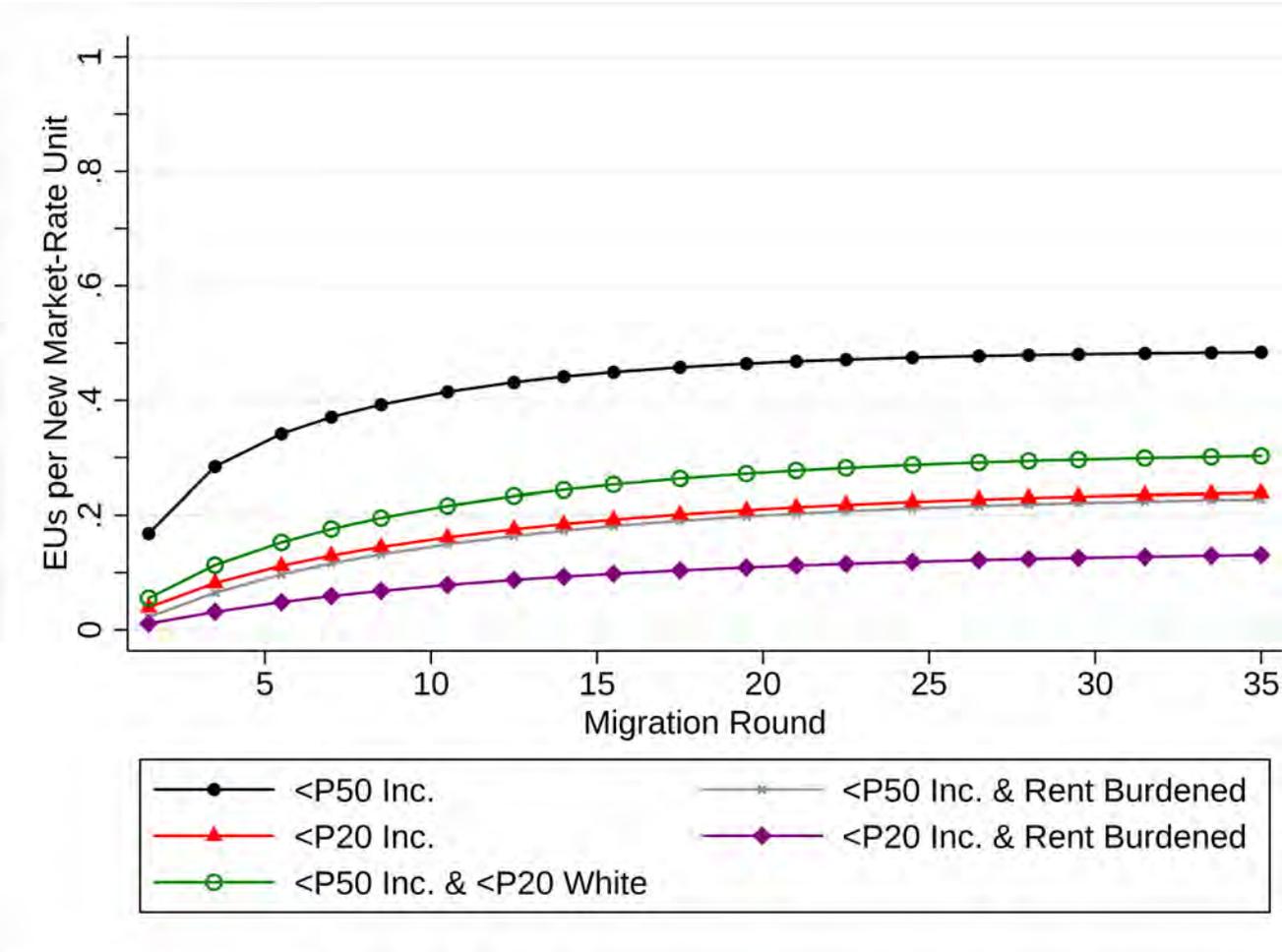


**Panel B: Round 3**



NOTE: Each line shows the probability that a unit in a given tract income decile was filled by a person who originated in a lower-income tract. The full sample line includes all units in the income decile, while the other line includes only units that were in round 1 (Panel A) or round 3 (Panel B) of the sequence of origin units.

Figure A.8: Equivalent Unit Creation within Principal Cities



NOTE: This figure repeats Figure 6, but includes only equivalent units in the principal city of each CBSA.

Table A.1: MSA Characteristic Deciles

<b>MSA</b>	<b>Median income</b>			<b>Percent white</b>			<b>Median rent burden</b>		
	P20	P50	P90	P20	P50	P90	P20	P50	P90
Atlanta	41,375	61,496	108,295	14.9%	51.0%	85.7%	24.7%	29.5%	40.1%
Boston	57,613	83,865	135,858	49.9%	78.9%	94.8%	25.2%	29.7%	39.9%
Chicago	40,024	62,601	113,019	9.5%	56.8%	87.6%	24.5%	30.1%	44.6%
Dallas	40,733	61,300	118,241	19.1%	48.8%	82.0%	24.3%	28.7%	38.1%
Washington	64,810	95,690	160,833	18.5%	50.4%	83.0%	24.2%	28.8%	40.2%
Denver	49,918	72,031	117,917	43.6%	73.3%	89.2%	24.9%	29.6%	38.5%
Houston	36,932	57,136	112,357	7.7%	34.5%	73.8%	24.2%	29.1%	39.1%
Minneapolis	52,471	72,357	111,406	62.0%	81.4%	94.5%	24.0%	28.6%	37.7%
New York	45,677	72,657	129,479	8.1%	50.1%	87.1%	26.7%	32.3%	46.8%
Philadelphia	41,897	68,152	116,053	28.1%	74.0%	92.3%	25.0%	30.7%	46.0%
Seattle	55,530	79,040	117,500	50.9%	68.8%	86.0%	24.7%	29.2%	37.3%
San Francisco	62,731	96,210	157,045	17.5%	40.6%	74.0%	24.2%	29.2%	39.2%

NOTE: This table shows CBSA deciles of the characteristics used to define submarkets. Characteristics are drawn from the 2013–2017 ACS.

Table A.2: Equivalent Unit Totals

Round	Number of Equivalent Units in:				
	<P50 Inc.	<P50 Inc. & Rent Burdened	<P20 Inc.	<P20 Inc. & Rent Burdened	<P50 Inc. & <P20 White
5	0.539	0.186	0.164	0.077	0.238
10	0.659	0.313	0.262	0.141	0.361
15	0.690	0.379	0.318	0.181	0.422
20	0.699	0.414	0.350	0.208	0.453
25	0.701	0.432	0.369	0.225	0.470
30	0.702	0.442	0.380	0.236	0.478
35	0.702	0.447	0.386	0.244	0.482
100	0.702	0.453	0.396	0.258	0.488

NOTE: This figure shows the expected number of equivalent units created by a new market-rate unit, cumulative across rounds of the migration chain. An equivalent unit is created in, for example, a below-median income submarket when a migration chain reaches such an area for the first time, thus reducing demand for that submarket by one. Tract characteristics are taken from the 2013–2017 ACS, and all quantiles are computed within CBSAs. Income is median household income, and rent burdened is defined as in the top quintile of rent burden for the CBSA. These numbers are plotted in Figure 6.