Suburban Housing and Urban Affordability: Evidence from Residential Vacancy Chains*

Robert French, Harvard Kennedy School

Valentine Gilbert, Harvard Kennedy School†

April 18, 2024

Please click here for the latest draft.

This paper shows how different housing submarkets are linked by residential vacancy chains – the series of moves across housing units initiated by the construction of new housing. Using administrative data on the residential histories of the U.S. population, we compare the characteristics of vacancies created by new suburban single-family homes to those created by new urban multifamily housing. We find that vacancy chains are short, with 90% ending within three rounds of moves; and, consequently, that each new suburban home leads to only .015 moves in low-income urban neighborhoods. We then conduct a simulation exercise to understand what the observed patterns of vacancy chains imply about the welfare and price effects of new housing supply. We show that the geographic distribution of moves created by vacancy chains is correlated with the geographic distribution of welfare and price effects, and that the number of vacancies created in a neighborhood is as strong a predictor of price effects as are model-derived cross-neighborhood substitution effects. Our results imply that new suburban housing supply has little effect on urban housing affordability or on the welfare of low-income urban households.

*Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2733. (CBDRB-FY24-P2733-R10951).

†Job Market Paper. Email: jvgilbert@g.harvard.edu. I thank my advisors, Ed Glaeser, Stefanie Stantcheva, and Adrien Bilal, for their valuable feedback, patience, and support. All errors are our own. We thank participants of the Economics and Social Policy Workshop for helpful comments. We thank Jim Davis, Sari Kerr, Shawn Ratcliffe, and Shital Sharma for help with the Census data. Oleg Kolev provided excellent research assistance. I gratefully acknowledge receiving financial support from the John R. Meyer Dissertation Fellowship of the Harvard Joint Center for Housing Studies while working on this research.

1. Introduction

Over the past thirty years, real home prices and rents in dense urban centers in the U.S. grew more rapidly than housing costs in suburban metropolitan neighborhoods (Howard and Liebersohn 2021).[1] At the same time, the vast majority of housing supply growth has been in low-density suburbs, which accounted for less than half of the metropolitan area housing stock in the U.S. in 1990 but more than 80% of the increase in housing supply between 1990 and 2018 (Baum-Snow 2023).[2] This pattern of rising prices in dense urban centers and expanding supply on the urban periphery reflects both a secular increase in the demand for urban amenities and the difficulty of building in already developed central neighborhoods (Couture et al. 2021; Couture and Handbury 2020; Glaeser and Gyourko 2003; Baum-Snow and Han 2021; Gyourko, Saiz, and Summers 2008). But the pattern also raises an important question about how U.S. cities can best address what is widely regarded as a housing-affordability crisis. Namely, can continued suburban expansion alleviate rising housing costs in the urban center, or will cities have to grow denser to become more affordable?

In this paper, we seek to answer this question by examining how different housing submarkets are connected by residential mobility. If residential mobility is high, then an increase in the supply of housing in one submarket can decrease housing costs in other disparate submarkets.[3] If, on the other hand, there is little mobility between housing submarkets, the benefits of expanding the supply of one type of housing will be concentrated among the residents of that type of housing. These two possibilities are represented in two sides of an ongoing debate about urban housing policy, with one side advocating for dramatically expanding the overall supply of housing and the other arguing that increasing the supply of market-rate housing will do little to make cities more affordable for low- and middle-income households that are unable to afford these new market-rate housing units (Been, Ellen, and O'Regan 2019; Gray 2021; Demsas 2022; Dougherty 2020; Friedrich et al. 2023).[4]

[1]Between 1990 and 2018, real rents in the 10% of tracts closest to the city center increasing by an average of 45%, substantially more than the average increase in real metropolitan rents of 30%. We make these and subsequent aggregate-level calculations using National Historical Geographic Information System (NHGIS) data (Manson et al. 2023).

[2]By contrast, the supply of housing in the 10% of tracts closest to city centers increased by less than one million units, accounting for less than 3% of the growth in metro area housing supply.

[3]In the extreme case of perfect mobility, then spatial equilibrium implies that all households are equally well off and thus that expanding the supply of housing of one type will benefit all residents of the city equally.

[4]The pro-supply and supply-skeptical sides of this debate are respectively summarized by the con-

Importantly, expanding the supply of one type of housing can make other types of housing more affordable even if there is little residential mobility directly between the two submarkets, so long as the submarkets are connected by a chain of residential moves that pass through other submarkets. This paper directly studies these *residential vacancy chains* – the series of moves initiated by the construction of a new housing unit.[5] The first migration round, or "link", in a vacancy chain consists of moves into a newly constructed housing unit, which potentially leave the movers' origin units vacant. The second link consists of moves into these vacated units – moves that generate their own set of vacancies. The chain continues on in this way until it ends, either because the origin unit is not vacated[6], because the vacated origin unit remains vacant or is demolished, or mover's origin unit lies outside the market under consideration, as in the case of a move originating outside the U.S.

In the first part of this paper, we motivate our focus on vacancy chains with a simple model and then present new descriptive facts about the vacancy chains initiated by different kinds of new housing. We show that vacancy chains are relatively short and that new suburban housing supply generates few moves in urban neighborhoods. In the second part of the paper, we conduct a simulation exercise to understand the economic consequences of the descriptive facts we document. We find that the number of vacancies created in a neighborhood is strongly correlated with the price and welfare effects of new housing. These simulation results, when applied to our descriptive findings, imply that the geographic and distributional incidence of the benefits of new housing supply depend importantly on where and what kind of new housing is built.

To fix ideas, we begin by presenting a simple stylized model of a differentiated housing market and derive a general expression for the effect of additional supply in one submarket on prices in other submarkets. We show that even in this simple model, this price effect can be decomposed into a direct effect and an indirect effect, with the indirect effects depending on a chain of cross-submarket residential substitution terms.[7] The expression for the indirect price effect in our model illustrates the po-

trasting headlines, "Build Build Build Build Build Build Build Build Build Build Build Build Build Build" (Dougherty 2020) and "More Building Won't Make Housing Affordable" (Friedrich et al. 2023).

[5]Vacancy chains may be initiated by other events, such as a death that creates a new vacancy or the consolidation of two households. Because we are interested in understanding the effects of new housing construction, we focus only on vacancy chains initiated by the creation of a new housing unit.

[6]The origin unit of a move is not vacated when a new household is formed, as in the case when a roommate moves out to live on her own.

[7]More precisely, the indirect effects depend on the product of residential diversion ratios. The residential diversion ratio between neighborhoods m and n captures the share of households that leave m in response to rising housing costs that substitute towards n.

tential importance of vacancy chains as a mechanism connecting different housing submarkets.

Our descriptive characterization of residential vacancy chains is one of the main contributions of our paper. We use newly available administrative data from the Census Bureau on the residential histories of the entire U.S. population from 2000 to 2021 and on the inventory of U.S. residential housing units in 2022. We use these data to identify 1.5 million new single family suburban and multifamily urban housing units built between 2009 and 2018 in the 17 most populous metropolitan areas in the U.S. We then construct the vacancy chains initiated by these units, tracing their paths through different kinds of neighborhoods. We document that while vacancy chains that grow long enough do connect disparate housing submarkets, vacancy chains are generally quite short, with 90% ending within three migration rounds.[8] We also show that the majority of vacancies created by a new unit are created within one year of the initial move into that unit, implying that the vacancy chains we construct would not grow substantially longer if followed over a longer period of time.

We find that each unit of new high-income urban multifamily and new low-density suburban single family housing creates an average of .9 vacancies that subsequently become occupied within four years.[9] Because new housing units are typically more expensive, the number of vacancies created in below-median income tracts is much lower: New high-income urban multifamily housing generates about .15 such vacancies and low-density suburban single family homes generate .25 such vacancies. The number of vacancies created in low-income high-density tracts is even lower still, with high-income urban multifamily and low-density suburban single family housing respectively creating only .03 and .015 vacancies in below-median income tracts in the top decile of population density.

While the connectivity between the submarkets for new suburban single family housing and for housing in low-income high-density tracts is especially low, the large increase in the supply of suburban homes means that new suburban construction has generated more vacancies in low-income high-density tracts than has new high-income urban multifamily construction. The 1.2 million new low-density suburban single family units identified in our data collectively created about 17 thousand vacancies in low-

[8] Existing work on residential vacancy chains by Mast (2021) and Bratu, Harjunen, and Saarimaa (2023) similarly documents the connections between disparate submarkets created by residential mobility. The fact that vacancy chains end quickly is a new fact that we are able to establish because of the comprehensive data we use.

[9] To avoid counting demolished units and units that are unavailable for occupancy because of renovation, we only count vacated units that subsequently become occupied.

income high density tracts, compared to the 11 thousand such vacancies created by the 356 thousand new high-income urban multifamily units identified in our data.

These descriptive results are suggestive but cannot tell us about to the effects of new suburban and urban housing construction on prices and the welfare of residents without additional structure. In the second part of this paper, we conduct a simulation exercise that connects the observed characteristics of vacancy chains to the unobserved price and welfare consequences of new housing construction. We conduct these simulations using a model and preference parameters taken from Bayer, Ferreira, and McMillan (2007) and data drawn from the IPUMS 1990 5% sample[10].

The simulation exercise is conceptually simple: We first simulate an initial equilibrium set of prices and matches between households and housing units; then, iterating many times, we add a small number of new housing units to a randomly chosen neighborhood and simulate the new equilibrium prices and matches. The difference between the initial equilibrium and the new equilibrium implies a set of vacancy chains, price effects, and welfare effects, which we analyze to understand what vacancy chains can tell us about the price and welfare effects of new housing. This exercise is the second main contribution of our paper.

We show that, despite using a relatively small sample of data for our simulation, our initial equilibrium prices and assignment of households to housing units replicate the key stylized patterns found in the underlying data. We then simulate the effect of a 5% increase in housing supply in one neighborhood at a time, running 1,000 simulations in total. The simulated welfare and price effects of new housing are economically meaningful – the average elasticity of the returns to living in the city with respect to an increase in supply is 1, and the average elasticity of the urban rent premium with respect to supply is -0.3.

Underlying these average effects is considerable variation in the impact of new housing supply. We show that the number of vacancies created in a neighborhood by vacancy chains initiated by new housing is strongly correlated with this variation. We then compare the predictive power of these vacancies with the direct and indirect cross-neighborhood substitution effects implied by the model underlying our simulation. A key result is that the observed number of vacancies is as predictive of variation in the price effects of new housing as are the model-derived substitution effects. The fact that vacancy chains are relatively easy to observe, whereas the model-derived

[10]We use the IPUMS 1990 5% sample rather than the Census Bureau microdata used for our descriptive analysis because it more closely corresponds to the data used to estimate the preference parameters in Bayer, Ferreira, and McMillan (2007).

substitution effects require the estimation of a large number of own- and cross-price demand parameters, makes them especially useful for predicting the non-local price effects of new housing.

Related Literature. This paper builds on and contributes to several strands of the extensive literature on housing supply and affordability. Our work is most closely related to a small literature on vacancy chains that goes back to White (1971). The data required to observe vacancy chains means that earlier work was mostly theoretical or relied on statistical modeling of vacancy chains (Marullo 1985; Turner 2008). The availability of more detailed data on residential histories has only recently made it possible to construct vacancy chains, as in recent work by Mast (2021) and Bratu, Harjunen, and Saarimaa (2023) who study residential vacancy chains in the U.S. and Finland, respectively. Our data allow us to contribute to this literature by documenting new facts about vacancy chains. In addition, we contribute to this literature by providing insights into the economic implications of these descriptive patterns.

Previous work has examined other ways in which housing submarkets interconnect, either through filtering (Rosenthal 2014; Liu, McManus, and Yannopoulos 2022), search (Piazzesi, Schneider, and Stroebel 2019; Landvoigt, Piazzesi, and Schneider 2015), aggregate substitution between housing submarkets (Nathanson 2020), or the hyperlocal effects of new housing construction (Asquith, Mast, and Reed 2021; Damiano and Frenier 2020; Diamond and McQuade 2019; Pennington 2021; Li 2019). Vacancy chains represent an important micro-level mechanism underlying these higher-level mechanisms.

We also contribute to the literature on the city-wide effects of increases in housing supply (e.g., Anenberg and Kung (2020); Molloy, Nathanson, and Paciorek (2022)). Our approach allows us to contribute to this literature by going beyond city-wide averages to better understand the geographic and sociodemographic incidence of the benefits of new housing.

Finally, this paper relates to extensive literatures on supply constraints (Song 2021; Gyourko, Saiz, and Summers 2008; Saiz 2010; Baum-Snow and Han 2021; Kulka, Sood, and Chiumenti 2022) and urban housing affordability (Couture et al. 2021; Couture and Handbury 2020; Su 2021; Handbury 2021). The consequences of urban supply constraints on urban housing costs depend crucially on residential mobility across submarkets. This, in turn, has significant implications for housing policy. If households are able to move easily between submarkets, then policies that increase supply in areas with few constraints may be effective. In contrast, limited mobility between submarkets would recommend policies that relax existing constraints. We contribute to these literatures by documenting the extent to which increases in the supply of housing in one submarket ripple across other submarkets.

The rest of this paper proceeds as follows: Section 2 presents a simple model of differentiated housing demand to fix ideas. Section 3 describes our data and the construction of vacancy chains. Section 4 presents new descriptive facts about vacancy chains initiated by new housing construction in the U.S. Section 5 describes our simulation exercise and results. Section 6 concludes.

2. A Simple Model

In this section, we develop a simple static model of residential demand for housing in different submarkets. Despite the model's simplicity, we are able to express the effect of an increase in the supply of housing in one submarket on housing prices in other submarkets as a function of direct and indirect effects, with the indirect effects consisting of a chain of residential substitution effects mediated by intermediary submarkets. This chain of substitution effects is naturally interpreted as a residential vacancy chain, thus motivating the descriptive analysis of residential vacancy chains in Section 4.

Setup. We model the housing market of a single city as being made up of housing submarkets $n \in \mathcal{N} = \{1, \ldots, N\}$. Housing supply in each submarket is perfectly inelastic, with supply in submarket n denoted S_n. We model demand for housing in submarket n with the reduced form $D_n(\mathbf{p})$, where $\mathbf{p} = (p_1, \ldots, p_N)$ is the vector of prices for housing in all submarkets. We assume that $\frac{\partial D_n(\mathbf{p})}{\partial p_n} < 0$ and $\frac{\partial D_n(\mathbf{p})}{\partial p_m} \geq 0$ for $m \neq n$. We denote the own-price effect on demand for submarket n by $\epsilon_n \equiv -\frac{\partial D_n}{\partial p_n}$; and the cross-price effect on demand for submarket n with respect to prices in $m \neq n$ by $\gamma_{nm} \equiv \frac{\partial D_n}{\partial p_m}$. Prices in equilibrium adjust to clear each submarket.

In addition, we denote the "residential diversion ratio" of m to n by $\lambda_{nm} \equiv \frac{\gamma_{nm}}{\epsilon_m}$. This measure captures the share of the change in demand for housing in submarket m that comes from substitution to or from submarket n when price p_m changes. In the event of a *decrease* in p_m, λ_{nm} represents the share of *increased* demand for housing in m due to substitution *away* from n. When aggregate residential demand is microfounded on household preferences that feature unit housing demand, this interpretation of λ_{nm} is further refined to the share of households that move to submarket m that come from submarket n.

Cross-market effects. We are interested in how an increase in supply in one submarket affects prices in other submarkets. Without loss of generality, we consider the impact of an exogenous increase in supply in submarket 1 on prices in submarket 2.

For clarity of exposition, first consider the cross-market price effects when $N = 3$. Differentiating the market clearing conditions and solving the model yields the following:

$$\frac{dp_2}{dS_1} \propto -\epsilon_2^{-1}\left(\underbrace{\lambda_{21}}_{\textit{direct effect}} + \underbrace{\lambda_{23}\lambda_{31}}_{\textit{indirect effect}}\right).$$

This expression shows that the cross-market price effect depends on a *direct effect*, which is proportional to the residential diversion ratio of 2 to 1, and an *indirect effect*, which is proportional to the product of the residential diversion ratios of 3 to where 1 and of 2 to 3. The interpretation of this indirect effect is as follows: The increased supply in 1 leads households to substitute housing demand away from 3 towards 1; this, in turn, leads households to shift housing demand away from 2 towards 3. The natural interpretation of this chain of substitution effects is a residential vacancy chain.

The fact that this cross-market price effect depends on a direct effect and on indirect effects extends to the general case with N neighborhoods. We demonstrate this by first deriving a general expression for the cross-price effects that takes a recursive form.

PROPOSITION 1. *Given a set of neighborhoods $\mathcal{N} = \{1, \ldots, N\}$, the price-effects of an increase in S_1 take the form:*

$$\frac{dp_i}{dS_1} = -\epsilon_i^{-1}\frac{\Phi_i(\mathcal{N})}{\sum_{j\in\mathcal{N}}\lambda_{1j}\Phi_j(\mathcal{N})}, \tag{1}$$

where $\Phi_i(\mathcal{N})$ is defined recursively:

$$\Phi_{i\neq 1}(\mathcal{N}) = \sum_{j\in\mathcal{N}\setminus i}\lambda_{ij}\Phi_j(\mathcal{N}\setminus i) \tag{2}$$

$$\Phi_1(\mathcal{N}) = -\sum_{j\in\mathcal{N}\setminus 1}\lambda_{Nj}\Phi_j(\mathcal{N}\setminus 1). \tag{3}$$

The recursive form of this expression shows that the cross-market price effects of an increase in S_1 on submarket i depend on the cross-market effects that would exist in submarkets $j \neq i$ if submarket i were removed from $\mathcal{N}$. These cross-market effects are then transmitted to i based on the residential diversion ratio of i for j.

Given that we are able to express the cross-market effects in the case when $N = 3$ as a function of a direct and indirect effect, and given the general expression in the proposition above, it follows from inspection that the price effects consist of a direct effect and a set of indirect effects for any $N \geq 3$. Further, the indirect effects themselves consist of direct and indirect effects, resulting in indirect effects running between submarkets 1 and 2 that are mediated by up to $N - 2$ other neighborhoods. While the simple setup of this model doesn't allow us to make explicit predictions about vacancy chains, it illustrates the intuition for how an increase in the supply of housing in one submarket impacts prices in another submarket through multiple chains of residential substitution between neighborhoods. This motivates our descriptive analysis of vacancy chains in Section 4.

3. Data

3.1. Data sources

Administrative data from the Master Address File. The primary data sources we use to construct vacancy chains are derived from the Census Bureau's Master Address File (MAF). The MAF database is an inventory of all known living quarters in the United States and was created for the 2000 Census. It is updated semi-annually from the US Postal Services delivery sequence file. Additional updates occur through partnerships with local and state governments, address canvassing activities, and other sources. The MAF defines the base sampling frame for the American Community Survey, decennial censuses, and other Census Bureau data products.

We use the 2022 MAF Extract (MAF-X), a snapshot of the MAF database in which we observe the inventory of US housing units in 2022. Housing units in the MAF-X are assigned a MAFID, a unique identifier that can be used to link records across data sources. In addition, we observe housing units' addresses and geographic locations.

In addition, we use the 2000-2021 MAF Auxiliary Reference Files (MAF-ARF) to construct residential histories of the US population at an annual frequency (Genadek and Sullivan). The MAF-ARF is derived from several underlying data sources, including individual tax returns, Selective Service registration data, and Medicare enrollment. Each year of the MAF-ARF is at the individual-level and consists of an individual-level unique identifier and an associated MAFID.

Two key housing unit characteristics for our analysis are the age of the unit and the number of units in the unit's building. While we do not directly observe these characteristics, we are able to impute them from the MAF-X and MAF-ARF.

To impute the year in which a unit first becomes occupied, we use the first year in which a MAFID is associated with a PIK in the MAF-ARF as that unit's first year of occupancy. We construct vacancy chains initiated by new housing in each year from 2009 to 2018. Because our residential histories extend back to 2000, units identified as being new had no associated PIKs for at least nine years before first appearing in the MAF-ARF.

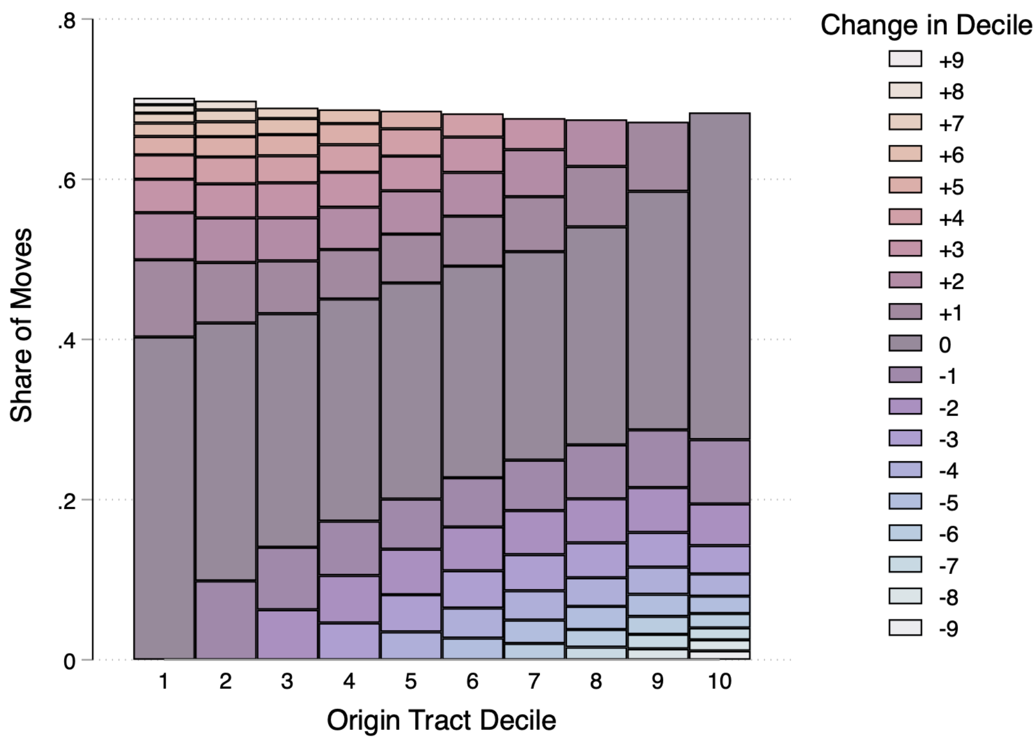
To impute the number of units in a unit's building, we simply take the number of valid MAFIDs in the same census tract with the same street address. In doing so, we exclude units that are indicated to be trailers or mobile homes, or have an address indicating that units are located on separate lots.

Additional data sources. We combine these administrative data sources with data from the American Community Survey (ACS) covering 2005-2018. We use the ACS to measure the tract-level characteristics with which we characterize the vacancies created by new housing units and the types of new housing constructed.

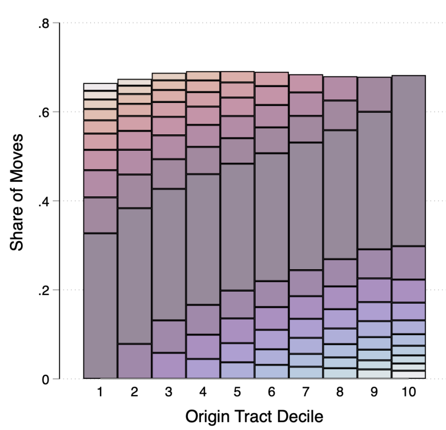
3.2. Moves across submarkets

Connections between housing submarkets. Figure 1 illustrates connections between housing submarkets within Core-Based Statistical Areas (CBSAs) with a 2010 population of three million or greater between 2010 and 2017.

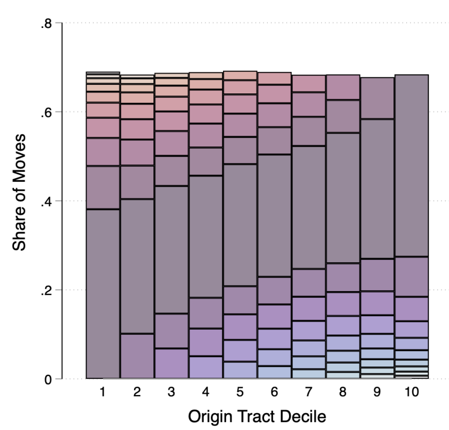
Figure 1A shows the share of individuals aged 25 and older moving from a tract at a given within-CBSA decile of household income that increase or decrease their tract's decile by a given amount. Individuals who leave the CBSA account for the difference between the total share of movers who change their tract's decile and 1. The figure shows that, while there is some stickiness in the kinds of tracts that individuals move to, tracts at different deciles of household income are still connected by a substantial number of moves. For example, while 41% of all moves out of top-decile tracts are also to top-decile tracts, about 27% of moves are to lower-decile tracts within the same CBSA. In addition, the figure illustrates how vacancy chains can connect submarkets even when there are few direct moves between them. For example, while there are very few direct moves between the bottom- and top-decile tracts, these two submarkets are connected indirectly through chains of moves from bottom-decile tracts that pass



A. Median Household Income



B. Median Two-Bedroom Rent



C. Percent College-Educated

FIGURE 1. Moves between tracts within CBSAs

Note: This figure shows the distribution of changes in tract characteristics conditional on origin tract characteristics for individuals aged twenty-five and older who moved between 2010 and 2017 and originated from a tract in a CBSA with a 2010 population of three million or more. In each panel, each column of stacked bars represents movers out of an origin tract at the given within-CBSA decile of the indicated characteristic. The size of each bar indicates the share of moves out of the given origin-tract decile that result in the change of tract decile indicated in the legend of panel A. The stacked bars sum to less than one because of moves out of the CBSA. We calculate tract characteristics using the 2013-2017 ACS.

through tracts at other deciles and end at top-decile tracts. An example of one such path is the series of moves from bottom-decile tracts to top-decile tracts that increase the mover's tract decile by one.

Panels B and C of Figure 1 illustrate similar patterns of migratory flows across submarkets defined by median rents and college share.

3.3. Constructing vacancy chains

New housing gives rise to vacancy chains by initiating a series of moves in which households move into newly available units and vacate their origin units. While the idea is simple, some complications arise when considering how to construct a vacancy chain and connect moves over time. First, not all moves leave a unit vacant, although they may leave a room within the unit vacant. For example, if a roommate moves out, this may initiate a series of moves even though the unit was not vacated. For our analysis, we simplify things by only considering moves that vacate the origin unit. If a unit is not vacated, the chain ends.

Second, units may sit vacant for some time before becoming occupied. When constructing and describing vacancy chains, we would like to describe not only the composition of neighborhoods and movers that are part of the chain but also how the chain evolves over time. To do so, we construct vacancy chains over different time horizons. Figure 2 illustrates how we construct a hypothetical vacancy chain initiated by a new housing unit that was first occupied in 2010. To construct the first link of the chain, we identify the occupants of the new housing unit and trace them back to their 2009 origin unit or units. For origin units that were vacated in the first link of the chain, we search for new occupants in 2010 and trace them back to their 2009 origin unit or units, constructing the second link of the chain.

We continue in this fashion until the chain ends, either because no origin unit is found, an origin unit isn't left vacant, or a vacated unit doesn't become occupied within the chosen time horizon. The vacancy chain illustrated in Figure 2 would end after two rounds of moves if we were to construct the chain over a one-year horizon. Over a two-year horizon, we are able to extend the chain by searching for a new occupant of the vacated housing unit in 2011. We then continue building the chain over a one-year horizon, with 2011 as the reference year. This generalizes for longer horizons straightforwardly.

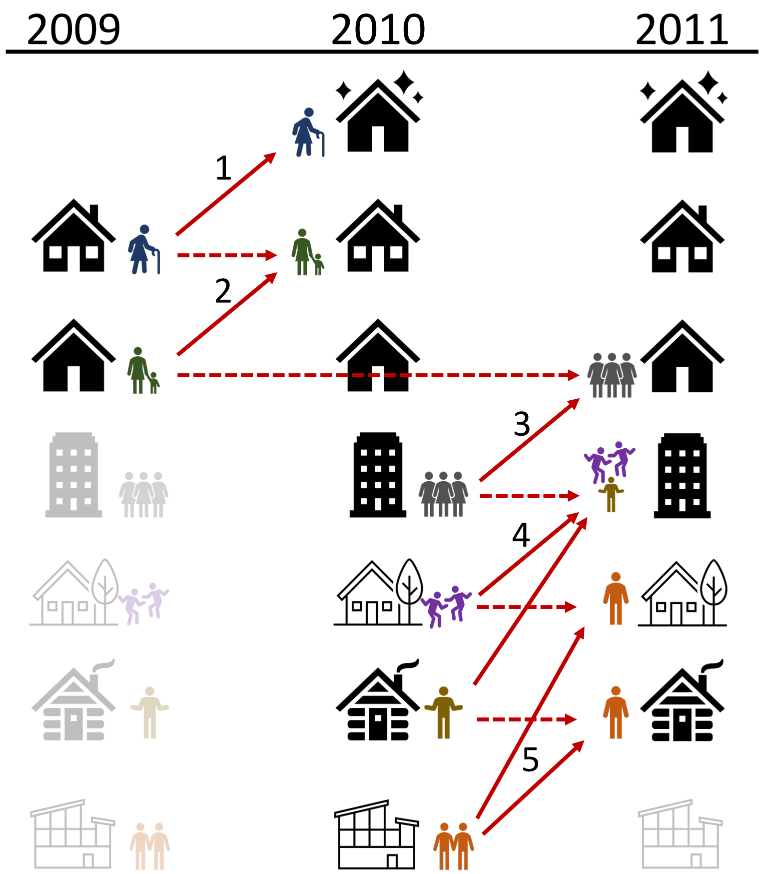


FIGURE 2. Vacancy Chain Construction

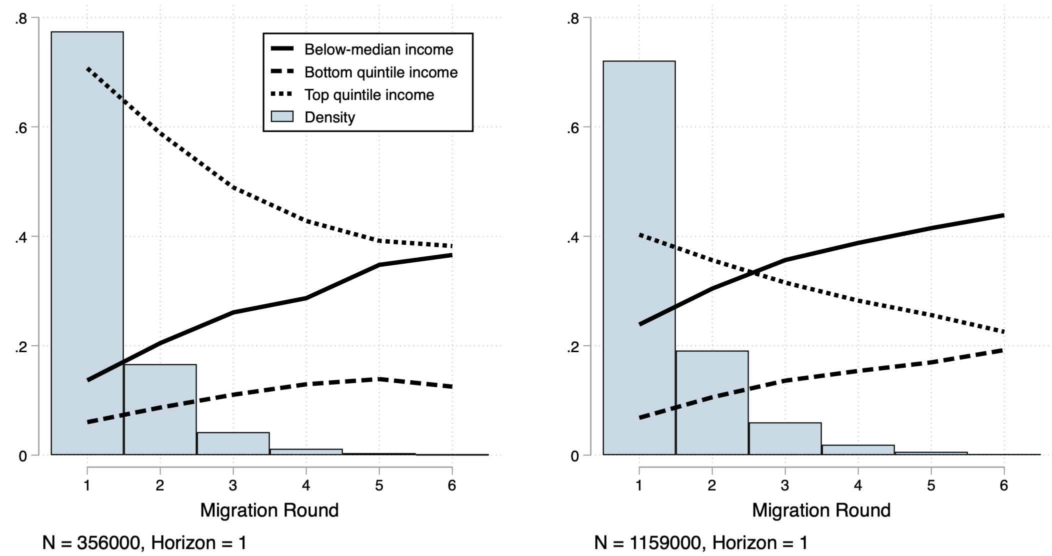
Describing vacancy chains. To characterize vacancy chains, we consider the number of effective vacancies created in different kinds of neighborhoods. Effective vacancies are calculated as a weighted sum, where the weights are inversely proportional to the number of distinct chains connected to a given vacancy. For example, if a unit is vacated by a move in which some household members move to a unit that is part of one vacancy chain and the remaining household members move to a unit that is part of a separate chain, we attribute half of the resulting vacancy to each chain. In addition, we assign a weight of zero to units that are vacated but are not observed to be filled within the time horizon under consideration. This is to avoid counting vacated units that are demolished or are unavailable for occupancy due to renovation or redevelopment.

We focus on vacancy chains initiated by two types of new housing in CBSAs with populations greater than three million: low-density suburban single family homes and high-income urban multifamily buildings. The first type consists of single family homes located in below-median density tracts outside the metropolitan area's principal city. New construction in these tracts accounts for 80% of the increase in housing stock in the US since 1980 (Baum-Snow 2023). The second type of new construction we consider consists of units in 20+ unit buildings located in above-median income tracts within five miles of the metropolitan area's central business district, which corresponds to the type of new housing studied in the existing literature on vacancy chains (Mast (2021); Bratu, Harjunen, and Saarimaa (2023)). Our primary analysis sample consists of 1,159,000 initiated by low-density suburban single family homes and 356,000 vacancy chains initiated by high-income urban multifamily units.

4. Descriptive Facts

We now turn to our descriptive characterization of vacancy chains. We begin by considering how the composition of vacated units changes as the vacancy chain grows longer. Figure 3 shows the share of effective vacancies created in tracts with a given characteristic, conditional on the migration round and over a one-year horizon. Panel A shows these shares for vacancy chains initiated by high-income urban multifamily housing. The first round of moves into these new high-income units create vacancies in predominantly high-income tracts, with 71% of vacancies created in top quintile income tracts, 14% in below-median tracts, and only 8% in bottom quintile tracts. This is unsurprising, given that these kinds of new units are typically very expensive.

As the chain progresses, however, the share of vacancies created in high-income



A. High-Income Urban Multifamily B. Low-Density Suburban Single Family

FIGURE 3. Origin Tract Incomes by Migration Round

Note: This figure shows the share of vacated units in each migration round located in a tract with a given characteristic. In each panel, migration round 1 represents vacancies created by the set of moves into the new units that initiated the chain; migration round 2 represents the vacancies created by the set of moves into the vacancies created in round 1; and so on. Chains are constructed over a one-year horizon. Tract characteristics for vacancy chains initiated in year t are taken from the 5-year ACS covering years $t-4$ to t and quantiles correspond to the national distribution. Income is median household income per capita.

tracts in each round declines and the share created in low-income tracts rises. By the sixth round of moves, 38% of vacancies are created in top quintile income tracts and 37% of vacancies are created in below-median income tracts.

Panel B of Figure 3 displays similar trends as the vacancy chains initiated by new single family homes in low-density suburban tracts grow longer. The share of vacancies in top quintile tracts declines from 40% in the first round of moves to 22% in the sixth round while the share of vacancies created in below-median income tracts increases from 24% in the first round to 44% in the sixth. A notable difference from the chains initiated by high-income urban multifamily housing is that the vacancies created by the initial set of moves is concentrated in tracts with lower median incomes. Again, this is unsurprising, since housing costs for new single family homes in low-density suburbs are typically lower than those for high-income urban multifamily units.

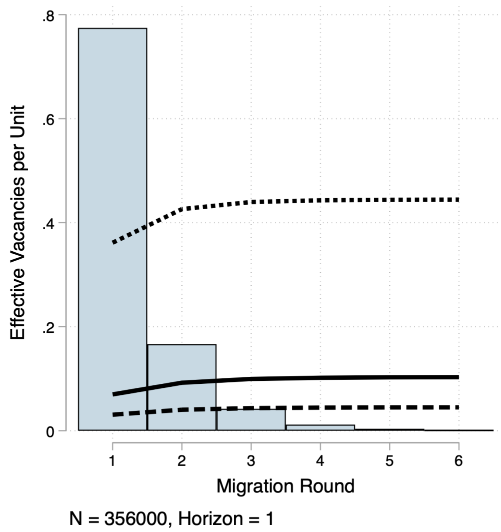
Overall, Figure 3 illustrates how different housing submarkets are connected by residential mobility. The fact that new housing units create vacancies in lower income tracts suggests that building new housing – even in high-income neighborhoods – can loosen demand for housing in lower segments of the market and lower housing costs.

Taken in isolation, this fact might suggest that a viable strategy for lowering housing costs for low-income renters is to build more housing of any kind. However, another salient fact illustrated in Figure 3 is that vacancy chains are relatively short. In both panels, the density of the total number of vacancies is shown by the blue histogram. Panel A shows that 77% of all vacancies created by high-income urban multifamily housing are created in the first round of moves, and each round of moves creates about 25% as many vacancies as the previous round. In panel B, a qualitatively similar pattern holds. An important implication of this pattern is that the location and type of new housing construction matters. Even though increases in the supply of new high-end housing can loosen demand in lower-end segments if the resulting vacancy chains go on long enough, vacancy chains are very unlikely to continue for many rounds and these supply increases are unlikely to have meaningful effects on costs.

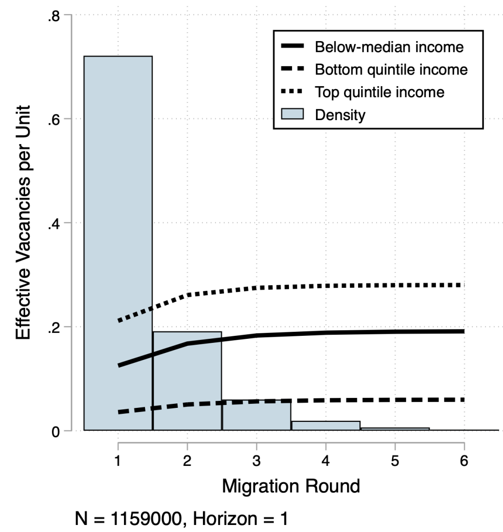
Given the short length of vacancy chains, we turn our focus to the cumulative number of effective vacancies created by vacancy chains. Figure 4 shows the cumulative number of effective vacancies created in each round of moves. Panels A and B show vacancies created over a one-year time horizon, with panel A showing vacancies created by high-income urban multifamily housing and panel B showing vacancies created by low-density suburban single family homes. In both panels, the number of vacancies quickly plateaus as the chain grows longer due to the relatively low probability of a chain advancing from one round to the next. This suggests that there would be very few additional vacancies created in migration rounds beyond the sixth round, and the number of vacancies created by the sixth round is a close approximation of the total number of vacancies created by each type of new construction.

Panels C and D show cumulative vacancies created over a four-year horizon, with panel C showing vacancies created by high-income urban multifamily housing and panel D showing vacancies created by low-density suburban single family homes. While chains constructed over this longer time horizon tend to be longer, the distribution of vacancies is still strongly skewed towards earlier migration rounds such that the cumulative number of vacancies levels off by the sixth round.

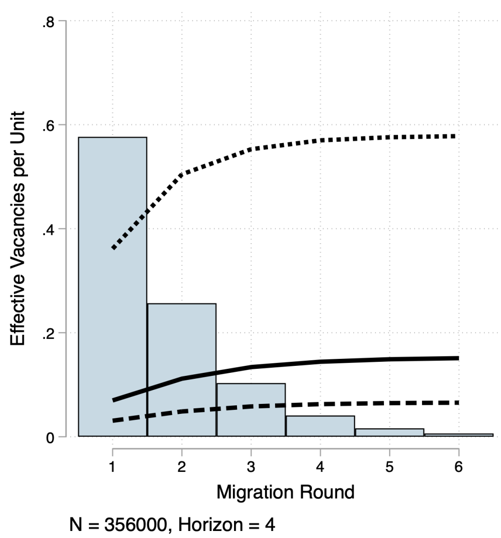
Figure 5 shows how the cumulative number of effective vacancies created by both kinds of new construction changes over time. Panels A and B demonstrate that both types of housing produce more vacancies in high-income neighborhoods than in low-income neighborhoods, though high-income urban multifamily housing produces substantially more high-income vacancies than does low-density suburban single family



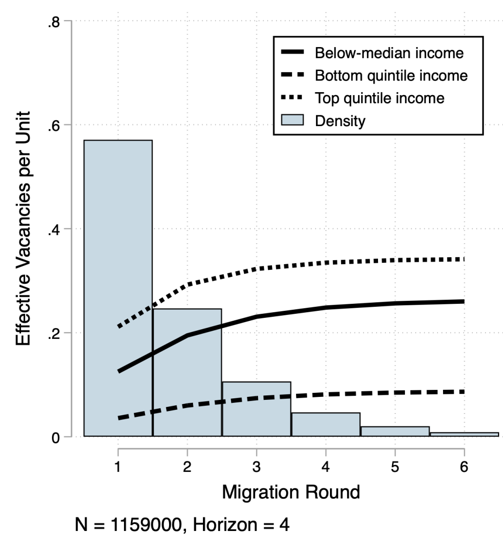
A. High-Income Urban Multifamily



B. Low-Density Suburban Single Family



C. High-Income Urban Multifamily



D. Low-Density Suburban Single Family

FIGURE 4. Cumulative Vacancies by Migration Round

Note: This figure shows the cumulative number of effective vacancies created in each migration round located in a tract with a given characteristic. In each panel, migration round 1 represents vacancies created by the set of moves into the new units that initiated the chain; migration round 2 represents the vacancies created by the set of moves into the vacancies created in round 1; and so on. Chains are constructed over a one-year horizon. Tract characteristics for vacancy chains initiated in year t are taken from the 5-year ACS covering years $t-4$ to t and quantiles correspond to the national distribution. Income is median household income per capita.

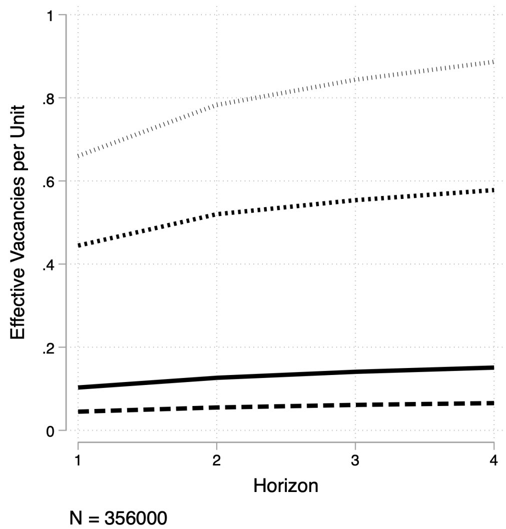
housing. In particular, panel A shows that high-income urban multifamily housing produces .44 and .58 vacancies in top quintile income tracts over a one- and four-year horizon, respectively. This represents about two thirds of the total number of vacancies created by this type of housing. In contrast, high-income urban multifamily housing produces only .1 and .15 vacancies in below-median income tracts over a one- and four-year horizon, respectively, which represents about 15% of the total number of vacancies created.

Panel B shows that low-density suburban housing produces .28 and .34 vacancies in top quintile income tracts over a one- and four-year horizon, respectively, which represents about 35% of the total number of vacancies created. New low-density suburban housing creates a comparable number of vacancies in below-median income tracts: .19 and .26 vacancies over one- and four-year horizons, representing about 25% of the total number of vacancies created.

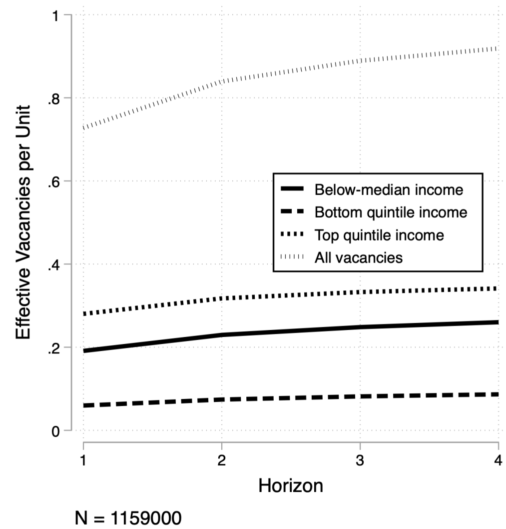
Both panels A and B show that the majority of vacancies created by new housing construction of either type are created within a one-year horizon. In addition, the number of new vacancies created in each year is diminishing. This pattern suggests that few additional vacancies are created over time horizons beyond four years and we are therefore capturing the majority of vacancies created by new housing construction.

Panels C and D of Figure 5 focus on how both types of new housing construction connect to the low-income high-density submarkets where the households most exposed to rising housing costs are most likely to live. Both panels show the number of vacancies created over time in low-income tracts that are either high-density or very high-density. We define low-income tracts as those in the bottom quintile of the national distribution of income; high-density tracts as those in the 19th ventile of the national distribution of population density; and very high-density tracts as those in the top ventile of the national distribution of population density.

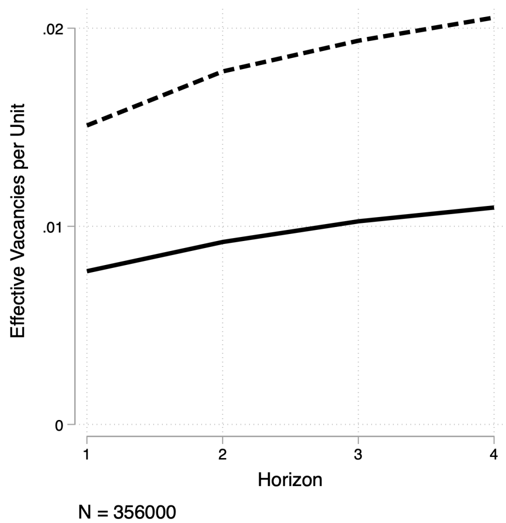
Panel C shows the number of effective vacancies created in low-income and high-density tracts by high-income urban multifamily housing. In general, the number of vacancies created is very low – over a four-year horizon, it takes 50 new high-income urban multifamily units to generate one vacancy in a low-income very high-density tract and 100 new units to generate a vacancy in a low-income high-density tract. Panel D shows that the number of effective vacancies created in these submarkets by new low-income suburban single family housing is even lower, requiring more than 100 new units to generate a vacancy in either high- or very high-density low income tracts over a four-year horizon.



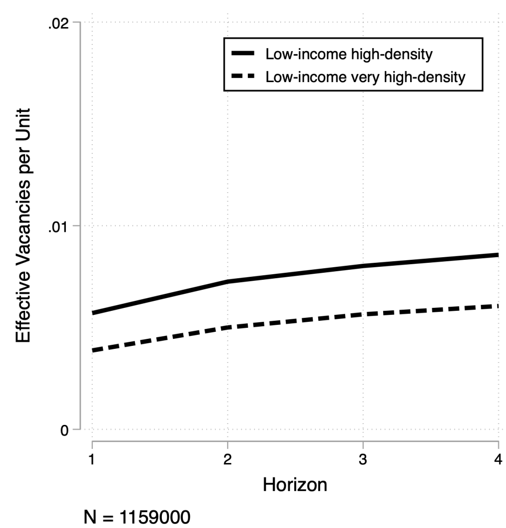
A. High-Income Urban Multifamily



B. Low-Density Suburban Single Family



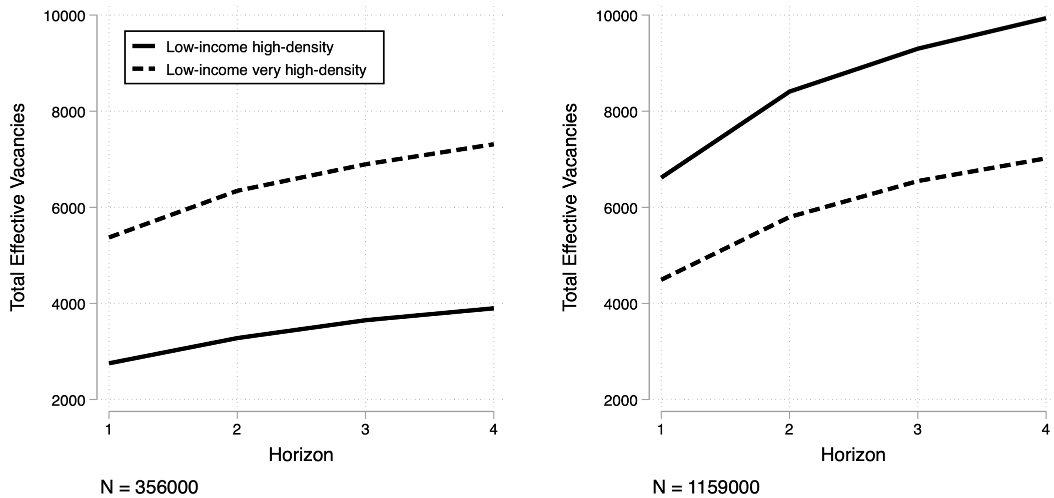
C. High-Income Urban Multifamily



D. Low-Density Suburban Single Family

FIGURE 5. Cumulative Vacancies per Unit over Time

Note: This figure shows the number of effective vacancies created over time in tracts with a given characteristic per unit of new housing. In each panel, migration round 1 represents vacancies created by the set of moves into the new units that initiated the chain; migration round 2 represents the vacancies created by the set of moves into the vacancies created in round 1; and so on. Each point represents the number of effective vacancies created over six rounds of moves over the indicated time horizon. Tract characteristics for vacancy chains initiated in year t are taken from the 5-year ACS covering years $t - 4$ to t and quantiles correspond to the national distribution. Income is median household income per capita. Low-income tracts are those in the bottom quintile of the income distribution, high-density tracts are in the 19th vingtile of the distribution of population density, and very high-density tracts are those in the top vingtile of the distribution.



A. High-Income Urban Multifamily B. Low-Density Suburban Single Family

FIGURE 6. Total Vacancies over Time

Note: This figure shows the total number of effective vacancies created over time in tracts with a given characteristic. In each panel, migration round 1 represents vacancies created by the set of moves into the new units that initiated the chain; migration round 2 represents the vacancies created by the set of moves into the vacancies created in round 1; and so on. Each point represents the number of effective vacancies created over six rounds of moves over the indicated time horizon. Tract characteristics for vacancy chains initiated in year t are taken from the 5-year ACS covering years $t-4$ to t and quantiles correspond to the national distribution. Income is median household income per capita. Low-income tracts are those in the bottom quintile of the income distribution, high-density tracts are in the 19th vingtile of the distribution of population density, and very high-density tracts are those in the top vingtile of the distribution.

While each unit of high-income urban multifamily housing generates more vacancies in low-income urban neighborhoods than a new unit of low-density suburban single family housing, Figure 6 shows that low-density suburban single family homes have created more total vacancies in low-income high-density tracts and a comparable number of vacancies in low-income very high-density tracts. This is due to the much larger number of new suburban housing units constructed between 2008 and 2018.

5. The Economic Content of Vacancy Chains

We have presented a detailed picture of how vacancy chains vary according to the type of new housing that is built and the location of construction. In this section, we conduct a simulation exercise that connects the observed characteristics of vacancy chains to unobserved price and welfare effects.

The simulation exercise is conceptually simple: We first simulate an initial equilibrium matching of households to housing units and a vector of prices that sustains that matching; then, iterating many times, we add a small number of new housing units to a randomly chosen neighborhood and simulate the new equilibrium prices and matching. The difference between the initial equilibrium and the new equilibrium implies a set of vacancy chains, price effects, and welfare effects, which we analyze to understand what vacancy chains can tell us about the price and welfare effects of new housing.

We use a modified version of the model estimated by Bayer, Ferreira, and McMillan (2007) and calibrate it using their parameter estimates. We sample neighborhoods, housing units, and households from the 1990 IPUMS 5% sample. Given these preferences and the sampled neighborhoods, housing units, and households, we apply a tatonnement algorithm to find an initial equilibrium consisting of a market-clearing set of prices and the corresponding matching of households to housing units. We then repeatedly draw new units and add them to the housing stock in the simulated data, recomputing the equilibrium prices and matching of households to units in each iteration. We show how the resulting simulated vacancy chain characteristics correlate with the characteristics of the new unit types and locations, as well as the price and welfare effects.

5.1. Model

The following is a modified version of the residential choice model presented in Bayer, Ferreira, and McMillan (2007), modified to make it easily replicated using the IPUMs data. There is a finite number of households indexed by i, housing units indexed by h, and neighborhoods indexed by n. The set of available housing units consists not only of units in the city, but also units in an outside option neighborhood oo.

Households choose a unit to live in to solve

$$\max_h V^i_{hn} = \underbrace{\alpha^i_X \mathbf{X}_h + \alpha^i_Z \mathbf{Z}_n - \alpha^i_p p_h + \xi^i_h}_{\equiv v^i_{hn}} + \epsilon^i_{nh},$$

where $\mathbf{X}_h$ is a vector of non-price housing unit characteristics that includes housing unit age bin indicators and the number of rooms; $\mathbf{Z}_n$ is a vector of neighborhood characteristics that includes racial and ethnic composition, college-educated share, and average income; p_h is the price of unit h, ξ_n is unobserved neighborhood quality; and ϵ^i_h is household i's idiosyncratic preference for unit h, where ϵ^i_h is drawn from a type-1 extreme value distribution. Units in the outside option are normalized such that $\nu^i_{h,oo} = 0$. In addition, households are indifferent between units in the outside option, so $\epsilon^i_h = \epsilon^i_{oo}$ for all units $h$in the outside option.

Household preferences are permitted to vary with the following observable household characteristics: The presence of children under 18; capital and non-capital income; capital income; race and ethnicity; educational attainment; employment status; and age.

We impute unobserved neighborhood quality ξ^i_n by estimating a hedonic regression and taking the mean residual variation in rents across across housing units within a neighborhood. We assume that this residual variation reflects a willingness to pay to live in n that is common across all households. Because the marginal utility of consumption is permitted to vary across households, this assumption implies that ξ^i_n varies across households.

Parameter estimates are computed from the tables in Bayer, McMillan, and Rueben (2004). This model is particularly well suited to our application because it models residential choices at the housing unit level and the parameters are estimated to maximize the likelihood of observing each household matched with the unit in which it resides. This is in contrast to many residential discrete choice models in which a continuum of households choose over neighborhoods and the parameters are estimated to match the neighborhood choice shares observed in the data.

5.2. Equilibrium and Iteration

We now describe the tatonnement algorithm we use to compute equilibrium prices for a given set of households and housing units. This is an implementation of the Hungarian algorithm (Demange, Gale, and Sotomayor, 1986; Easley and Klineberg, 2010).

We begin by setting all prices equal to zero. In each iteration of the algorithm, we find each household's utility-maximizing set of housing units given the current vector of prices, which we refer to as their preferred units. If there is a perfect matching of households to housing units in which each household matches with one of its preferred units, we have found an equilibrium. If there is no such perfect matching, then there must exist a *constricted set* of units $\mathcal{S}$ – a set of units such that: (a) the households that prefer units in $\mathcal{S}$ prefer no units outside of $\mathcal{S}$; and (b) there are more households that prefer units in $\mathcal{S}$ than there are units in $\mathcal{S}$. We identify the constricted set and raise prices for all units in $\mathcal{S}$ by one price increment. We then begin the next iteration of the algorithm and continue until a perfect matching is found.

Because the algorithm requires that valuations and prices have discrete support, we normalize the data in several ways. First, we convert all preferences into a willingness to pay by rescaling each household's preference parameters such that the marginal utility of consumption is unity. This implies that the scale of idiosyncratic preferences varies across households.

Second, we rescale the utility achieved with each choice to be integer-valued. We do so by dividing by the price increment used in the algorithm and rounding to the nearest integer.[11]

Finally, we normalize the value of the outside option to be equal to the minimum utility achieved by a household choosing an inside option when prices are equal to 0. We then shift all utilities by a constant such that the utility achieved by a household choosing the outside option is 0. After these normalizations, each household's utility when matched to housing unit h reflects their willingness to pay (in units implied by the price increment) to live in h rather than in the outside option.

Once we have computed an initial equilibrium, we repeatedly simulate the effects of new housing construction. In each iteration, we begin with the initial equilibrium and randomly sample a small number of new housing units and add them to the set of existing housing units in a single neighborhood. We then find the new equilibrium, construct the resulting vacancy chains, and calculate the price and welfare effects of the increase in supply.

One limitation of this exercise is that it does not allow for vacancy chains to end as a result of new household formation or because a unit remains vacant, both of which are important reasons that vacancy chains end in the observed data. For this exercise, vacancy chains can only end because they reach the outside option. Despite this limitation, the simulation exercise provides insights that help us interpret the descriptive patterns described in the previous section.

5.3. Data

Our simulation exercise uses microdata from the 1990 IPUMS 5% sample.[12] While we define neighborhoods at the tract level in our descriptive analysis of vacancy chains,

[11] Rounding to the nearest integer naturally introduces some error into the algorithm. Using smaller price increments leads to lower approximation error but at the expense of computation time.

[12] We use the 1990 IPUMS data to facilitate the use of the parameter estimates from Bayer, Ferreira, and McMillan (2007). One concern with using these data rather than more recent data is that the demographic composition and amenity value of cities have changed substantially since 1990. This might mean that our simulation results do not generalize to the time period used for our descriptive analysis of vacancy chains.

the most granular geographic units we observe in our simulation data are PUMAs. We, therefore, define neighborhoods at the PUMA level for this exercise. While PUMAs are much more populous than tracts, containing at least 100,000 individuals, they are geographically compact in dense metro areas, making them a reasonable proxy for neighborhoods.

The tatonnement algorithm we use to find equilibrium matchings and prices is computationally intensive and fails to converge when using a realistic number of households, housing units, and neighborhoods. Because of this, we use a reduced sample to simulate the effects of increased housing supply. For our primary specification, we construct a bootstrapped sample by randomly sampling 10 PUMAs from the Chicago CBSA with replacement. We then sample 100 housing units and 133 households from each of the sampled PUMAs. In addition, we take all housing units in our sampled PUMAs that are less than one year old as the pool from which we draw new housing units in our simulation.

Because there are more households than housing units in our sample, we augment the sample by adding additional units to represent the outside option. The value of these outside option units is normalized such that households have zero willingness to pay to live in them and are indifferent between all outside option units. When computing an equilibrium, we thus have a perfect matching when every housing unit in the CBSA is matched to a household and the remaining households are matched to units in the outside option.

Table 1 reports mean characteristics of the PUMAs used in the simulation exercise. The main point worth noting is that there is substantial variation in neighborhood characteristics, with PUMAs ranging from very low-income and low college share to high-income high college share. Figure 7 shows the locations of these neighborhoods in the Chicago Metropolitan Area. The PUMAs used in our simulation exercise are also geographically varied, with some located in high-density areas near the city center and others in more distant suburbs.

TABLE 1. Mean Characteristics of Sampled PUMAs

PUMA	ξ_n	Black	Hispanic	College	Household Income	Owner-occupied	Rooms	Age	Employed	Child Present	Built in 80s
1	-12.4	0.11	0.03	0.25	48,628	0.78	6.2	44.9	0.81	0.50	0.20
2	-82.7	0.22	0.04	0.24	45,301	0.77	6.0	48.8	0.71	0.43	0.10
3	-45.3	0.81	0.01	0.25	23,573	0.16	4.2	48.9	0.44	0.34	0.05
4	-17.1	0.00	0.02	0.27	52,834	0.84	6.2	47.8	0.77	0.41	0.36
5	-0.9	0.03	0.04	0.22	47,093	0.75	5.6	49.6	0.72	0.35	0.09
6	34.7	0.03	0.02	0.50	59,717	0.72	6.2	43.0	0.86	0.42	0.30
7	46.6	0.21	0.03	0.17	42,211	0.78	5.7	47.7	0.73	0.43	0.14
8	-50.4	0.49	0.19	0.10	31,548	0.64	5.3	50.7	0.58	0.44	0.02
9	-93.6	0.00	0.02	0.20	41,389	0.75	5.8	47.3	0.74	0.42	0.16
10	-8.9	0.79	0.16	0.04	20,989	0.29	5.2	47.1	0.44	0.55	0.08

Note: This table reports mean characteristics of the PUMAs sampled for our simulation exercise, as reported in the 1990 Census via IPUMS. The column ξ_n reports the mean residual from a hedonic regression of rent on PUMA- and unit-level characteristics. The mean individual-level characteristics Black, Hispanic, College, Age, and Employed are calculated based on the characteristics of the household head. Child Present indicates the mean number of households with a member under 18 years of age. All means are calculated using household weights.

5.4. Results

Initial Equilibrium. We begin by describing the initial equilibrium of our simulation exercise. Reassuringly, we observe that the patterns in this initial equilibrium are similar to those in the underlying data. Figure 8 shows that the simulated housing unit prices in our initial equilibrium are highly correlated with observed housing prices, with an increase in a unit's observed being accompanied on average by a one-for-one increase in that unit's simulated price.

Figure 9 shows that the matching of households to units preserves the sorting patterns observed in the underlying data. Each panel shows the mean of a given characteristic of sampled households conditional on the mean of the PUMA they reside in. The red squares show this relationship for the underlying data, while the blue circles show the relationship between the mean characteristics of sampled households conditional on the mean of the PUMA they match with in our initial equilibrium. The pattern of sorting that results from our simulation is remarkably similar to the pattern observed in the underlying data, with high-income households sorting to high-income PUMAs, college-educated households sorting to PUMAs with a high college share, Black households sorting PUMAs with a higher share of Black households, and Hispanic households sorting to PUMAs with a higher share of Hispanic households.

While the preference parameters we use in our simulation are estimated to match similar data, there are several reasons why it is not ex-ante obvious that our simulation exercise would be able to replicate these features of the underlying data so closely. First, and most importantly, the preference parameters estimated by Bayer, Ferreira, and McMillan (2007) are estimated via maximum-likelihood, taking prices as given and with no structure placed on how households match with housing units. By contrast, we simulate equilibrium prices via a tatonnement algorithm to find a one-to-one matching of households to housing units such that no household wants to switch units.

Second, the preference parameters are estimated using data on households and housing units in San Francisco, while our simulation exercise uses data from Chicago. If there were unobserved heterogeneity in preferences across cities, applying the preferences of San Franciscans to the residential choices facing Chicagoans might have resulted in a simulated equilibrium that failed to match the observed patterns of residential choices.

Finally, we conduct our simulation exercise using only a small subset of the data. The fact that households in our sample have a limited choice set might have resulted in a different pattern of sorting than in the observed data. Overall, the similarity between our simulation results and the observed data gives us more confidence when applying

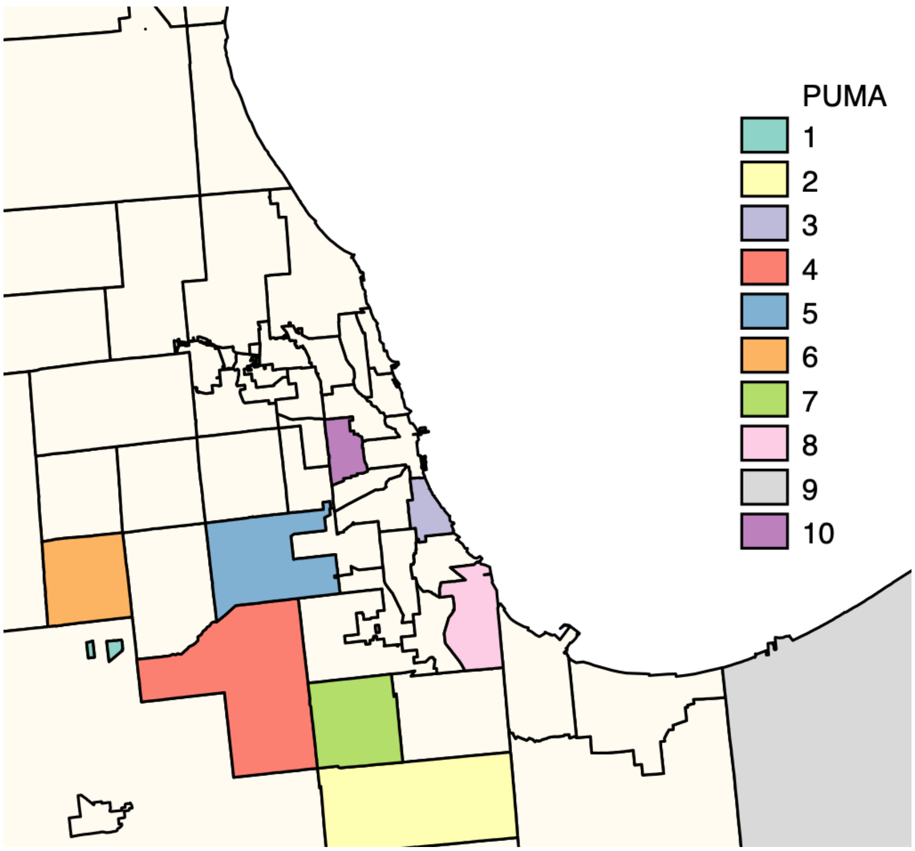


FIGURE 7. Chicago Metropolitan Area

Note: This figure shows the location of the ten 1990 PUMAs used in our simulation exercise. PUMA numbers correspond to those in Table 1.

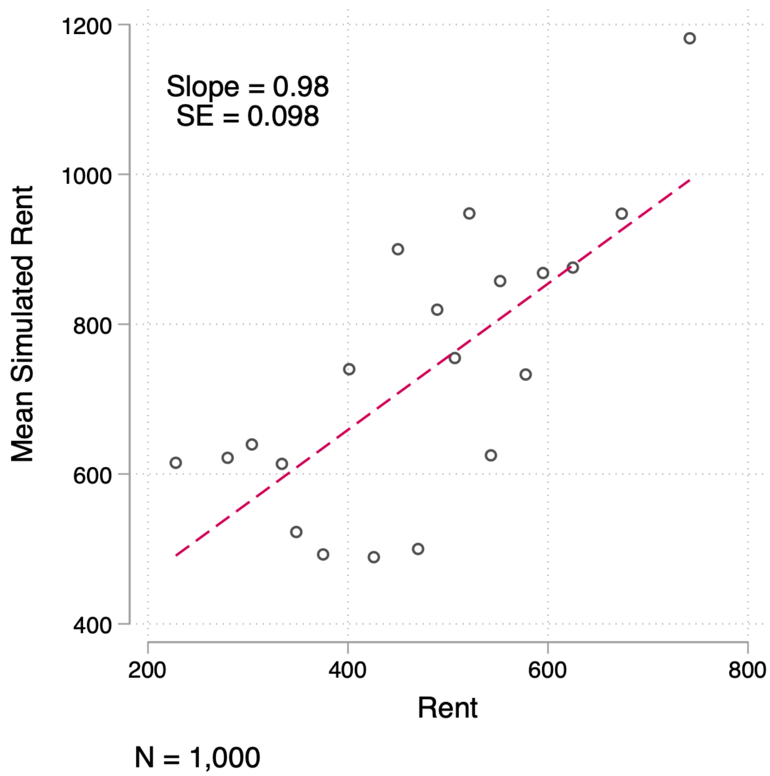
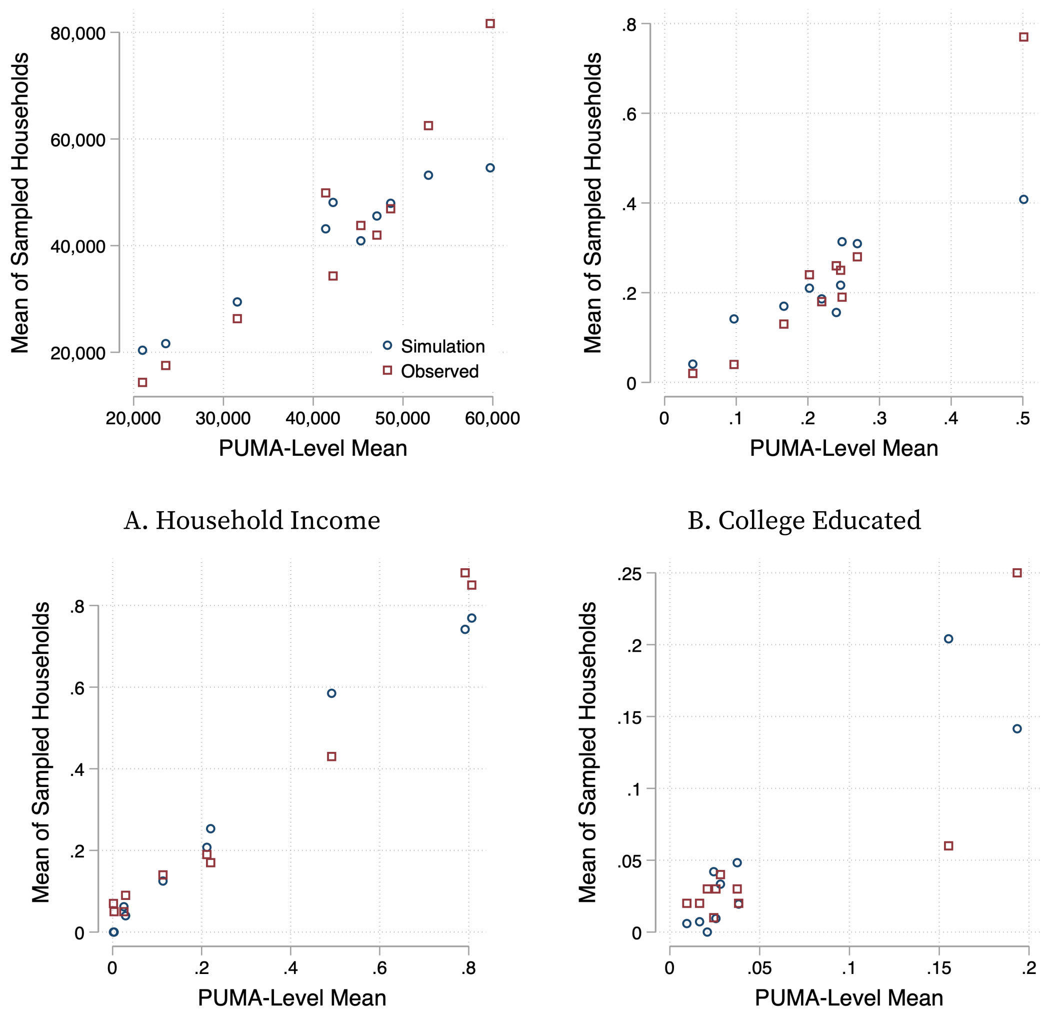


FIGURE 8. Simulated Unit Prices and Observed Rent

Note: This figure shows the mean simulated price of housing units conditional on their observed rent in the IPUMS 1990 5% sample. We estimate a hedonic regression to impute the rent of owner-occupied housing units in the IPUMS sample.

our simulation results to interpret the descriptive facts on vacancy chains presented in the first part of this paper.

Price and Welfare Effects. We now turn to our main objects of interest for this exercise – the simulated price and welfare effects of new housing supply. Figure 10 shows the distribution of these effects generated by 1000 simulations. Panel A shows the distribution of welfare effects as a percent of the initial level of aggregate welfare. The dashed line indicates the mean welfare effect of 0.5%. Given that we normalize the utility of the outside option to be zero and that, in each simulation, we add five housing units to the existing sample of 1,000, this implies an elasticity of the returns to living in the city of 1. Panel B shows the distribution of price effects, which approximately mirrors the distribution of welfare effects. The dashed line corresponds with the mean price effect of -.14%, which implies that the elasticity of the urban rent premium with respect to



A. Household Income

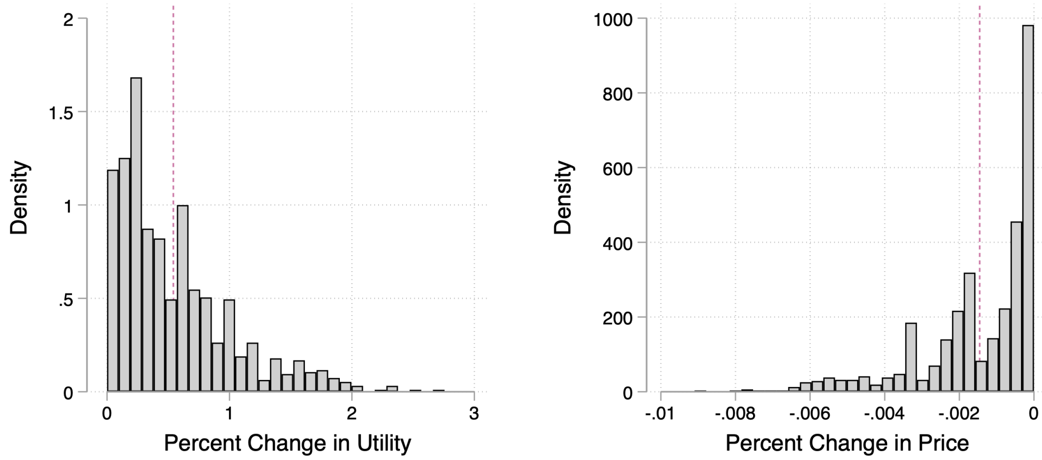
B. College Educated

C. Black

D. Hispanic

FIGURE 9. Simulated and Observed Residential Sorting

Note: This figure shows the mean characteristics of households used in our simulation exercise conditional on the PUMA-level mean of those characteristics in the PUMAs they are assigned to in the initial simulated equilibrium. PUMA-level means are estimated using the IPUMS 1990 5% sample. Each panel represents the 1,000 households in our sample that are matched with a sampled housing unit in the initial equilibrium.



A. Welfare Effects   B. Price Effects

FIGURE 10. Simulated Welfare and Price Effects of New Housing

Note: This figure shows the distribution of welfare and price effects from new housing calculated over 1000 simulations. The dashed red line indicates the mean effect size.

supply is 0.3.

Table 2 shows the incidence of these effects on different types of households. Columns two and four, respectively, show aggregate utility and mean prices for each group in our simulation's initial equilibrium, while columns three and five show the change in aggregate utility and average prices for each group. The welfare effects reported in columns one and two show that households that move and local households (i.e., those residing in the PUMA that receives new housing supply) experience the largest percent increases in welfare – 2.1% and 1.7% respectively. The fact that 43% of the aggregate welfare gains accrue to movers is attributable mostly to better matches. While local households experience relatively large percent increases in welfare, 83% of the aggregate welfare effect accrues non-locally, which suggests that the non-local price effects of new housing supply are economically important.

How does the variation in welfare and price effects documented in Figure 10 and Table 2 correlate with the vacancy chains that result from the addition of new housing? Figure 11 shows the mean simulated price and welfare effects of new housing conditional on the number of vacancies created in a neighborhood. The mean welfare effects are calculated for households that lived in a PUMA in which a given number of vacancies was created, while mean price effects are calculated for housing units in a PUMA with a given number of vacancies. Both price and welfare effects are strongly correlated with

TABLE 2. Price and Welfare Effects

	# of Households	Utility	ΔUtility	Price	ΔPrice
All	1,330	2,612	14.1	547.5	-0.008
			(12.3)		(0.009)
Low-Income Households	665	1,119	4.9	528.1	-0.005
			(5.0)		(0.007)
Low-Income PUMAs	500	993	4.6	655.7	-0.007
			(4.6)		(0.008)
Movers	240	281	6.0	599.8	-0.010
	(10)	(18)	(3.7)	(14.8)	(0.012)
Stayers	1,095	2,331	8.2	536.3	-0.007
	(10)	(18)	(9.4)	(3.4)	(0.009)
Local	100	211	2.4	728.2	-0.018
		(64)	(3.3)	(433.7)	(0.027)
# of Simulations: 1,000					

Note: This table reports means of initial utility and prices and mean welfare and price effects for different PUMAs and households over 1000 simulations. Standard errors are reported in parentheses.

the number of vacancies. Households residing in PUMAs in which no vacancies were created experienced an increase in welfare of only .26%, while those in PUMAs with five or more vacancies experienced an average increase in welfare of 1.1%. Similarly, units in PUMAs with no vacancies saw a fall in prices of less than .001% while those with five or more vacancies saw a more than three-fold greater fall in prices of .0024%.

Demand Substitution and Vacancy Chains. We also examine how our simulated price effects compare to those predicted by the underlying residential choice model. To do so, we compute the individual own- and cross-price partial derivatives of demand implied by the model. Following the notation introduced in the earlier section, we denote the own-price partial derivative of demand for neighborhood i by $\epsilon_i \equiv -\frac{\partial D_i}{\partial p_i}$ and the cross-price partial derivative of demand for neighborhood i with respect to the average price of units in j by $\gamma_{ij} \equiv \frac{\partial D_i}{\partial p_j}$. We denote the residential diversion ratio of j for i by $\lambda_{ji} \equiv \frac{\gamma_{ji}}{\epsilon_i}$.

To better understand what vacancies reveal about price and welfare effects, we estimate a series of regressions in which we regress the simulated change in prices on these substitution terms and the number of simulated vacancies. To make the estimates easier to interpret, we normalize all variables to be mean zero with unit variance. Table 3

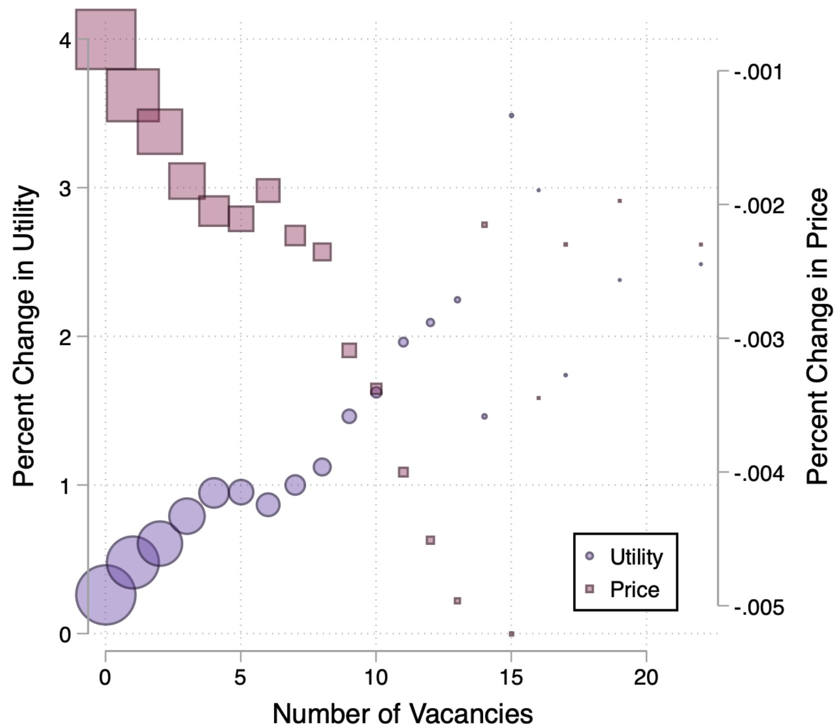


FIGURE 11. Simulated Price and Welfare Effects Conditional on Number of Vacancies

Note: Welfare effects are calculated at the household level and are conditional on the number of vacancies in the household's PUMA of residence in the initial equilibrium. Price effects are calculated at the housing unit level. Point sizes are proportional to the number of observations.

reports these regression estimates. Column 1 reports estimates from a regression of the simulated change in prices in PUMA j on the direct and indirect substitution chains between PUMAs i and j, where i is the PUMA in which new housing was added. We include indirect substitution effects that pass through up to two different PUMAs. We find that both direct and indirect substitution effects are highly significant predictors of the price effects of new housing supply – a one standard deviation increase in direct substitution between i and j leads to a .79 standard deviation increase in the magnitude of the price effect while a one standard deviation increase in indirect substitution mediated by one and two other neighborhoods leads, respectively, to .33 and .5 standard deviation increases in the magnitude of the price effect.

Column 2 adds the number of vacancies created in PUMA j as a regressor. We find that the number of vacancies strongly predicts variation in the simulated price effects, with a one standard deviation increase in vacancies (i.e. an increase of 2.6) predicting a .27 standard deviation increase in the magnitude of the price effect. Column 3 adds

TABLE 3. Regression Estimates

	(1) $\Delta Price_j$	(2) $\Delta Price_j$	(3) $\Delta Price_j$	(4) $\Delta Price_j$	(5) $\Delta Price_j$
$\epsilon_j^{-1}\lambda_{ji}$	-0.785***	-0.513***	-0.384***		
	(0.033)	(0.035)	(0.037)		
$\epsilon_j^{-1}\sum_k \lambda_{jk}\lambda_{ki}$	-0.332***	-0.321***	-0.298***		
	(0.019)	(0.019)	(0.019)		
$\epsilon_j^{-1}\sum_{k,\ell} \lambda_{jk}\lambda_{k\ell}\lambda_{\ell i}$	-0.497***	-0.407***	-0.293***		
	(0.036)	(0.035)	(0.037)		
$Vacancies_j$		-0.272***	0.228***	-0.282***	0.424***
		(0.014)	(0.055)	(0.010)	(0.054)
$Vacancies_j \times \epsilon_j^{-1}$			-0.532***		-0.718***
			(0.057)		(0.054)
Constant	0.107***	0.087***	0.060***	-0.056***	-0.056***
	(0.013)	(0.013)	(0.013)	(0.010)	(0.010)
R^2	0.077	0.110	0.118	0.075	0.091
Observations	10,000	10,000	10,000	10,000	10,000

Note: All variables are transformed to be mean zero with unit variance. $\epsilon_i \equiv -\frac{\partial D_i}{\partial p_i}$ is the own-price partial derivative of demand for neighborhood i; $\gamma_{ij} \equiv \frac{\partial D_i}{\partial p_j}$ is the cross-price partial derivative of demand for neighborhood i with respect to the average price of units in j. $\lambda_{ji} \equiv \frac{\gamma_{ji}}{\epsilon_i}$ is the residential diversion ratio of j for i.

the number of vacancies interacted with the inverse own-price effect for PUMA j. This, too, is a highly significant predictor of variation in price effects and is associated with the largest variation in effect sizes of any regressor.

Columns 4 and 5 consider the predictive power of the number of vacancies alone. Notably, variation in the number of vacancies alone is just as predictive of variation in price effects as the direct and indirect substitution terms in column 1, explaining 7% of the variation in price effects. In column 5, adding the interaction between the number of vacancies and the inverse own-price elasticity of demand for PUMA j explains three-quarters of the variation explained by the full set of regressors in column 3.

Overall, the results reported in Table 3 show that the number of vacancies created by vacancy chains strongly predicts the incidence of price effects generated by new housing supply. While the direct and indirect substitution effects have independent predictive power, these terms are much harder to observe. With just ten PUMAs, forty-five distinct pairwise cross-price terms and ten distinct own-price terms are required to calculate these effects. In a realistic setting with many more PUMAs, the number of parameters to estimate quickly becomes infeasibly large. In contrast, the number of vacancies created in vacancy chains is easily observed, regardless of the number of neighborhoods, and is as predictive of variation in price effects as the substitution parameters.

6. Conclusion

The effect of new housing supply in one submarket on housing costs in other submarkets depends crucially on how residential mobility connects these submarkets. The impact that the large increase in suburban housing supply over the past four decades has had on the costs facing low-income renters in the urban core thus depends on whether the chains of moves it initiated reached low-income high-density neighborhoods. Our results show that they did not – instead, residential vacancy chains initiated by new low-density suburban single family housing end quickly, before they can reach the urban neighborhoods in which residents are most exposed to rising housing costs. This descriptive feature of vacancy chains, when viewed in light of our simulation results, suggests that the non-local price effects of new housing supply are concentrated in nearby submarkets and that the incidence of the benefits of additional housing therefore depends crucially on what kind of housing is built and where.

These results have important implications for housing policy that seeks to increase housing affordability. While many supply advocates argue that increasing supply of any kind will be effective at decreasing housing costs for all, this paper suggests that a more targeted approach is required if policymakers want to reduce costs in the least affordable neighborhoods or for the most rent-burdened households. Our results also suggest that a more targeted approach can be effectively guided by the distribution of vacancies created by new supply in a given submarket. Our simulation results show that the distribution of these vacancies is as predictive of variation in price effects as the cross-neighborhood substitution effects derived from individual demand elasticities. While housing policy will have to be guided by the costs of construction in different neighborhoods, as well as potential effects on local amenities, the observed number of vacancies connected to different kinds of new housing can help policymakers evaluate the non-local price effects of a given policy.

References

Anenberg, Elliot, and Edward Kung. 2020. "Can more housing supply solve the affordability crisis? Evidence from a neighborhood choice model." *Regional Science and Urban Economics* 80: 103363.

Asquith, Brian J., Evan Mast, and Davin Reed. 2021. "Local Effects of Large New Apartment Buildings in Low-Income Areas." *The Review of Economics and Statistics*: 1–46.

Baum-Snow, Nathaniel. 2023. "Constraints on City and Neighborhood Growth: The Central Role of Housing Supply." *Journal of Economic Perspectives* 37 (2): 53–74.

Baum-Snow, Nathaniel, and Lu Han. 2021. "The Microgeography of Housing Supply.": 90.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4): 588–638. Publisher: The University of Chicago Press.

Bayer, Patrick, Robert McMillan, and Kim Rueben. 2004. "An Equilibrium Model of Sorting in an Urban Housing Market."Technical Report w10865, National Bureau of Economic Research. Cambridge, MA.

Been, Vicki, Ingrid Gould Ellen, and Katherine O'Regan. 2019. "Supply Skepticism: Housing Supply and Affordability." *Housing Policy Debate* 29 (1): 25–40.

Bratu, Cristina, Oskari Harjunen, and Tuukka Saarimaa. 2023. "JUE Insight: City-wide effects of new housing supply: Evidence from moving chains." *Journal of Urban Economics* 133: 103528.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst. 2021. "Income Growth and the Distributional Effects of Urban Spatial Sorting.": 87.

Couture, Victor, and Jessie Handbury. 2020. "Urban revival in America." *Journal of Urban Economics* 119: 103267.

Damiano, Anthony, and Chris Frenier. 2020. "Build Baby Build?: Housing Submarkets and the Effects of New Construction on Existing Rents.": 48.

Demsas, Jerusalem. 2022. "Housing Breaks People's Brains." Section: Ideas.

Diamond, Rebecca, and Tim McQuade. 2019. "Who Wants Affordable Housing in Their Backyard? An Equilibrium Analysis of Low-Income Property Development." *journal of political economy*: 55.

Dougherty, Conor. 2020. "Build Build Build Build Build Build Build Build Build Build Build Build Build Build." *The New York Times*.

Friedrich, Michael, Tracy Rosenthal, Tracy Rosenthal, Ben Sandman, Ben Sandman, Sean T. Byrnes, Sean T. Byrnes, John Semley, John Semley, Scott W. Stern, Scott W. Stern, Scott Bradfield, and Scott Bradfield. 2023. "More Building Won't Make Housing Affordable." *The New Republic*.

Glaeser, Edward L, and Joseph Gyourko. 2003. "The Impact of Building Restrictions on Housing Affordability.": 19.

Gray, M. Nolan. 2021. "America Needs More Luxury Housing, Not Less." Section: Ideas.

Gyourko, Joseph, Albert Saiz, and Anita Summers. 2008. "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." *Urban Studies* 45 (3): 693–729. Publisher: SAGE Publications Ltd.

Handbury, Jessie. 2021. "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities." *Econometrica* 89 (6): 2679–2715. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA11738.

Howard, Greg, and Jack Liebersohn. 2021. "Why is the rent so darn high? The role of growing demand to live in housing-supply-inelastic cities." *Journal of Urban Economics* 124: 103369.

Kulka, Amrita, Aradhya Sood, and Nicholas Chiumenti. 2022. "How to Increase Housing Affordability? Understanding Local Deterrents to Building Multifamily Housing."

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider. 2015. "The Housing Market(s) of San Diego." *American Economic Review* 105 (4): 1371–1407.

Li, Xiaodi. 2019. "Do New Housing Units in Your Backyard Raise Your Rents?": 58.

Liu, Liyi, Doug McManus, and Elias Yannopoulos. 2022. "Geographic and temporal variation in housing filtering rates." *Regional Science and Urban Economics* 93: 103758.

Manson, Steven, Jonathan Schroeder, David Van Riper, Katherine Knowles, Tracy Kugler, Finn Roberts, and Steven Ruggles. 2023. "National Historical Geographic Information System: Version 18.0."

Marullo, Sam. 1985. "Housing Opportunities and Vacancy Chains." *Urban Affairs Quarterly* 20 (3): 364–388. Publisher: SAGE Publications.

Mast, Evan. 2021. "JUE Insight: The effect of new market-rate housing construction on the low-income housing market." *Journal of Urban Economics*: 103383.

Molloy, Raven, Charles G. Nathanson, and Andrew Paciorek. 2022. "Housing supply and affordability: Evidence from rents, housing consumption and household location." *Journal of Urban Economics* 129: 103427.

Nathanson, Charles G. 2020. "Trickle-down housing economics.": 46.

Pennington, Kate. 2021. "Does Building New Housing Cause Displacement?: The Supply and Demand Effects of Construction in San Francisco." SSRN Scholarly Paper 3867764, Social Science Research Network. Rochester, NY.

Piazzesi, Monika, Martin Schneider, and Johannes Stroebel. 2019. "Segmented Housing Search.": 87.

Rosenthal, Stuart 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.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply *." *Quarterly Journal of Economics* 125 (3): 1253–1296.

Song, Jaehee. 2021. "The Effects of Residential Zoning in U.S. Housing Markets.": 67.

Su, Yichen. 2021. "The Rising Value of Time and the Origin of Urban Gentri…cation.": 83.

Turner, Lena Magnusson. 2008. "Who Gets What and Why? Vacancy Chains in Stockholm's Housing Market." *European Journal of Housing Policy* 8 (1): 1–19. Publisher: Routledge.

White, Harrison C. 1971. "Multipliers, Vacancy Chains, and Filtering in Housing." *Journal of the American Institute of Planners* 37 (2): 88–94. Publisher: Routledge _eprint: https://doi.org/10.1080/01944367108977362.