Age Structure and Housing Affordability

Erica Moszkowski* Gregor Schubert† Shiyan Wei‡

October 21, 2024

Abstract

Changes in age structure across different locations coinciding with a lack of housing affordability have recently led to public discussions over the ability of young households to acquire "starter homes" if older households who own a substantial share of the housing stock choose to "age in place." To quantify the impact of demographic structure with regard to age on housing market affordability, we develop a quantitative spatial equilibrium model that allows for changes in demand for different housing types over the lifecycle in partially segmented housing markets. We develop a shift-share IV approach based on long-run age structures to estimate the model and quantify the degree to which changes in housing demand by different age groups spills over to housing affordability for other groups. We aim to show whether long-run changes in age structure can affect housing affordability.

Keywords: Demographics, home ownership, spatial models, age structure.

JEL Classification: G51, D83, R31, D10.

*Board of Governors of the Federal Reserve System. The analysis and conclusions set forth in this paper are those of the authors and do not indicate concurrence by the Board of Governors, the Federal Reserve System, or other members of their staffs. Contact: erica.l.moszkowski@frb.gov

†UCLA, Anderson School of Management. Contact: gregor.schubert@anderson.ucla.edu

‡UC-Irvine, Paul Merage School of Business. Contact: shiyanw4@uci.edu

1 Introduction

Housing affordability is an issue of increasing concern in major U.S. cities (Favilukis et al., 2023), which has recently been exacerbated by increases in housing demand associated with increases in remote work during the pandemic (Kmetz et al., 2022). However, the impact of these declines in affordability varies substantially across different demographic groups. In particular, one dimension often highlighted in public discourse is the contrast in housing wealth between older generations that were able to acquire housing before the large run-up in house prices in recent decades and the inability of younger households to acquire "starter homes." As new housing supply becomes increasingly scarce in desirable locations (Ganong and Shoag, 2017), competition between different age groups for the same housing stock could play an increasing role in driving house price growth. Strong house price growth in certain cities or housing segments resulting from changes in the size and income of different age groups could affect the ability of younger and lower-income households to access homeownership and/or productive labor markets (Mankiw and Weil, 1989; Green and Hendershott, 1996). However, this issue could be mitigated by the fact that preferences for different housing types vary over the lifecyle (Green and Lee, 2016), or by adjustments to housing supply from a price-responsive construction sector.

In this paper, we ask how demographic change, birth cohort- and age-specific preferences, and housing supply responses combine to affect housing affordability for different age groups and across geographic locations. First, we document changes in ownership patterns across age cohorts over the last 4 decades. Then, we provide causal reduced-form estimates of the house price impact of shifts in age structure in different cities and housing market segments. Finally, using the instrument developed for causal analysis, we specify and estimate a spatial equilibrium model that allows us to quantify demand spillovers between different age groups, as well as the equilibrium consequences of shifts in age structure on location choices and homeownership rates of different types of households. Our model allows us to explain to what degree differences in age composition can explain regional differences in past declines in housing affordability, as well as which housing types and locations are likely to see increasing demand in the future.

In order to establish that age structure is likely to play an important role in local housing demand, we first document, using Census and American Community Survey data for 1980-2020 that the share of homes of different sizes owned by older households has changed substantially over time. We find that the share of owner-occupied large homes (at least 4 bedrooms) that is owned by households headed by individuals at least 65 years old has increased substantially over time. The total number of these older households owning large homes more than tripled between 1980 and 2020, while the overall number of households in the US grew by only 57%. We show that this change can be explained by a steadily increasing propensity among older households to own large homes, accompanied by a rapid increase in the number of old households, particularly during the 2010s.

Next, we develop an approach to estimate causal effects of changes in the age structure on local housing markets. This requires overcoming the concern that local age structure is endogenous with regard to recent housing market dynamics, as populations of different ages respond differently to changes in local house prices. To address this issue, we use the fact that regional mobility is limited and that the age structure of the local population is therefore closely related to local variations in birth cohort sizes in earlier decades. We show that the size of a county's age cohort in later decades is strongly correlated with the size of the corresponding younger cohort

in the same county in earlier decades, even for lags of up to 40 years. We exploit this fact to develop an instrumental variable approach that exploits exogenous variation in the size of recent age cohorts that is associated with the size of historical local age cohorts, and thus unlikely to be driven by recent housing market dynamics. This instrument, in combination with variation in demand for different housing types over the lifecycle, allows us to generate exogenous variation in housing demand for different housing market segments within a city. We use this approach in a reduced-form setting to estimate the causal effect of changing age structures on local house prices in different cities. This approach also allows us to estimate separate housing supply elasticities by housing segment, which we believe are new in this literature.

In order to understand how these changes in age structure and lifecycle housing demand combine to affect location choices and affordability for different age groups and cities, we develop a quantitative spatial equilibrium model that captures these different dimensions of housing demand. We adapt the model of Ahlfeldt et al. (2015) to incorporate differential demand for different housing types by different age groups and estimate the model using data for a panel of U.S. cities. We estimate housing demand parameters by age group for different housing types using the historical age structure instrument, and also use Bartik (1991) shift-share shocks based on local industry exposure to national wage trends, as well as differences in local housing supply constraints to estimate the remaining model parameters. A key advantage of our structural model is that it allows us to account for both birth cohort and age effects (Green and Hendershott, 1996; Green and Lee, 2016).

We use our quantitative spatial model to quantify two important policy counterfactuals: (1) The degree to which housing affordability for younger households would improve or deteriorate if an increasing number of older households choose to "age in place", i.e. hold on to their large houses, compared to a scenario where older households return to the ownership patterns of the 1980s. (2) The expected change in house prices and population growth across U.S. cities in future decades if housing supply constraints remain the same and if housing supply is loosened through re-zoning.

2 The changing age structure of homeownership

In this section, we show that the share of large homes owned by seniors has increased dramatically over the last several decades. This could be driven by either an increase in the number of senior households, or an increase in the propensity of senior households to own large homes. We find that both have occurred, but that the majority of the change in the share of large homes owned by seniors since 2000 is primarily due to seniors' increased propensity to own large homes.

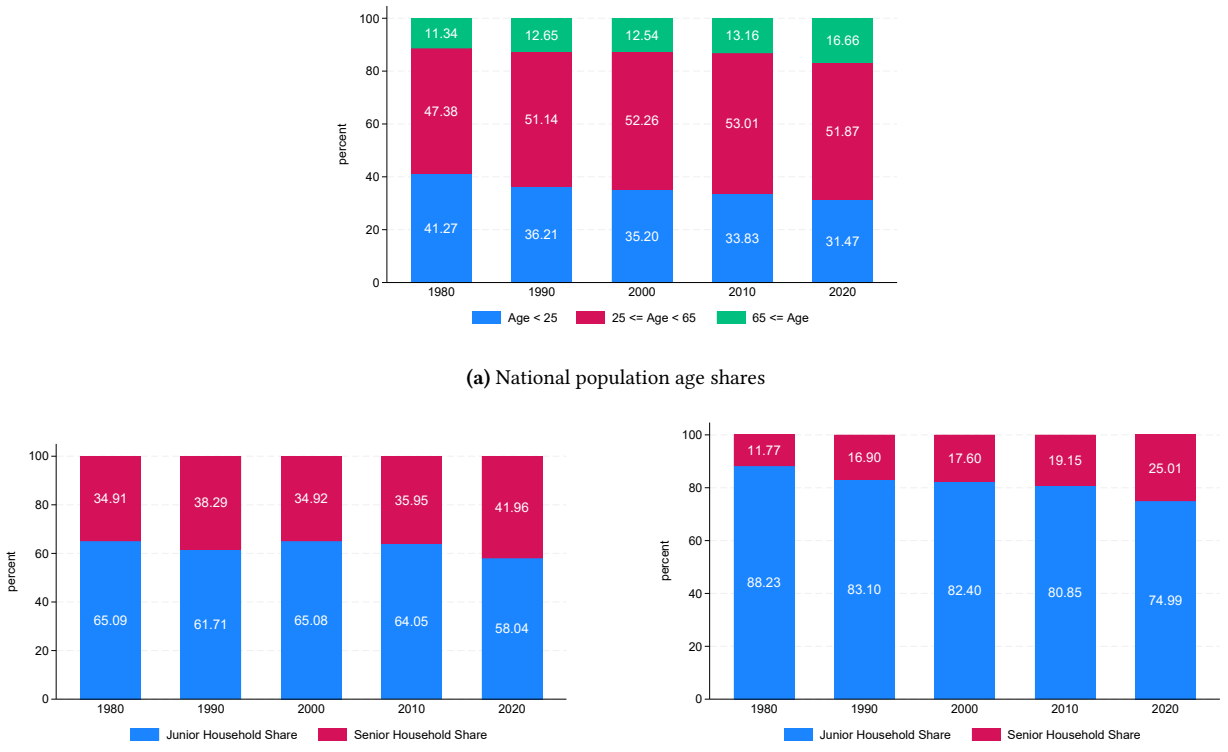
We use data from the Decennial Census and American Community Survey (ACS) as provided by IPUMS-NHGIS (Manson et al., 2023). We use county-level data on population by age from the Decennial Census for census years from 1970-2020[1]. We use the term "seniors" to refer to individuals aged 65 and above. We use the Decennial Census and ACS to calculate the number of households in each age bracket that live in each type of house (a house type is defined by the number of bedrooms it contains). When we refer to "large houses", we refer to those with four or more bedrooms. Data on housing characteristics, including including the number of bedrooms in the house and the age of the head of household, comes from the Decennial Census prior to

[1]We proxy for 2020 demographics using the ACS 5-year data for 2017-2021.

2010 and from non-overlapping 5-year vintages of the micro level American Community Survey (ACS) dataset from 2010 onward (Ruggles et al., 2024). Both sources sample 5% of the total U.S. population.

Figure 1 shows the evolution of the US population age distribution from 1980-2019 alongside the evolution of the distribution of the age of household heads for small (two bedroom) and large (four-plus bedroom) homes. It is clear that the ownership of large homes is becoming disproportionately concentrated among seniors: in 1980, seniors accounted for 11% of the population and owned 15% of large homes. By 2019, they accounted for 16% of the population (a 45% increase from 1980) and owned 29% of large homes (a 93% increase from 1980). The share of 2-bedrooms owned by seniors likewise has increased, but by less: seniors owned 35% of 2-bedrooms in 1980, and 42% of 2-bedrooms in 2020 (a 20% increase). Furthermore, the share of two-bedrooms owned by seniors increased by more than a few percentage points only from 2010-2020, coinciding with the large Baby Boomer generation reaching the age of 65.

Figure 1. Population share and large house ownership



(a) National population age shares

(b) Share of two-bedrooms owned by seniors

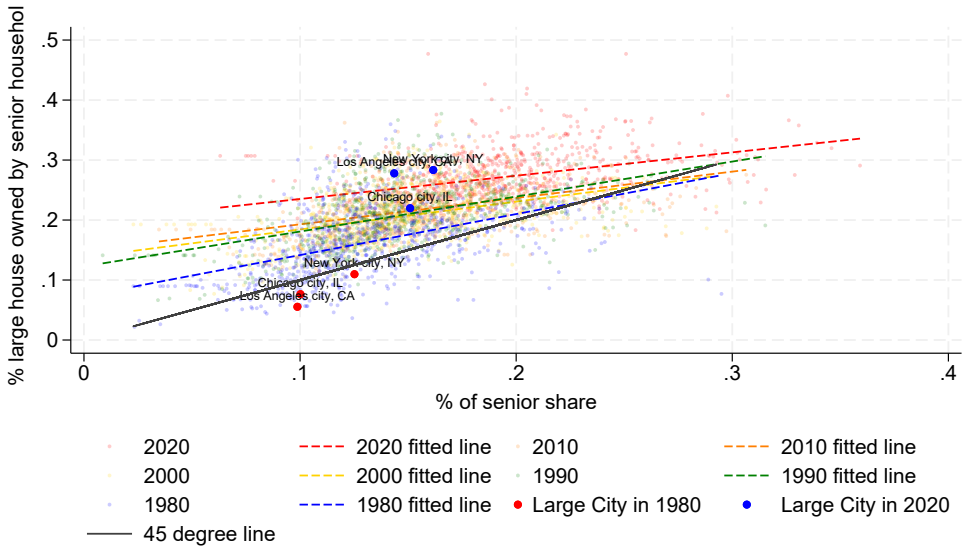
(c) Share of large houses (4BR+) owned by seniors

Figure 2 shows that this phenomenon appears to be occurring across all commuting zones.[2] This figure plots the senior share against the share of large homes owned by seniors for all commuting zones in the United States, for each of our data vintages. The correlation between the

[2]We conduct most of our analysis at the commuting zone level, since these areas best capture the geographic scale at which households make joint labor and housing decisions. Administrative boundaries, such as county borders, often subdivide the market for housing or labor supply in arbitrary ways. We aggregate the household count data to the commuting zone level using Autor et al. (2019)'s crosswalk file.

senior share and senior large home ownership share is positive. More notably, the fitted line shifts up over time: for the same senior share, the senior large home ownership share is higher in 1990 than in 1980, higher in 2000 than 1990, and so on.

Figure 2. Large house ownership pattern change by year



Note: Weighted by local population. Labeled the 2000 top 3 large cities by total population

Next, using the method of Oaxaca (1973) and Blinder (1973), we quantitatively decompose the increase in the senior large house ownership share into three factors: (A): changes in the number of senior households, (B) changes in the probability that senior households own large houses, and (C) the interaction between the two.

To make this precise, let D_{it} be the number of large houses in location i at time t owned by senior households, and let N_{it} be the number of seniors living in location i at time t. Then, we can write D_{it} as the product of the number of seniors and the number of large houses owned by each senior η_{it}:

$$D_{it} = \eta_{it} N_{it}$$

The Oaxaca-Blinder method decomposes period-to-period changes in D_{it} into three factors that represent changes in the "endowment" of the location with seniors, changes in the propensity to own large houses, and the interaction between these two forces:

$$\bar{D}_{i,t} - \bar{D}_{i,t-1} = \underbrace{\eta_{i,t-1}\left(\bar{N}_{i,t} - \bar{N}_{i,t-1}\right)}_{\textit{A: Endowment Effect}} + \underbrace{\left(\eta_{i,t-1} - \eta_{i,t}\right)\bar{N}_{i,t-1}}_{\textit{B: Coefficient Effect}} + \underbrace{\left(\eta_{i,t-1} - \eta_{i,t}\right)\left(\bar{N}_{lt} - \bar{N}_{i,t-1}\right)}_{\textit{Interaction Effect}} \quad (1)$$

Table 1 shows the results of this decomposition at the national level in the US. The top row shows the change in the number of large houses owned by seniors in each 10 year period (in thousands). There has been an increase in the ownership of large homes by older households in each decade. The subsequent rows show the change in the number of large houses owned by seniors that can be attributed to each of the factors. The importance of these different factors

changes over time: from 1980 to 2000, changes in the propensity to own large homes among old households explain more than half the observed increases in the number of old households owning large homes. Over the 2000-2020 period, seniors' propensity to own large homes continues to increase, but the relative effect of changes in the number of seniors becomes much larger. In the 2010-2019 period, the increase in the number of seniors accounts for 75% of the increase in the number of large homes owned by seniors.

Table 1 Oaxaca-Blinder Decomposition of National Large House (4BR+) Ownership by Senior Households ('000s)

	1980-1990	**1990-2000**	**2000-2010**	**2010-2020**
Δ large houses owned by HHs 65+ yrs old	987	741	1134	2311
Decomposition				
A: Due to Δ HHs 65+ yrs old	405	318	368	1341
B: Due to Δ propensity to own large houses	434	370	681	729
C: Interaction b/w propensity and Δ HHs	147	54	86	241

While these results are not causal, they provide some intuition for where the variation in ownership is coming from that is relevant for policy purposes. In the next section, we use the intuition from this decomposition to develop an age structure IV which we can use to identify the causal effect of changes in age structure on housing markets.

3 Age structure and house prices

The descriptive statistics presented in section 2 raise a broad question: how does the disproportionate increase in the number of large homes owned by seniors affect other demographic groups? Whether there is any effect at all is theoretically ambiguous. If housing is supplied elastically and younger households generally demand smaller homes (due to budget constraints or preferences), then younger households' ability to access homeownership should not be dramatically affected. Moreover, if financial markets for real estate efficiently incorporate information about demographic trends, predictable changes in the local age structure should be priced in advance, and should not be associated with contemporaneous changes in house prices in the affected market segments.

However, in the presence of housing supply constraints and a lack of sufficient intertemporal arbitrage capital, seniors' increasing demand for large homes may constitute a demand shock which drives up home prices not just for large homes, but also for smaller starter homes. If supply constraints vary across housing segments, households which would otherwise demand the most inelastically supplied housing are likely to be most affected. If demographic changes play an important role in local house price dynamics, this represents a factor influencing housing affordability that is currently underappreciated in policy discussions.

In this section, we show that long-run demographic trends causally impact local house prices. Causal interpretation of our results requires us to address the endogeneity of both age structure and demand for particular housing types, as households will change both where and in what type of housing they live if affordability changes. Specifically, we need to identify variation in demand that is uncorrelated with (1) local shocks to the desirability of a particular location which might in turn affect that location's age structure and (2) local changes in preferences for different

housing types which may be induced by changes in relative prices. To address this endogeneity, we construct a shift-share instrument that uses historical variation in demographic cohort sizes together with secular variation in preferences for different housing types at the national level. This instrument is plausibly exogenous with regard to these identification issues and can therefore identify the causal effects of changes in local housing demand on house prices in different segments of the market.[3]

We first describe the construction of our instrument, and then use the instrument to estimate the causal impact of segment-level demand shocks on segment-level prices. These estimates will deliver segment-specific housing supply elasticities which we use to assess which segments are most affected by demand shocks related to age structure.

3.1 Constructing demographic housing demand shocks

To identify the causal effect of changes in age structure on local housing markets, we need to find an exogenous factor driving differential demand for different local segments of the housing market. We construct exogenous variation in demand for housing in specific size segments by combining three different drivers of housing demand into a "shift-share" instrument that is plausibly exogenous at the local level: (1) Local age cohort sized predicted by historical age structure; (2) National changes in the propensity to own different housing segments conditional on owning; (3) National variation in headship rates by age group.

Age structure shocks. First, we want to identify exogenous variation in the local age composition. One of the key identification concerns is that people of particular ages will endogenously migrate between cities in response to house prices in preferred housing market segments. Therefore, to identify exogenous shifts in the local age structure—and the associated changes in demand for different housing types—we use the fact that historical cohort size predicts the later size of age cohorts as local residents age. We take advantage of the fact that geographic mobility is relatively low. As a result, historical birth rates and the size of younger cohorts together with the passage of time can predict the local size of older population cohorts in the long run. At the same time, historical age cohort sizes are unlikely to be affected by later housing market dynamics in the commuting zone and the ex ante predictable part of the contemporaneous age structure thereby avoids any reverse causality concern that contemporaneous demographics may be shaped by housing market dynamics.

Building on earlier work that highlighted the relationship between age structure and housing demand (e.g. Mankiw and Weil (1989); Green and Hendershott (1996); Green and Lee (2016)), we thus develop an instrumental variable approach that exploits exogenous variation in the size of contemporaneous age cohorts that is associated with the size of historical local age cohorts. That is, we can think of the number of local residents in a particular age group at a particular moment in time as:

$$N_{iat} = \left(1 - \sum_{l \neq i} m^{i \to l}_{a-k,(t-k,t)}\right) N_{i,a-k,t-k} + \sum_{l \neq i} m^{l \to i}_{a-k,(t-k,t)} N_{i,a-k,t-k},$$

[3]Our empirical approach for identifying housing demand by segment from historical differences in age structure relates to a small literature that studies the effect of age structure on financial outcomes. Crouzet et al. (2024) use the share of young vs. old potential customers in an area as drivers for firms to adopt Fintech innovations that are more likely to appeal to young customers.

where $m^{i \to l}_{a-k,(t-k,t)}$ is the share of all residents in location i at time $t-k$ that were $a-k$ years old that end up in location l at time t (subsuming death rates into mobility in this location framework). Here, the first term represents all formerly young historical residents that remained in the same location and are therefore now part of its older age groups, while the second term captures all those who have left.

As long as long-run mobility rates are sufficiently low, the second term will be small, and we can predict later demographic composition N_{iat} using only the local size of the same birth cohort in earlier years $N_{i,a-k,t-k}$. As long as we are interested in studying housing market dynamics in later years that are, conditional on control variables, uncorrelated with earlier age distributions other than through the housing demand channel, we can construct an instrument exploiting this historical variation. That is, we can use the historical counts of population by age group in each commuting zone to predict the local population in each age bracket in later decades.

We predict the contemporaneous local population in each age cohort using historical data in regressions of the form

$$\ln N_{iat} = \sum_{h=20,30} \left(\beta^a_{t-h} \ln N_{i,a-h,t-h} + \gamma^a_{t-h} \ln N_{i,t-h} \right) + \varepsilon_{iat}, \tag{2}$$

where the predictive variables include the 20-year and 30-year lagged sizes of the age cohorts that will become age group a at time t (the $N_{i,a-h,t-h}$ terms) as well as the lags of the total CZ population (the $N_{i,t-h}$ terms). For earlier decades and younger cohorts only the subset of variables is included that is available in our data.[4]

How well does historical age structure predict current demographics? None of our predictive regressions use data that predates the predicted cohort size by less than 20 years. Nonetheless, the predictive power of historical demographics is strong: The R-squared statistics of the estimations corresponding to equation 2 (which are shown for each age group and decade in Appendix Table 4) are above 95% for each age group and predicted decade from 1990-2020.

However, levels of age cohort size are more likely to be correlated with static unobserved location characteristics, such as being a retirement community. A more stringent test of the predictive power of historical age structures is whether this approach can predict *changes* in local age cohort sizes based on historical age structures. Table 2 shows that even changes in age structure are highly predictable using 20-year and 30-year lagged data. While the predictive power shown in the table is high on average–supporting the notion that we can use this variation in an instrument–it varies across age groups. In particular, in line with lifecycle mobility patterns, the size of very old (85+) and middle-aged cohorts (40-60 years old) tends to be more predictable using historical data. This is likely because geographic mobility weakens the persistence of the local age structure over time and both residents in their 20s and households near retirement age often relocate relative to where they lived 20 years earlier, while residents of working age who are likely to still live with their children are more likely to remain in the same city.

Overall, this approach allows us to generate predictions of local age cohort size that may affect housing demand and are plausibly exogenous with regard to contemporaneous housing market dynamics.

Lifecycle housing type preferences. Second, we estimate the variation of lifecycle housing

[4]This means that for cohorts that are 29 years old or younger, no 30-year lags of the specific age cohort size are included. For 1980 dependent variables, no 30-year lags are included.

Table 2
Predictive power of lagged demographics for changes in current age cohort size

The table shows the R-squared estimates from regressions of the form:

$$\Delta N_{iat} = \alpha_{at} + \beta_{at}\Delta\hat{N}_{iat} + \varepsilon_{iat},$$

estimated for each age group a at time t across commuting zones i. The columns show variation in the predicted period, while the rows represent different age cohorts whose sizes are being predicted.

	Year of Predicted Δ in Age Cohort Size			
	1990	2000	2010	2020
Age 20-24	.3846153	.2877941	.725975	.2680939
Age 25-29	.3508256	.4558129	.2830179	.0000665
Age 30-34	.3751904	.2514277	.032724	.1488763
Age 35-39	.7265397	.2387985	.0696514	.0778804
Age 40-44	.8147276	.5243419	.0525452	.2908121
Age 45-49	.6555247	.8538098	.7971416	.5524695
Age 50-54	.65348	.8582835	.890256	.2792799
Age 55-59	.7729855	.6940616	.9378503	.6650317
Age 60-64	.3831462	.5277428	.9569827	.8249102
Age 65-69	.2660826	.668747	.9006567	.9298187
Age 70-74	.2215532	.0857913	.5700551	.9673829
Age 75-79	.3474668	.47644	.6067051	.9305124
Age 80-84	.3743715	.859136	.9251407	.7932748
Age 85+	.795666	.9413264	.9537462	.9379485

demand with regard to different housing types at the national level. Conceptually, this common national variation in housing demand for different housing types by different age groups can be thought of as a "shifter" in a shift-share instrument. As they are common across locations, they are, by construction, not affected by contemporaneous local housing market dynamics, but generate variation over time in how local age groups might affect different housing market segments.

In each year t we compute the share of homes lived in by households with a head in age group a that are in housing segment j at the national level. The estimated lifecycle housing type composition $\hat{\omega}_{jat}$ reflects changes over time in how different age groups view different housing types.

Headship rates by age group. Third, we estimate the lifecycle-driven headship propensity ζ_{at} of different age groups. That is, we compute the ratio of the number of household heads of a given age to all persons in that age group at the national level. Accounting for headship rates is important, as they vary substantially both across age groups and over time, and they determine whether a particular age cohort gets to express their preferences for different housing types through a choice of residence. For example, in 2020 we find that 53% of 45-49-year-olds head their own household, while the headship rate for 25-to-29-year-olds is only 38%. Moreover, the propensity of 25-29-year-olds to start their own households has declined over time (the headship rate for this group in 1980 was 47%). While local headship rates are endogenous with regard to local housing market characteristics, for instance through affordability affecting household formation, national trends in headship rates are not affected by local dynamics but may capture long-run cultural or economic trends that then affect local markets through their age structure exposure to this "shifter".

Age structure housing demand instrument. Finally, we combine these three estimated components into a "shift-share" instrument that has the form

$$Z_{ijt}^{age} = \sum_{a} \hat{N}_{iat}^{\text{hist}} \times \hat{\zeta}_{at} \times \hat{\omega}_{jat} \tag{3}$$

where $\hat{N}_{ait}^{\text{hist}}$ is the predicted number of households headed by local age cohort a based on historical cohort sizes (from equation 2), $\hat{\zeta}_{at}$ is the estimated national headship rate of each age group, and $\hat{\omega}_{tj}^{(a)}$ is the predicted likelihood of age group a consuming housing of type j based on contemporaneous national trends. We refer to Z^{age} as our "age structure instrument" or "age structure variable".

The age structure variable encodes a location-by-type housing market's local exposure to different age cohorts' housing demand, as predicted by historical age structures in combination with national trends in household formation and lifecycle demand for different housing types. Its reliance on only long lags and national trends implies that it represents plausibly exogenous variation in local demand for different housing types. The instrument uses no local contemporaneous information and is therefore, conditional on controlling for appropriate fixed effects or persistent local characteristics, unlikely to be correlated with contemporaneous local housing market dynamics other than through the age structure-driven housing demand.

3.2 Age structure effects on house prices

With our age structure variable in hand, we can now estimate the effect of changes in local age structure on house prices. We aggregate Zillow's county-and-segment housing price indices (ZHVI) to the commuting zone level. Since our age structure variable is exogenous to contemporaneous local variation, an OLS regression will deliver the reduced-form effect of exogenous changes in demand driven by age structure on local, type-specific housing prices:

$$\ln p_{ijt} = \alpha_{jt} + \gamma \ln Z_{ijt}^{age} + \chi' \Gamma_{it} + \varepsilon_{ijt}, \tag{4}$$

Here, Γ_{it} represents other commuting zone characteristics that are often associated with house price changes, which we include as control variables: mean population density from the Census, the share of local land available for construction (Lutz and Sand, 2023), and average gross income growth per capita constructed from IRS data. We estimate equation 4 using OLS and the results are shown in column 1 of Table 3 in levels and in column 3 in a specification in changes. The estimated coefficients suggests that a 10 pp increase in predicted lifecycle-driven demand in that housing market (relative to national demand for that housing type and controlling for other commuting zone characteristics) is associated with a 1.2-1.9 pp increase in house prices in that market (we use the word "market" to refer to a housing segment in a particular location).

3.3 IV Estimates of Housing Supply Elasticities

The mechanism for age structure to translate into house price changes is via a shift in the housing demand curve (a move along the housing supply curve in each market segment). Therefore, we can also estimate the elasticity of housing supply directly by instrumenting for the quantity of housing units in each local housing market segment with the age structure instrument. The housing supply curve of interest is

$$\ln p_{ijt} = \alpha_t + \alpha_j + \beta \ln Q_{ijt} + \chi' \Gamma_{it} + \varepsilon_{ijt}, \tag{5}$$

where Γ_{it} represents control variables as defined above, and Q_{ijt} is the count of local housing units in each market segment.

Columns 2 and 4 of Table 3 show the result of implementing this IV estimation, either in levels or in changes. When estimating it in levels, we use $\ln Z_{ijt}^{age}$ to instrument for log housing units, and first difference both the latter and the instrument for the specification in changes in column 4. As shown by the first-stage F-statistics in both columns, the demographic IV is a strong instrument for market-level demand.

The estimated inverse supply elasticity in column 4 is 0.49 (corresponding to a supply elasticity of 2.04), which is well in line with other long-run elasticity estimates in the literature. For example, Saiz (2010) finds a metro area average of 1.75, consistent with our sample including more low-density commuting zones that are likely to have more elastic supply. However, as far as we know, the fact that our estimates are estimated at the level of *location-by-size* markets is new in this literature.

Table 3 Lifecycle housing demand effects on house prices by segment

Dependent variable:	Log ZHVI		Δ Log ZHVI	
Estimation:	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Log Housing Demand Shock	0.12***			
	(3.14)			
Log Housing Units		0.12***		
		(3.08)		
Δ Log Housing Demand Shock			0.19**	
			(2.27)	
Δ Log Housing Units				0.49**
				(2.32)
Observations	7,110	7,110	4,053	4,053
1st-stage KP F-stat.		10,893		159
# Bedrooms × Year FE	X	X	X	X
Density & land availability controls	X	X	X	X
Avg. income growth per capita	X	X	X	X

Note: The table shows panel regressions of log mean Zillow housing value indices in each commuting-zone-by-number-of-bedrooms market over time on the log of the lifecycle housing demand instrument (column 1) and on log occupied housing units in that market instrumented with the housing demand shock (column 2). The lifecycle housing demand instrument reflects the local demand for housing by segment that is predicted by the demographic structure of the location 20 years earlier and national trends in lifecycle housing demand. The estimation includes data for the years 2000, 2010, and 2020. Columns 3 and 4 show the results of estimating the same specifications in changes. All regressions are weighted by the CZ population in the year 2000 and control for bedroom-by-year FEs and commuting zone characteristics as indicated (see text for definitions). T-statistics based on heteroskedasticity-robust standard errors clustered at the CZ level shown in parentheses: * p<0.10, ** p<0.05, *** p<0.01.

This characteristics of the age structure instrument also allows us to differentiate the effects of demand shocks by housing type. As a result, we add to the existing literature on housing supply elasticities by estimating how the house price effects of demand shocks (and thus the slope of the supply curve) vary across housing types differentiated by size, and over time. The results of allowing the inverse supply semi-elasticity coefficients to vary along the unit size and year dimension are shown in Figure 3. The figure reveals a couple of new insights: On the one hand, the figure shows that supply semi-elasticities vary across housing types: units are more elastically supplied for the smallest and largest housing types, and less elastically for 2-bedroom and 3-bedroom units. This might be in part explained by the prevalence of these housing types interacting with construction capacity that is fungible between housing types but constrained at the city level: 2/3-bedroom units are the most common unit types in the housing market—and as a result any proportional change in demand for these categories requires a city to build a much larger absolute number of housing units than a proportional change in demand for the less common smaller and larger units. As a result, a given percentage increase in demand for 2- or 3-bedroom units would move further along the common city-level construction cost curve than the same percentage increase in demand for small and large units.

On the other hand, the figure shows that price responses to demand shocks are increasing over time – or, in other words, that housing supply is becoming more inelastic over the 2000-2020 period in every size segment. This is in line with a similar pattern observed by Sinai (2010) in earlier decades. Moreover, we can observe a flattening of the elasticity differences between the smallest and largest units: the house price response is about 41% larger for 5-bedroom units than for 1-bedroom units in 2000, but this gap shrinks to 7% by 2020. This pattern of faster decreases over time in the supply response for small units might explain the common assertion that the availability of "starter homes" has declined.[5]

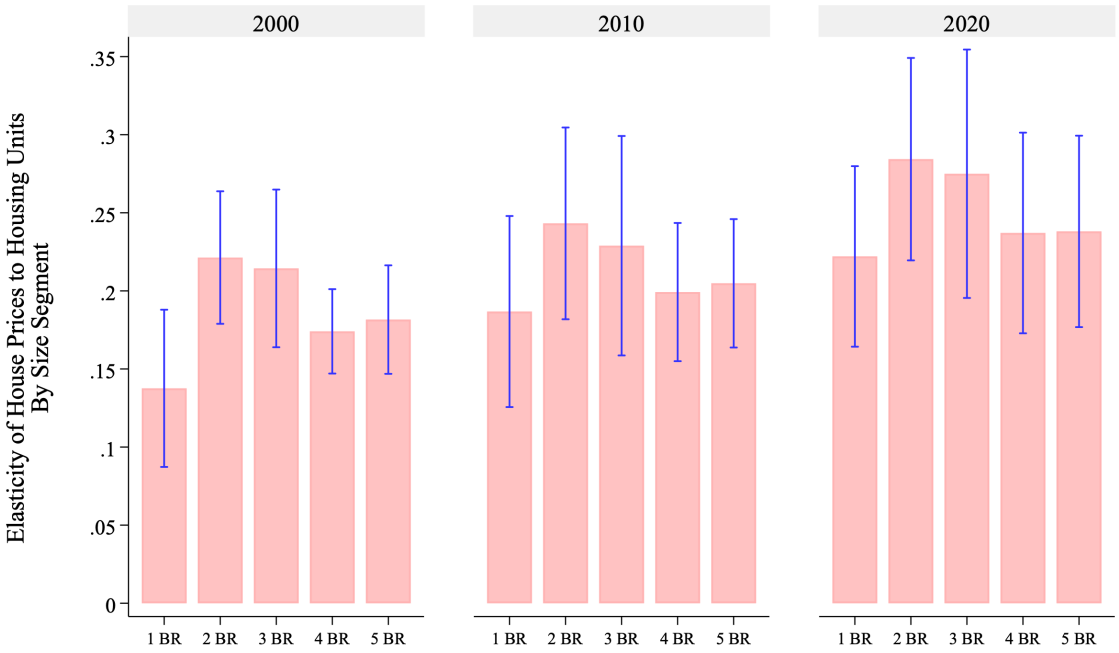
We can also use the change in the demographics-driven housing demand pressure in different housing markets to explain housing price growth. Figure 4 plots housing price growth at the market level from 2010-2020 against the change in demographics explained by historical age structure (that is, $\ln Z_{ijt}^{age} - \ln Z_{ij,t-1}^{age}$) over the same period for 2 bedroom homes. We can see that demographic shifts explained by historical age structure can predict cross-sectional differences in price growth of the same housing type segment across different commuting zones. There is a strong positive relationship between the change in the demographic demand driver and house price growth over this period (the corresponding R-squared is 8%), with an inverse elasticity corresponding to the slope of the fitted line in the graph of 0.54 (t-stat = 3.4).

3.4 Discussion

How does housing-type specific demand for housing aggregate into city-level house price changes? There is a substantial literature that estimates the effect of overall changes in local population on housing costs (e.g. Saiz (2003); Howard (2020); Schubert (2021)). However, the perspective of potentially segmented markets by housing type and variation in housing demand by age introduced in this paper allows for effect heterogeneity that prior research did not consider: specifically, if housing supply elasticities vary across different housing types, an inflow of particular demographic groups can have larger concentrated effects on house prices of particular

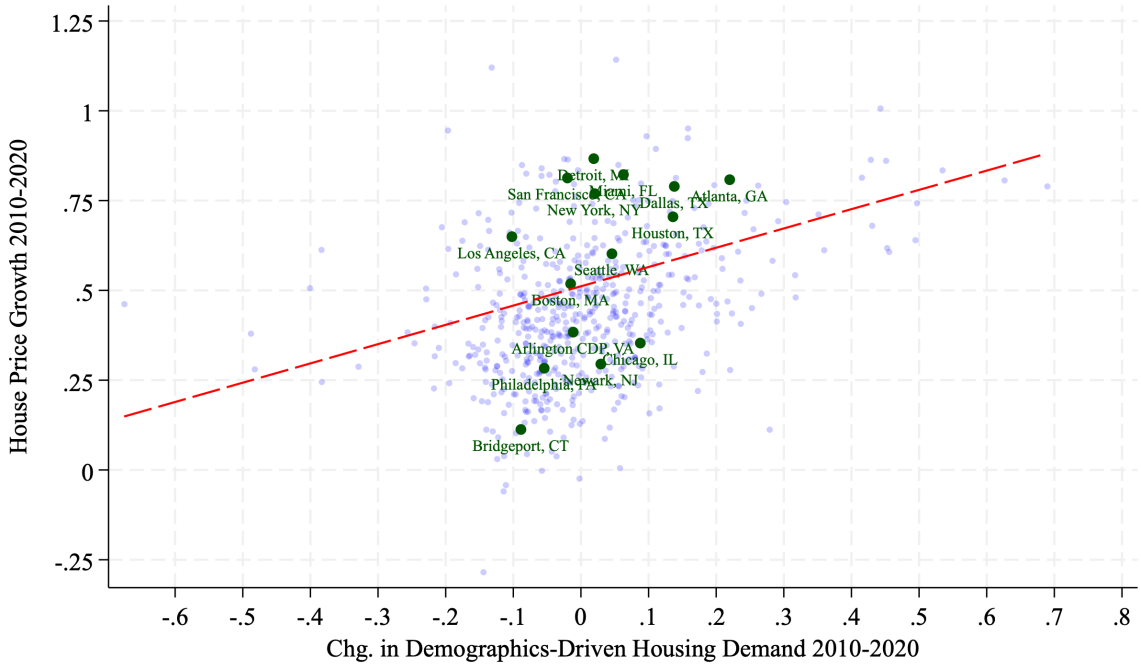
[5]See, for example: https://www.nytimes.com/2022/09/25/upshot/starter-home-prices.html

Figure 3. House price response to demand shocks by size segment.



Note: This figure shows the coefficient estimates from an IV panel regression of log mean Zillow housing value indices in each commuting-zone-by-number-of-bedrooms market in the years 2000, 2010, and 2020 on log occupied housing units in that market. The instrument is the change in lifecycle housing demand driven by the local age structure.

Figure 4. Demographic demand changes and house price growth: 2-bedroom markets 2010-2020. The figure shows the relationship between the change in log demographic housing demand predicted by historical age structure in the 2-BR housing segment in each CZ and the log change in the ZHVI house price index for that segment.



housing types than for the city overall. Similarly, aggregate city-level house prices can increase more for the same change in population, if the growing population groups either demand particularly inelastic housing types or have higher overall housing demand than other demographic groups.

Our housing supply estimates suggest a story of spillovers across housing segments which are not competed away due to supply constraints. Seniors, who make up a growing share of the American population, own an increasingly disproportionate share of large homes (Figure 2). This demand shock in the large-house segment puts upward pressure on prices for all housing types, but has an especially strong effect on those segments which have the most inelastic supply: 2 and 3 bedroom homes (Figure 3). In a standard model, increased prices in that segment will translate into negative welfare effects on households that would otherwise purchase those homes (in this case, younger households), either by causing those households to pay more or by inducing them to substitute to less-preferred types of housing.

How large are the welfare losses for young households triggered by senior households' increased demand for large homes? Given America's current age structure, how much new housing of each type would need to be supplied in each commuting zone to relieve upward pricing pressure on starter homes? Quantitatively answering these questions requires a model which accommodates spillovers across space and housing segments. The next section introduces our structural model, which allows households to substitute across geography and housing segments.

4 Model

4.1 Setup

We model consumer choices of residential locations and housing types. Our model closely follows Ahlfeldt et al. (2015), except that rather than modeling worker home and work location choices within a city, we model consumers' joint choice of city and house type.

Households choose from locations indexed by $i \in \mathcal{S} \equiv \{1, \dots, S\}$. Each location i has productivity Ψ_i and land area H_i^{supply}, represented by vectors $\mathbf{\Psi}$ and $\mathbf{H}^{\text{supply}}$, respectively. In each market, there are $j \in \mathcal{J} \equiv \{1, \dots, J\}$ types of housing available for purchase.

Households belong to age groups indexed by $a \in \mathcal{A} \equiv \{1, \dots, A\}$, with the national population share of age a given by $g(a)$.

4.2 Households

4.2.1 Individual household demand

Household utility depends on age, location, housing floor space, and final goods consumption. Utility is Cobb-Douglas in consumption and housing:

$$U_{ijo} = B_{ia} z_{ijo} \left(\frac{c_{ijo}}{\beta_a} \right)^{\beta_a} \left(\frac{\ell_{ijo}}{1 - \beta_a} \right)^{1-\beta_a}, \quad 0 < \beta_a < 1 \tag{6}$$

where i indexes home location, j indexes housing type, and o indexes households. The parameter β_a is a preference parameter trading off housing and non-housing consumption for workers of age a. c_{ijo} is consumption of the final good, ℓ_{ijo} is housing floor space, and B_{ia} are exogenous residential amenities capturing characteristics of location i for different age groups a. z_{ijo} is a Fréchet shock with cumulative distribution function

$$F(z_{ijo}) = \exp(-T_{ia} E_{ja} z_{ijo}^{-\epsilon_a}), \quad T_{ia}, E_{ja} > 0, \epsilon_a > 1 \tag{7}$$

T_{ia} determines average utility from the home location for consumers of age a. E_{ja} determines average utility from the housing type for consumers of age a. ϵ_a governs the dispersion of z_{ijo}, with $\mathbb{E}[z_{ijo}] = \Gamma\left(\frac{\epsilon_a - 1}{\epsilon_a}\right) (T_{ia} E_{ja})^{\frac{1}{\epsilon_a}}$.

Assigning the final good as the numeraire, the budget constraint of the worker is

$$w_{ia} = c_{ijo} + p_{ij} \ell_{ijo} \tag{8}$$

where p_{ij} is the price per square foot of residential floor space in city i for housing type j and w_{ia} is the wage in location i for households of age a. We assume that individuals work in the same location as they live, since our empirical specification will use commuting zones as locations.

An individual household's demand for floor space ℓ_{ijo}^* and final goods c_{ijo}^* follow the usual Cobb-Douglas demand functions:

$$\ell_{ijo}^* = (1 - \beta_a) \frac{w_{ia}}{p_{ij}} \tag{9}$$

$$c_{ijo}^* = \beta_a w_{ia} \tag{10}$$

Indirect utility for a consumer o living in location i and house type j is therefore:

$$u_{ijo} = B_{ia} z_{ijo} w_{ia} p_{ij}^{\beta_a - 1} \tag{11}$$

where B_{ia} represents the exogenous residential amenities capturing the characteristics of location i for different age groups a, and z_{ijo} is an idiosyncratic shock.

Since indirect utility is a monotonic function of the idiosyncratic shock z_{ijo}, which follows a Fréchet distribution, the indirect utility for workers living in city i and housing type j also follows a Fréchet distribution. The probability that a consumer of age a maximizes their utility by living in city i and housing type j is given by:

$$\pi_{ija} = \frac{T_{ia} E_{ja} p_{ij}^{-\epsilon_a(1-\beta_a)} (B_{ia} w_i)^{\epsilon_a}}{\sum_{s=1}^{S} \sum_{h=1}^{J} T_{sa} E_{ha} p_{sh}^{-\epsilon_a(1-\beta_a)} (B_{sa} w_{sa})^{\epsilon_a}} \tag{12}$$

So a consumer of age a is more likely to live in location i if it has good amenities (B_{ia}), high average utility (T_{ia}), high wages (w_{ia}), and low housing prices (p_{ij}). Similarly, a consumer of age a is more likely to choose house type j if it provides higher average utility (E_{ja}) or has lower prices p_{ij}.

4.2.2 City-house type housing demand

Total demand for housing type j in city i is given by

$$Q_{ij}^{demand} = \sum_a \underbrace{\ell_{ija}^*}_{\text{floorspace/HH}} \cdot \underbrace{\pi_{ija}}_{Pr(ij|a)} \cdot \underbrace{g(a) \cdot Pop}_{\text{\# age-a HHs}} \tag{13}$$

4.2.3 City-level labor supply

Assume that if $a \geq a^{retire}$ then the worker is retired, and that all other workers are employed. City labor supply is therefore:

$$L_i^{supply} = \sum_{a=1}^{a^{retire}} \frac{T_{ia} E_{ja} p_{ij}^{-\epsilon_a(1-\beta_a)} (B_{ia} w_{ia})^{\epsilon_a}}{\sum_{s=1}^{S} \sum_{h=1}^{J} T_{sa} E_{ha} p_{sa}^{-\epsilon_a(1-\beta_a)} (B_{sa} w_{sa})^{\epsilon_a}} \cdot g(a) \cdot Pop \tag{14}$$

4.2.4 Conditional location choice probabilities

Conditional on living in home type j, the probability a consumer of age a lives in city i is

$$\pi_{ija|ja} = \frac{\pi_{ija}}{\sum_i \pi_{ija}} = \frac{T_{ia} B_{ia}^{\epsilon_a} p_{ij}^{-\epsilon_a(1-\beta_a)} w_{ia}^{\epsilon_a}}{\sum_{s=1}^{S} T_{sk} B_{sa}^{\epsilon_a} p_{sj}^{-\epsilon_a(1-\beta_a)} w_s^{\epsilon_a}} \tag{15}$$

Notice that E_{ja}, the average preference age-k households have for housing type j, cancels out. This will aid in our identification of locational preferences.

4.2.5 Conditional house type choice probabilities

Conditional on living in i, the probability of a consumer of age k lives in a house of type j is

$$\pi_{ija|ia} = \frac{\pi_{ija}}{\sum_j \pi_{ija}} = \frac{E_{ja} p_{ij}^{-\epsilon_a(1-\beta_a)}}{\sum_{h=1}^{H} E_{hk} p_{ih}^{-\epsilon_a(1-\beta_a)}} \tag{16}$$

Notice that $T_{ia}B_{ia}^{\epsilon_a}$, the average preference age-k households have for living in city i, cancels out. This will aid in our identification of preferences over housing types.

4.2.6 Population mobility

Free movement across cities implies that each agent's choice delivers them the reservation level of utility in the wider economy $\bar{U}$:

$$\mathbb{E}_z[\max_{ij} u_{ijo}] = \gamma_k \Big(\sum_i \sum_j -T_{ia}E_{ja} \Big(B_{ia} w_i \Big)^{\epsilon_a} p_{ij}^{-\epsilon_a(1-\beta_a)} \Big)^{\frac{1}{\epsilon_a}} = \bar{U} \tag{17}$$

4.3 Final goods producers

Firms use labor L and commercial floor space q_i to produce the final consumption good. We assume a representative firm in each location i and Cobb-Douglas production technology with constant returns to scale:

$$y_i(L_i) = A_i L_i^{\alpha} q_i^{1-\alpha} \tag{18}$$

where A_i is firm productivity. Firm profits (taking prices and wages as given, and final goods as numeraire) can be expressed as:

$$\Pi_i = y_i(L_i) - L_i w_i - r_i q_i \tag{19}$$

where r_i is the rent on their commercial floor space. Firms choose labor and floor space inputs to maximize profits, leading to factor input demand:

$$L_i^{demand} = \Big(\frac{A_i \alpha}{w_i} \Big)^{\frac{1}{1-\alpha}} q_i \tag{20}$$

$$q_i^{demand} = \Big(\frac{A_i(1-\alpha)}{r_i} \; Big \Big)^{\frac{1}{\alpha}} L_i \tag{21}$$

Plugging 21 into 20 yields and expression for commercial floor space rents as a function of parameters and wages:

$$r_i = (1-\alpha) A_i^{\frac{1}{1-\alpha}} \Big(\frac{\alpha}{w_i} \Big)^{\frac{\alpha}{1-\alpha}} \tag{22}$$

So, all else equal, firms which are more productive and pay lower wages are willing to pay more for their commercial real estate.

4.4 Agglomeration

We account for agglomeration effects in both production of the final good and in residential amenities. We assume no spatial spillovers in agglomeration between cities.

We parameterize firm productivity A_i as a function of employment density in the city:

$$A_i = a_i Y_i^{\lambda}, \qquad Y_i = \frac{L_s}{H_s} \tag{23}$$

λ controls the curvature of agglomeration with respect to population density in determining overall productivity (relative importance of agglomeration forces relative to fundamentals). a_i represents the exogenous production fundamentals of location i.

Similarly, residential amenities B_{ia} depend on fundamentals b_{ia} and residential externalities Ω_{ia}:

$$B_{ia} = b_{ia}\Omega_{ia}^{\eta_k}, \qquad \Omega_i = \frac{\pi_i^{loc} \cdot Pop}{H_i} \tag{24}$$

where π_i^{loc} is the share of the overall population living in city i:

$$\pi_i^{loc} = \sum_k \sum_{j=1}^{H} \pi_{ija} f(k) \tag{25}$$

4.5 Developers

We assume a competitive construction sector for the construction of all floor space in location i. Equilibrium in the floor space market will require no-arbitrage between the development of commercial and residential floor space after the tax equivalent of land use regulations (which we assume are paid by developers). If all uses are present in a given city (which is the case in all US commuting zones) then no-arbitrage between these uses requires:

$$\begin{aligned} p_{ij} &= \xi_{ij} r_i \quad \forall i \in \mathcal{S}, j \in \mathcal{H} \\ \theta_{ij} &\in (0,1) \quad \forall i \in \mathcal{S}, j \in \mathcal{H} \\ \sum_j \theta_{ij} &< 1, \quad \forall i \in S \end{aligned} \tag{26}$$

where θ_{ij} denotes the share of floor space used for house type $j \in \mathcal{H}$, ξ_{ij} captures the tax equivalent of land use regulations that drive a wedge between residential house prices and commercial land prices, and $1 - \sum_j \theta_{ij}$ is the share of floor space dedicated to commercial use. We allow the wedge ξ_{ij} between residential type j and commercial floor prices to vary across locations.

We assume that the production function of floor space square footage is Cobb-Douglas in capital K and land H:

$$Q_i^{supply} = \underbrace{K_i^{\mu}}_{\text{capital}} \underbrace{H_i^{1-\mu}}_{\text{land}}$$

Let $P_i \equiv \max\{r_i, \max_j\{p_{ij}\}\}$ the price per square foot of the highest-profit use of floor space for developers.

Developer profit maximization and price-taking together implies that the dual cost function for floor space is

$$P_i = \underbrace{\mu^{-\mu}(1-\mu)^{-(1-\mu)} r_k^{\mu}}_{\equiv \chi} (r_i^{land})^{1-\mu}$$

where r_i^{land} is the price of land and r_k are rents paid to capital. Since the price for capital is the same across all locations, we can define $\varphi_i \equiv K_i^{\mu}$ and $\chi \equiv \mu^{-\mu}(1-\mu)^{-(1-\mu)} r_k^{\mu}$. This makes clear that floor space is a monotonic function of land area:

$$\underbrace{Q_i^{supply}}_{\text{floor space}} = \varphi_i \ \underbrace{H_i^{1-\mu}}_{\text{land input}}$$

and the price of housing is a monotonic function of the price of land:

$$P_i = \chi \cdot (r_i^{land})^{1-\mu}$$

Now we can solve for land demand H_i in terms of φ_i. First observe that:

$$\varphi_i = K_i^{\mu} \implies \varphi_i^{\frac{1}{\mu}} = (K_i^{\mu})^{\frac{1}{\mu}} = K_i$$

Then we can solve for total land demand in each real estate market as a function of the price of land:

$$H_i^{demand} = \left(\frac{\varphi_i P_i (1-\mu)}{r_i^{land}} \right)^{1/\mu} \tag{27}$$

The total supply of floor space created for each segment is given by:

$$Q_{ij}^{supply} = \theta_{ij} \phi_i H_i^{1-\mu} \quad \forall j \in \mathcal{H} \tag{28}$$

$$Q_{i,commercial}^{supply} = (1 - \sum_j \theta_{ij}) \varphi_i H_i^{1-\mu} \quad \forall j \in \mathcal{H} \tag{29}$$

4.6 Equilibrium

We now show that there is a unique mapping from the observed variables and a guess of parameter values to unobserved location characteristics. We will be unable to separately identify age-specific locational amenities B_{ia} from age-specific locational preferences T_{ia}, since they enter the model isomorphically. Therefore we define the following composites:

$$\tilde{B}_{ia} = B_{ia} T_{ia}^{1/\epsilon_a} \tilde{b}_{ia} = b_{ia} T_{ia}^{1/\epsilon_a} \tag{30}$$

Given the model's parameters $\{\alpha, \mu, \lambda, \boldsymbol{\beta}, \boldsymbol{\epsilon}, \boldsymbol{\eta}\}$, reservation level of utility in the wider economy $\bar{U}$, and vectors of exogenous location characteristics $\{\mathbf{a}, \tilde{\mathbf{b}}, \mathbf{E}, \varphi\}$, the general equilibrium of the model is given by a set of vectors $\{\pi_{\mathbf{ija}}, \mathbf{p}, \mathbf{r}^{\mathbf{land}}, \mathbf{r}, \mathbf{w}, H^{supply}\}$, such that:[6]

1. Population mobility (equation 17) holds

[6]Bold characters represent vectors of location-, house type-, or age- specific quantities. For example, $\boldsymbol{\alpha} \equiv \{\alpha_i\}_{i \in S}$

2. Housing choice probabilities follow equation 12

3. Housing markets clear (housing floor space demand, given by equation 13, equals housing floor space supply, given by equation 28)

4. The commercial real estate market clears (commercial floor space demand, given by equation 21, equals commercial floor space supply, given by equation 29).

5. Land markets clear (land demand, given by equation 27, equals exogenous land supply H_i^{supply})

6. Labor markets clear (labor demand, given by equation 20, equals labor supply, given by equation 14)

4.7 Identification of location-specific structural residuals

Ahlfeldt et al. (2015) show that given known values of the parameters and the observed data, there exist unique vectors of the unobserved location characteristics $\{\mathbf{E}, \tilde{\mathbf{b}}, \psi\}$ that are consistent with the data being an equilibrium of the model. We follow their approach, using the conditional choice probabilities $\pi_{\mathbf{ija}}, \pi_{\mathbf{ija|i}}, \pi_{\mathbf{ija|j}}$ to identify E_{ja}, $\tilde{b}_{ia}$, and ψ_i (up to a normalization). Because we intend to estimate our model at the commuting zone level, we assume consumers live and work in the same location. This assumption causes these equations to differ slightly from those in Ahlfeldt et al. (2015), but the intuition for identification is the same.

From equation 16, the log difference in adjusted production fundamentals in each city relative to the geometric mean can be written as

$$\underbrace{\ln\left(\frac{E_{ja}}{\bar{E}_a}\right)}_{\equiv \hat{E}_{ja}} = \ln\left(\frac{\pi_{ija|i}}{\bar{\pi}_{ia}}\right) + \epsilon_a(1-\beta_a)\ln\left(\frac{p_{ij}}{\bar{p}_i}\right) \tag{31}$$

where a bar over a variable represents its geometric mean across house types j.

From equation 15, the log difference in adjusted residential fundamentals in each city relative to the geometric mean can be written as

$$\underbrace{\ln\left(\frac{\tilde{b}_{ia}}{\bar{\tilde{b}}_a}\right)}_{\equiv \hat{b}_{ia}} = \frac{1}{\epsilon_a}\ln\left(\frac{\pi_{ija|ja}}{\bar{\pi}_{ja}}\right) - \eta_a \ln\left(\frac{\Omega_i}{\bar{\Omega}}\right) - \ln\left(\frac{w_{ia}}{\bar{w}_a}\right) + (1-\beta_a)\ln\left(\frac{p_{ij}}{\bar{p}_j}\right) \tag{32}$$

where a bar over a variable represents its geometric mean across locations i.

From equation 15, the log difference in productivity fundamentals in each city relative to the geometric mean can be written

$$\underbrace{\ln\left(\frac{\psi_i}{\bar{\psi}}\right)}_{\equiv \hat{\psi}} = (1-\alpha)\ln\left(\frac{r_i}{\bar{r}}\right) - \lambda \ln\left(\frac{Y_i}{\bar{Y}}\right) + \alpha \ln\left(\frac{w_i}{\bar{w}}\right) \tag{33}$$

5 Quantitative Estimation

This section estimates the quantitative model's agglomeration and preference parameters (λ, η, ϵ) using a GMM approach. We take advantage of the shift-share instrument developed in section 3.1.

5.1 Moment conditions

We use three sets of moment conditions to identify our three sets of parameters. All three moment conditions impose that the changes in residential and production fundamentals (period-to-period changes in the quantities defined by equations 31, 32, and 33) from decade to decade are uncorrelated with exogenous changes in local residential concentration and business activity.

Our first moment condition (equation 34) asserts that changes in the exogenous quality of type-j housing for age-a households, E_{ja}, are orthogonal to changes in demand for housing of type j as predicted by historical age structure. This pins down the preference dispersion parameter ϵ_a, conditional on our calibrated β_a. This dispersion parameter could equivalently be estimated from a regression of changes in age-specific housing-type preferences E_{ja} on changes in house prices, instrumenting for house prices with predicted age structure.

$$m_{ja}^{H} = \mathbb{E}[\Delta \ln(\hat{E}_{ja}(\Lambda)) \times \Delta Z_{ij}^{age}] = 0 \quad \forall j \in \mathcal{H}, a \in \mathcal{A} \tag{34}$$

Our second moment condition (equation 35) asserts that changes in exogenous location quality $\ln(\tilde{b}_{ia}/\bar{\tilde{b}})$ for households of age a are orthogonal to changes in predicted demand to live in location i as predicted by historical age structure. This is equivalent to estimating the residential spillover parameters η_a from a regression of changes in locational amenities $\tilde{B}$ on changes in residential amenities (Ω), instrumenting for residential amenities with predicted changes in demand attributable to changes in age structure Z_i^{age}. The exclusion restriction is that predicted changes in age structure affects changes in amenities Ω_{ia}, but not locational fundamentals b_{ia}.

$$m_a^R(\mathbf{X}_i, \Lambda) = \mathbb{E}\left[\Delta \ln(\hat{b}_{ia}) \times \Delta Z_{ij}^{age}\right] = 0 \tag{35}$$

The third moment condition asserts that changes in exogenous location productivity is orthogonal to predicted changes in employment. This is equivalent to estimating the productivity spillover parameter λ from a regression of changes in productivity ϕ on changes in agglomeration spillovers (Y), instrumenting for agglomeration spillovers with employment predicted from historical age structure. We calculate predicted employment from historical age structure: $Z_i^{emp} = \sum_{a<a^{retire}} \hat{N}_{ia}^{hist}$.

$$m_{ia}^C(\mathbf{X}_i, \Lambda) = \mathbb{E}\left[\Delta \ln\left(\hat{\phi}\right) \times Z_i^{emp}\right] = 0 \tag{36}$$

5.2 Identification of agglomeration parameters

The key parameters we estimate are: (1) λ, which governs productivity spillovers within each commuting zone, (2) η_a, which governs residential spillovers for households of age a within each commuting zone, and (3) ϵ_a, which governs the degree of preference dispersion across households of age a.

We calibrate the remaining parameters. Specifically, we set firms' commercial floor space expenditure share $1 - \alpha$ to 0.2 and the share of land in construction costs $1 - \mu = 0.25$. We also set households' housing expenditure share $1 - \beta_a$ according to Foster (2015) for each age group a.

Conditional on the calibrated parameters, the estimated parameters are identified from variation in the data. All of the estimated parameters are identified by our moment conditions, and intuition largely follows Ahlfeldt et al. (2015). Because the structure of the data allows us to separately recover production and preference fundamentals using observed data, the parameters $\lambda, \eta_a, \epsilon_a$ could be estimated from a set of regressions. η_a can be identified from a regression of changes in locational amenities $\hat{\tilde{b}}_{ia}$ on changes in our age structure instrument Z_{ia}^{age}. Similarly,λ could be estimated from a regression of changes in locational production fundamentals $\hat{\psi}$ on changes in employment density, instrumenting for employment density with a version of our demographics-driven demand shock ΔZ^{age} which predicts the size of the work force from demographic trends.

6 Counterfactuals

We plan to use the quantitative model to quantify two important policy counterfactuals: (1) The degree to which housing affordability for younger households would improve or deteriorate if an increasing number of older households choose to "age in place", i.e. hold on to their large houses, compared to a scenario where older households return to the ownership patterns of the 1980s. (2) The expected change in house prices and population growth across U.S. cities in future decades if housing supply constraints remain the same and if housing supply is loosened through re-zoning.

To implement the first counterfactual, we plan to substitute estimated $\hat{E}_{ja}$ values for j = large houses and a = seniors with a forecast value of $\hat{E}$. We plan to generate this forecast by estimating a time trend in our historical $\hat{E}$ estimates for this group. We will also use our age structure instrument to predict the sizes of future age cohorts. Given these values, we will solve the model and study changes in welfare for younger age groups.

[[TO BE COMPLETED]]

References

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The Economics of Density: Evidence From the Berlin Wall. *Econometrica*, 83(6):2127–2189.

Autor, D., Dorn, D., and Hanson, G. (2019). When work disappears: Manufacturing decline and the falling marriage market value of young men. *American Economic Review: Insights*, 1(2):161–178.

Bartik, T. J. (1991). Who benefits from state and local economic development policies?

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, pages 436–455.

Crouzet, N., Ghosh, P., Gupta, A., and Mezzanotti, F. (2024). Demographics and technology diffusion: Evidence from mobile payments. *Available at SSRN*.

Favilukis, J., Mabille, P., and Van Nieuwerburgh, S. (2023). Affordable housing and city welfare. *The Review of Economic Studies*, 90(1):293–330.

Foster, A. C. (2015). Consumer expenditures vary by age. *Bureau of Labor Statistics*, 4(14).

Ganong, P. and Shoag, D. (2017). Why has regional income convergence in the us declined? *Journal of Urban Economics*, 102:76–90.

Green, R. and Hendershott, P. H. (1996). Age, housing demand, and real house prices. *Regional Science and Urban Economics*, 26(5):465–480.

Green, R. K. and Lee, H. (2016). Age, demographics, and the demand for housing, revisited. *Regional Science and Urban Economics*, 61:86–98.

Howard, G. (2020). The migration accelerator: Labor mobility, housing, and demand. *American Economic Journal: Macroeconomics*, 12(4):147–179.

Kmetz, A., Mondragon, J., and Wieland, J. (2022). Remote work and housing demand. *FRBSF Economic Letter*, 2022(26):1–5.

Lutz, C. and Sand, B. (2023). Highly disaggregated land unavailability. *Available at SSRN 3478900*.

Mankiw, N. G. and Weil, D. N. (1989). The baby boom, the baby bust, and the housing market. *Regional science and urban economics*, 19(2):235–258.

Manson, S., Schroeder, J., Riper, D. V., Knowles, K., Kugler, T., Roberts, F., and Ruggles, S. (2023). Ipums national historical geographic information system: Version 18.0 [dataset]. http://doi.org/10.18128/D050.V18.0.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International economic review*, pages 693–709.

Ruggles, S., Flood, S., Sobek, M., Backman, D., Chen, A., Cooper, G., Richards, S., Rodgers, R., and Schouweiler, M. (2024). Ipums usa: Version 15.0 [dataset]. minneapolis, mn: Ipums. https://doi.org/10.18128/D010.V15.0.

Saiz, A. (2003). Room in the kitchen for the melting pot: Immigration and rental prices. *Review of Economics and Statistics*, 85(3):502–521.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296.

Schubert, G. (2021). House price contagion and us city migration networks. *Job Market Paper. Cambridge, Mass.: Joint Center for Housing Studies, Harvard University*.

Sinai, T. (2010). Feedback between real estate and urban economics. *Journal of Regional Science*, 50(1):423–448.

A Appendix Figures

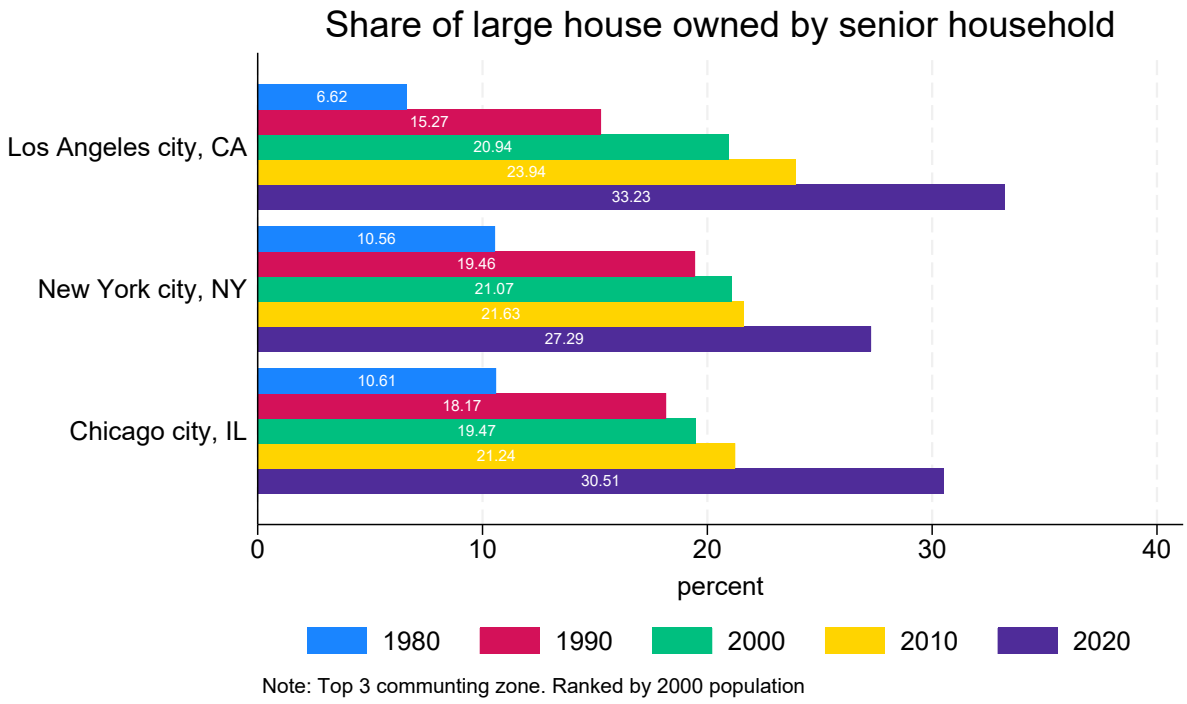
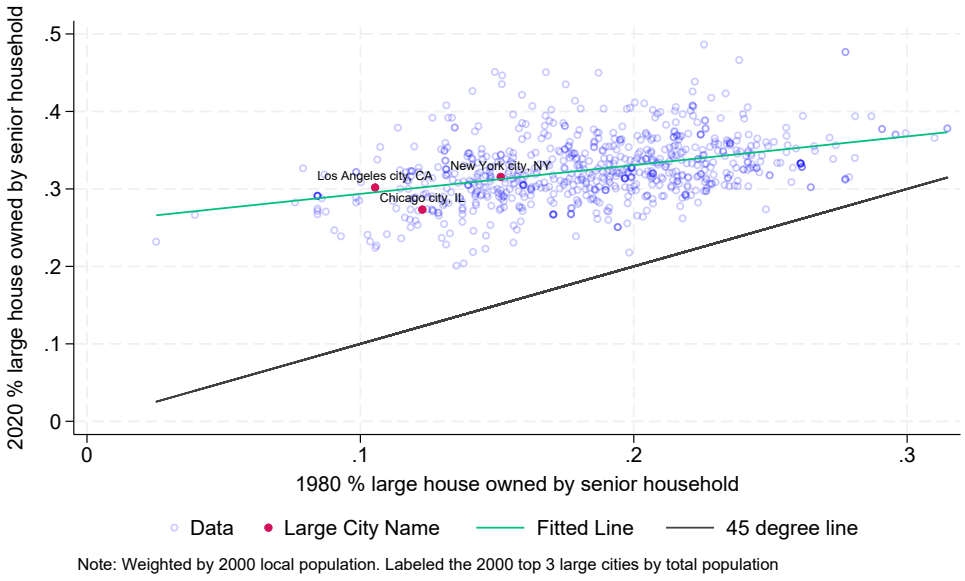
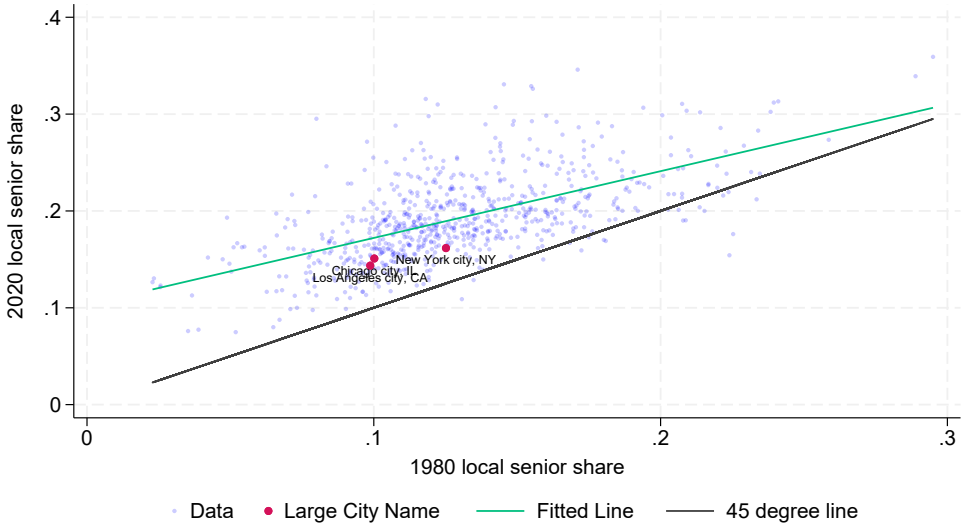


Figure 5. Caption



Note: Weighted by 2000 local population. Labeled the 2000 top 3 large cities by total population

Figure 6. Large house ownership transition



Note: Weighted by 2000 local population. Labeled the 2000 top 3 large cities by total population

Figure 7. Senior share transition

B Appendix Tables

Table 4
Predictive power (R-squared) of lagged demographics for levels of current age cohort size

The table shows the R-squared estimates from regressions of the form: $N_{iat} = \alpha_{at} + \beta_{at}\hat{N}_{iat} + \varepsilon_{iat}$ estimated for each age group a at time t across commuting zones i. The columns show variation in the predicted period, while the rows represent different age cohorts whose sizes are being predicted.

	Year of Predicted Age Cohort Size				
	1980	1990	2000	2010	2020
Age 20-24	.9564746	.9712221	.9670393	.9717454	.9758227
Age 25-29	.9615297	.9841488	.9823291	.9856898	.9855105
Age 30-34	.9613398	.9848946	.9876163	.9885588	.9896188
Age 35-39	.9657676	.9834606	.9890602	.989762	.9875987
Age 40-44	.9732971	.9838797	.9908777	.9918644	.9901475
Age 45-49	.9792337	.9853541	.9927969	.9941361	.9925023
Age 50-54	.981303	.9885153	.9920653	.9940799	.9931905
Age 55-59	.9788354	.9882949	.9909262	.9929501	.9923261
Age 60-64	.9738262	.9849271	.9893627	.9909263	.9924666
Age 65-69	.970838	.9818472	.9892153	.9873017	.9923365
Age 70-74	.9733883	.9847789	.9912982	.9903822	.987693
Age 75-79	.9794852	.9897237	.9935086	.9934934	.9880677
Age 80-84	.9847105	.992541	.9928147	.9941593	.9881803
Age 85+	.9839216	.9867062	.987198	.9859362	.9795654